

Joint Music-Making and Temporal Coordination in Joint Action

Thomas Wolf

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 Supervisor: Natalie Sebanz

Budapest, Hungary

2019

Declaration of Authorship

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or which have been accepted for the award of any other degree or diploma at Central European University or any other educational institution, except where due acknowledgment is made in the form of bibliographical reference.

The present thesis includes work that appears in the following articles:

Wolf, T., Sebanz, N., & Knoblich, G. (2018). Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing?. *Cognition*, 178, 103-108.

Wolf, T., Sebanz, N., & Knoblich, G. (in prep.). Adaptation to Extraordinary Coordination Patterns in Joint and Individual Actions

Wolf, T., Vesper, C., Sebanz, N., Keller P. E., & Knoblich, G. (under review). The Bug Within: Why Do People Rush during Rhythmic Interactions?

Abstract

When humans engage in joint action, they bring about changes in the environment together by coordinating in space and time. Even such simple joint actions as shaking hands require sophisticated temporal coordination. This is even more obvious for complex forms of joint action, such as joint music-making. Indeed, joint music-making is a domain that demands an exquisite degree of precision in temporal coordination. It also poses additional challenges, which arise from the need to predict and adapt to each other's actions – often on different instruments and with different skill levels – while performing extraordinarily complex patterns and adhering to context-dependent aesthetic criteria and performance targets. In this thesis, I present three studies in which I investigated how expert and novice musicians deal with three such challenges. The findings illuminate the basic mechanisms underpinning humans' remarkable ability to coordinate the timing of their actions both in musical and in non-musical joint actions.

In the first study, I investigated expert pianists' ability to adjust their temporal predictions to the systematic, but suboptimal, timing deviations of novice pianists. In a music coordination task, expert pianists had access to different pieces of information about their co-performer and the co-performer's part. The results indicate that experts use information about the novice's performance style during easier passages and information about the novice's part (i.e. the score) during passages that are difficult to perform.

In the second study, I asked participants to adapt to an unusual coordination pattern under various coordination conditions. The primary question was whether the weaker coupling between limbs in interpersonal coordination (e.g., the two hands of two different individuals) during joint performances allows for better adaptation to difficult coordination patterns than the stronger coupling between limbs in intrapersonal coordination (e.g., the two hands of one

person). The results show that while strong coupling between limbs facilitates precise coordination in simple coordination patterns, this advantage disappears in more difficult patterns.

The third study focuses on a particular performance bias, namely the tendency to gradually increase tempo during joint music-making ('rushing'). The central question was whether this bias is specific to joint performance, or whether it also occurs during solo music-making. The results indicate that rushing is indeed specific to joint performance. Various hypotheses concerning the underlying mechanisms of rushing are discussed and tested. Of these mechanisms, the findings speak in favor of a combination of human-specific period correction mechanisms, and evolutionarily ancient synchronization mechanisms found even among distantly related species of chorusing insects.

In this thesis, I treat the domain of joint music-making as a microcosm in which to study humans' remarkable ability to precisely coordinate their actions in time. The three studies focus on some of the challenges that humans face when trying to coordinate their actions in time during joint music-making, but the findings also have broader significance: they provide us with new insights into the general mechanisms of temporal coordination in humans, and offer new starting points and constraints for research on joint action.

Acknowledgements

There are many people without whom this thesis would not have been possible and there are many people without whom writing this thesis would have been much harder and definitely much less enjoyable.

First of all, I want to thank my supervisors, Günther and Natalie, for their guidance and support. I feel incredibly lucky and honored to have been working with you. Thank you also for inspiring me to become the best researcher I can be and for providing me with countless opportunities to grow.

I also want to thank John for his advice, encouragement and mentoring and for being there when I need a shoulder to lean on. You are so much more than an academic role model and you constantly open up and introduce me to new and wonderful worlds, such as the world of academia a long time ago.

A very special thanks goes to my parents. Thank you for your wisdom, skills and creativity, which have helped me to overcome countless obstacles throughout my academic career, not the least with the design and building of experimental setups. Thank you also for having continued to financially support me for such a long time.

As this thesis was to a large part inspired by my experiences of joint music-making, I would like to thank my sisters, who introduced me to music in the first place. I also want to thank Julia, Marika and Marlene, with whom I have shared so many musical experiences. Thank you, John, Cordula, Tanja, Gina, Francesca, and Denis for sweetening my PhD with joint music-making and thank you Barbara, Lisi and Zsuzsi for teaching me.

Thank you Dávid, for helping me tremendously with all the testing. Thank you, Andi, Eszter and Réka, for your crucial, administrative assistance and your incredible patience.

Thanks to all my colleagues at the Cognitive Science Department at CEU. Thank you, Helena, Martin and Mia for making me feel like I have a second home in the Jungle Office.

Thank you, Arianna, Anna, Barbu, Cordula, Laura, Oana, Pavel and Vero for your invaluable insights. Thank you, Francesca, Gina, Helena, Laura, Merryn, Nazlı and Simily for always being there for me. Thanks to the faculty and post-docs for fostering such a great scientific community. Thank you everyone, for the dancing, the board game evenings, the fredagsbars, the movie nights, the growling, the yoga sessions and the reading club.

Table of Contents

1	Introduction.....	1
1.1	Temporal Coordination in Joint Actions	1
1.2	Mechanisms for Temporal Coordination.....	4
1.2.1	Sensorimotor Synchronization.....	5
1.2.2	Error Correction in Information-Processing Accounts	6
1.2.3	Entrainment and Coupling Strength.....	7
1.2.4	Models for Interpersonal Coordination.....	8
1.3	Flexibility in Interpersonal Temporal Coordination.....	9
1.3.1	Challenge 1: Systematic Temporal Deviations	10
1.3.2	Challenge 2: Extraordinary Coordination Patterns	11
1.3.3	Challenge 3: Flexibility of Control	12
1.4	Joint Music-Making as Temporal Coordination.....	13
1.4.1	Music-Making is Universal and the Majority of it is Joint	14
1.4.2	Diversity of Practices and Functions of Joint Music-Making	15
1.4.3	Evidence for Flexibility in Joint Music-Making.....	16
1.4.4	Additional Advantages of Music-Making Paradigms.....	17
1.5	Research Questions: Flexible Temporal Coordination.....	18
1.5.1	From Expert-Expert to Expert-Novice Interactions (Study 1).....	18
1.5.2	Adaptation to Extraordinary Coordination Patterns (Study 2)	20
1.5.3	Joint Rushing: The Bug Within (Study 3)	21

2 Study 1: Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing? 23

2.1	Introduction	23
2.2	Methods	26
2.2.1	Participants.....	26
2.2.2	Apparatus and Material.....	26
2.2.3	Procedure and Design	27
2.3	Results	29
2.3.1	Mean Absolute Asynchronies	29
2.3.2	Variability of Asynchrony (SD).....	31
2.3.3	Cross-Correlation Analysis	32
2.3.4	Random Pairing Analysis	32
2.3.5	Autocorrelation Analysis	33
2.4	Discussion.....	33
2.5	Follow-Up Study: What is the Contribution of Pitch Information?	38
2.5.1	Introduction.....	38
2.5.2	Methods.....	39
2.5.3	Results.....	40
2.5.4	Discussion	42

3 Study 2: Adaptation to Extraordinary Coordination Patterns in Joint and Individual Actions 44

3.1	Introduction	44
3.1.1	Previous literature	46

3.1.2	Current Study	49
3.2	Experiment 1 - Methods	50
3.2.1	Participants.....	50
3.2.2	Apparatus and Material.....	51
3.2.3	Procedure and Design	51
3.2.4	Analysis.....	53
3.3	Experiment 1 - Results	55
3.3.1	Accuracy	55
3.3.2	Asynchrony in Baseline Trials	55
3.3.3	Asynchrony in Delay Trials	55
3.3.4	Variability in Baseline Trials	59
3.3.5	Variability in Delay Trials	59
3.3.6	Unsigned Asynchrony and its Variability Combined	60
3.4	Experiment 1b – Methods.....	61
3.4.1	Participants.....	61
3.4.2	Procedure and Design	62
3.5	Experiment 1b - Results	62
3.5.1	Accuracy	62
3.5.2	Asynchrony in Baseline Trials	63
3.5.3	Asynchrony in Delay Trials	63
3.5.4	Variability in Baseline Trials	64
3.5.5	Variability in Delay Trials	65
3.5.6	Unsigned Asynchrony and its Variability Combined	67
3.6	Experiment 2 – Methods.....	67
3.6.1	Participants.....	67
3.6.2	Procedure and Design	67
3.7	Experiment 2 – Results	68

3.7.1	Accuracy	68
3.7.2	Asynchrony in Baseline Trials	68
3.7.3	Asynchrony in Delay Trials	69
3.7.4	Variability in Baseline Trials	70
3.7.5	Variability in Delay Trials	70
3.7.6	Unsigned Asynchrony and its Variability Combined	72
3.8	Cross-Experiment Comparison.....	73
3.8.1	Adaptation performance.....	73
3.8.2	Questionnaire Data.....	74
3.9	Discussion.....	75

4 Study 3: The Bug Within: Why Do People Rush during Rhythmic Interactions? 81

4.1	Introduction	81
4.1.1	Previous research	82
4.1.2	Present study	83
4.2	Experiment 1.....	85
4.2.1	Methods.....	85
4.2.2	Analysis.....	86
4.2.3	Results.....	87
4.3	Experiment 2.....	89
4.3.1	Methods.....	89
4.3.2	Results.....	89
4.4	Experiment 3.....	91
4.4.1	Methods.....	91

4.4.2	Results.....	92
4.5	General Discussion	94
4.6	Methods	97
5	General Discussion.....	100
5.1	Joint Action Coordination in Expert-Novice Pairs.....	100
5.1.1	Future Research	102
5.2	Adaptation to Unstable Phase Relations.....	103
5.2.1	Future Research	105
5.3	Why do People Rush during Rhythmic Interactions?	106
5.3.1	Future Research	107
5.4	Joint Music-Making and Temporal Coordination in Joint Action.....	107
6	References.....	109

1 Introduction

1.1 Temporal Coordination in Joint Actions

Humans engage in various forms of joint actions, actions in which two or more individuals coordinate in time and space to bring about a change in the environment (Sebanz, Bekkering & Knoblich, 2006). Standard examples given for joint actions include shaking hands, dancing tango, lifting objects together, passing objects, throwing and catching objects, walking hand-in-hand, improvising in a jazz band or performing a symphony (Knoblich, Butterfill & Sebanz, 2011; McEllin, Knoblich & Sebanz, 2018; Vesper, Schmitz, Safra, Sebanz & Knoblich, 2016). Whereas being coordinated in time is a defining feature of joint actions, the requirements on how precisely joint action partners have to be coordinated in time vary vastly between different kinds of joint action. In this thesis I will focus on interactions that require temporal coordination to take place with the accuracy of a few hundred milliseconds or less (Keller, 2008; Sebanz & Knoblich, 2009). Jointly spreading a blanket evenly over a bed, for example, involves lifting and lowering the blanket in tight coordination (Sebanz & Knoblich, 2009). Precision on such a timescale can be very challenging even in *intrapersonal coordination*, when one person coordinates the movements of her limbs, such as when juggling balls using both hands (Zago et al., 2017). In intrapersonal coordination both hands' and arms' movements rely on the same internal processes. Joint action partners in *interpersonal coordination*, however, cannot rely on shared internal processes. Hence, our ability to achieve such precision even in interpersonal coordination, i.e. when coordinating with others with whom we do not share internal processes (Knoblich & Jordan, 2003) is truly remarkable. Empirical studies on rhythmic joint actions demonstrate that pairs regularly achieve temporal coordination at a precision close to a hundredth of a second (Keller, Novembre & Hove, 2014).

Much of the empirical psychological research on intrapersonal and interpersonal coordination has focused on studying the production of isochronous intervals (Repp, 2005; Repp & Su, 2013). However, humans exhibit an astounding flexibility regarding temporal coordination, both in terms of complex rhythms and in terms of complex actions being performed. Especially in joint music-making, humans have to overcome several coordination challenges. As Repp (2005) notes: “Ensemble playing requires not only in-phase synchronization with other players, but also precise alternation (anti-phase coordination) and other, more complex forms of rhythmic coordination”.

Complex forms of rhythmic coordination include, among others, polyrhythms and micro-rhythms. Polyrhythms are non-harmonically related beat subdivisions, such that for example one person plays two isochronously spaced notes, dividing the beat into two, while another person plays three isochronously spaced notes, dividing the beat into three (Summers, Rosenbaum, Burns & Ford, 1993). György Ligeti, for example, is known for especially challenging combinations. Whereas the 2:3 polyrhythm example above combines two rhythms, Ligeti simultaneously superimposed even more rhythms, as for example in a 4:5:7 polyrhythm in his Piano Etude No. 6 *Automne à Varsovie*, or even a six-stream polyrhythm in his piano concerto *Entrelacs* with the ratios 4:5:6:7:9:11 (Taylor, 2003). In some of his pieces, for example in his violin concertos I and III, Ligeti distributes these polyrhythms across multiple musicians of an ensemble, thereby bringing them into the realm of interpersonal coordination (Taylor, 2012).

Micro-rhythms are characterized by intended, minute shifts in time to create a certain feel. For example, the conventional micro-rhythmic pattern of a Viennese waltz realized in an ensemble requires bass players to act a little bit early so that their instruments emit their sounds on time, while the second violins are not supposed to play too early on the first, but on the second beat and too late on the third (Fuchs, 2019; Gabler, 1998; Gludovatz, 2009). These

micro-rhythmic deviations are common in many music styles around the world. They have been studied for example in Jazz, Funk, Cuban, Samba and have been investigated as the main contributor to the groove of music performances (see Davies, Madison, Silva & Gouyon, 2013). Malcolm Braff, a pianist, developed a systematic theory of micro-rhythms, along with terminology and ways to graphically represent micro-rhythms in scores and as geometric shapes. This allows his ensemble to master and to refer to different *phrasings*, i.e. micro-rhythmic patterns, and to perform continuous *morphings* between them (Braff, 2015a; Braff, 2015b).

Whereas polyrhythms and micro-rhythms can sometimes refer back to a common pulse or *pulsation* (Arom, 1991), some musical performances demand systematic departures from a common pulse. The compositional technique of *phasing*, for example in some of Steve Reich's compositions requires performers to speed up one of two rhythmic or tonal patterns by a small amount in order to gradually shift it through phase space until it realigns with a pattern that remains stable (Schwarz, 1980). Each time a performer speeds up a pattern she departs from a common pulse. Yet other musical interactions are defined by requiring the lack of temporal coordination, as for example between two groups from different communities in the Afro-Brazilian Congado ritual, in which the ability to resist falling into coordination is a sign of a group's spiritual power (Lucas, Clayton & Leante, 2011).

It is not only remarkable *what* humans are capable of doing in terms of temporal coordination, but it is also interesting to consider *why* humans have such exquisite coordination skills. Joint music-making, which, as we will see later, is a particularly ubiquitous form of interpersonal temporal coordination, has been argued to be an effective means for large scale social bonding, much more efficient than, for example, mutual grooming (Launay, Tarr & Dunbar, 2016). Interpersonal temporal coordination can therefore function as “social glue” (Honing, ten Cate, Peretz & Trehub, 2015) and has been argued to enhance cooperation and

pro-sociality (see Green et al., under review; Rennung & Göritz, 2016). Hence, investigating temporal coordination also has the potential to shed light on human sociality (Cross, 2008). Furthermore, observing tight temporal coordination and synchronization, in particular, are linked to aesthetic appreciation for example during the performance of dance choreographies (Vicary, Sperling, von Zimmermann, Richardson & Orgs, 2017) and probably also during synchronized sports, in which endless hours of repetition and related injuries are tolerated to achieve tight temporal coordination (Mountjoy, 2009).

Whereas this thesis might be motivated by *what* humans achieve in terms of temporal coordination and *why* they do it, its aim is to contribute to the investigation of *how* humans are able to achieve precise temporal coordination with the remarkable flexibility to adapt to a variety of coordination requirements. In this thesis, I will focus on interpersonal coordination unfolding on timescales for which reacting to another's behavior would be too slow to allow for sufficient accurate coordination. Hence a common necessary factor at this timescale is the ability to anticipate (Pecenka & Keller, 2011; Repp, 2005). As temporal coordination is a key feature of joint music-making (Keller, et al., 2014; Repp, 2005), inspiration for experimental paradigms will be drawn from the domain of music. In Chapter 1, I will first review existing mechanistic explanations for temporal coordination, then specify three challenges for temporal coordination linked to the flexibility we observe in coordinating humans, and rephrase finally derive research aims from these challenges, leading to an outlook of research questions and experimental paradigms presented in the main chapters of this thesis.

1.2 Mechanisms for Temporal Coordination

As a starting point, I will consider mechanistic explanations of sensorimotor synchronization, which originate from mathematical formulations of how humans synchronize to regular external events, such as the sounds of a metronome. There certainly are several

differences between sensorimotor synchronization to a regular metronome and the flexibility humans exhibit in joint music-making. One of these differences is the bi-directional flow of information between interaction partners. However, over the last two decades researchers interested in sensorimotor synchronization have tried to widen the scope of their models to bring them closer to cases of joint music-making (Repp & Su, 2013; Repp, 2005). Some of the implemented extensions are relevant for the aims of this thesis and will be discussed. To facilitate discrimination between different kinds of temporal coordination I will refer to cases of coordination with external non-responsive events such as metronome clicks as *extrapersonal coordination*, as opposed to *intrapersonal coordination*, where one person coordinates the movements of two or more of her own limbs, and *interpersonal coordination*, where one or more individuals coordinate their movements.

1.2.1 Sensorimotor Synchronization

When reaction times are too high to allow for sufficient coordination with external events, these events have to be anticipated (Pecenka & Keller, 2011; Repp, 2005). In the simplest case one person coordinates with external events which are produced by a device, such as a metronome and follow a steady, isochronous rhythm, where all intervals between metronome beats have the same duration. During such extrapersonal coordination our predictions about upcoming events can be based on preceding events and their regular timing. For such behavior sensorimotor synchronization research developed mathematical descriptions. The focus of this literature has traditionally been on extrapersonal synchronization. Repp's review (2005) of the sensorimotor synchronization literature for example, focuses almost exclusively on the synchronization of finger-tapping to external cues, with a limited set of coordination requirements (twice or half the period of a metronome and cases of anti-phase coordination, see Figure 1 in Repp, 2005).

Models of sensorimotor synchronization tend to fall into one of two groups, dynamical accounts or information-processing accounts (Pressing, 1999). Whereas the former group relies on non-linearly coupled oscillator equations (Haken, Kelso & Bunz, 1985; Schmidt & Turvey, 1994) the latter group works under the assumptions of clocks and utilizes linear equations (van der Steen, van Vugt, Keller & Altenmüller, 2014; Vorberg & Schulze, 2002; Vorberg & Wing, 1996). Additionally, another noteworthy distinction between the two groups is that dynamical accounts often focus on investigating and explaining phenomena of break-down, instability, and transitions (Kelso et al., 1986), whereas information-processing accounts traditionally focus on how stable coordination is maintained (Vorberg & Wing, 1996). Despite these differences, the two resulting frameworks are argued to be compatible, with the linear equations of information-processing accounts being described as a mathematically more tractable approximation of non-linear models, at least in the vicinity of stable states (Repp, 2005; Pressing, 1999). As Pressing (1999) points out, there is a close relationship between the concepts of clocks, which are a combination of an oscillator and a read-out mechanism, and oscillators. Following this, each model can be reformulated in terms of the other (see Loehr, Large & Palmer, 2011; Pressing, 1999). Pressing (1999) notes further that both kinds of models have “essential and complementary roles to play in behavioral description”. In the following I will summarize mechanisms and studies relevant for the purpose of this thesis.

1.2.2 Error Correction in Information-Processing Accounts

Independent of whether we assume sensorimotor synchronization to be based on oscillators or clocks, setting one’s internal frequency/tempo to match that of regular external events is not enough to maintain synchronization. Even when we start to tap along in perfect synchrony and the tempo is matched exactly, inherent variability in our taps (Wing & Kristofferson, 1973a; b) would lead to slight asynchronies between our taps and the metronome

clicks which, without correction mechanisms, would accumulate tap after tap. Hence, to stay in time with the metronome these slight asynchronies need to be corrected before they add up. Hence continuous *error correction* mechanisms have been proposed (Repp, 2005).

In information-processing accounts two correction mechanisms have been spelled out, *phase correction* and *period correction*. When we both tap at the same tempo and aim for synchronization, but our taps do not line up in time, we need to correct our relative phase. Phase correction mechanisms correct for asynchronies between the two taps without changing the tempo of internal timekeepers. To allow continuous synchronization even when one of us changes the tempo we need to correct the tempo setting of our internal timekeeper. This is done by so called period correction mechanisms, which correct for asynchronies by changing the tempo setting of the internal timekeeper (Semjen, Schulze & Vorberg, 2000). Interestingly, it seems that these correction mechanisms can lead to a continuous over-correction in interpersonal coordination. In one of the first tapping studies addressing interpersonal coordination condition, Konvalinka and colleagues (Konvalinka, Vuust, Roepstorff & Frith, 2010) found that participants responded to the deviations of each other's previous inter-tap-interval. This led to the oscillatory correction pattern of *hyper-following*, where each participant shortened their inter-tap-interval when the other's last interval was shorter and vice versa.

1.2.3 Entrainment and Coupling Strength

Whereas dynamical accounts avoid the term error correction, their concepts of entrainment and coupling strength fulfill a similar function. The stronger two oscillators are coupled the larger the pull towards in-phase coordination. Hence, parameters which represent coupling strength correspond to the parameters in information-processing accounts that specify the degrees to which error correction takes place (Pressing 1999; Repp, 2005). Whereas studies

situated in the framework of information-processing often focus on *planned coordination*, studies on entrainment have highlighted the effects of *emergent coordination* (Knoblich et al., 2011). Richardson and colleagues (Richardson, Marsh, Isenhower, Goodman & Schmidt, 2007) asked participants in rocking chairs to rock in their preferred frequency under different information coupling conditions, such as looking at each other or looking away from each other. The results show that under information coupling, i.e. looking at each other, participants were more likely to fall into synchronization with each other than without coupling, i.e. looking away from each other. This and similar studies have provided evidence that entrainment supports interpersonal coordination and in particular emergent, i.e. unintended interpersonal coordination (Schmidt & Richardson, 2008). However, entrainment's tendency to push coordination towards only a few specific, stable patterns – in-phase and anti-phase coordination – will not always have facilitating effects for temporal coordination. This should be the case when the desired coordination pattern deviates from these patterns which are stable in systems of coupled oscillators. Furthermore, dynamical accounts favor explanations on lower levels and in principle assume the same mechanisms for interacting metronomes and for interacting people (for a review of reformulations of social-psychological variables as control parameters see Schmidt & Richardson, 2008).

1.2.4 Models for Interpersonal Coordination

Information-processing accounts of sensorimotor synchronization have been developed further to specifically take into account cases of interpersonal coordination. Van der Steen and Keller (2013) proposed ADAM, an adaptation and anticipation model to account for “precise yet flexible” interpersonal coordination between humans. ADAM extends Vorberg and Wing's model of sensorimotor coordination to incorporate mechanisms of anticipation (see also van der Steen, Jacoby, Fairhurst and Keller (2015) and consists of two modules. Phase and period

correction mechanisms as described by Vorberg and Wing (1996) are part of the *adaptation module*. The additional *anticipation module* includes linear extrapolation from a couple of the most recent temporal intervals another individual has produced. Based on this extrapolation the next external interval is predicted and compared to one's own planned interval. Discrepancies between the two can then be corrected for to a certain, pre-specified degree. The anticipation module further allows to weight how much ADAM relies on anticipation versus simple tracking via a prediction/tracking parameter. Various versions of ADAM were compared with behavioral data collected in a finger tapping task to tempo changing sequences. Versions of ADAM that included both the adaptation and the anticipation module fit the data significantly better than an adaptation-only version. Hence, ADAM seems to be able to model humans' precise yet flexible sensorimotor synchronization abilities. However, ADAM's flexibility is limited by only taking into account the general linear trajectory of the last couple of intervals. As I will argue in the next section, the wide range of temporal coordination challenges humans face go beyond these specifications.

1.3 Flexibility in Interpersonal Temporal Coordination

Whereas mechanisms of sensorimotor synchronization might be of “fundamental importance” (Repp, 2005) for a wide range of interactions requiring temporal coordination, including joint music-making, they primarily take into account coordination on a local level. Error correction mechanisms, for example, are based on one preceding event and even the formalized parts of ADAM's Anticipation Module are only concerned with linear extrapolation from the last couple of synchronization events. Complex interpersonal coordination such as joint music-making however, is often governed by, and requires tracking, regularities on various levels (see Jones, 2009). Furthermore, sensorimotor synchronization research has focused on extrapersonal coordination and largely on one of many possible coordination

regimes, namely synchronization. Some studies of sensorimotor synchronization have investigated interpersonal coordination (Konvalinka et al., 2010) and some have included other coordination regimes in their investigation, as for example anti-phase coordination (Schmidt, Carello & Turvey, 1990) and coordinating at double or half the tempo (Loehr et al., 2011). Yet there is still a gap between models of sensorimotor synchronization and the wide variety of examples of temporal coordination I provided in the introduction. This leads to several open research questions. In the following, I will spell out three challenges that humans face in terms of temporal coordination during a range of rhythmic interactions. These challenges require flexibility that goes beyond what basic mechanisms of sensorimotor synchronization can provide. In the last section of this chapter, I will use these challenges to define open research questions that are addressed in this thesis.

1.3.1 Challenge 1: Systematic Temporal Deviations

Synchronizing movements with an external pacing signal which is completely determined by a set tempo, like the isochronous clicks of a metronome, is the paradigmatic case of sensorimotor synchronization (Repp, 2005). In joint actions however, the “external pacing signal” is produced by another human. This implies multiple sources of temporal variability and, therefore, deviations from a steady rhythm. Some of these deviations are indeed the product of noise, for example noise which originates from the central or from the peripheral motor system (see for example the two-level timing model of Wing & Kristofferson, 1973a; b). Correction mechanisms have been proposed to counteract deviations resulting from noise by correcting based on previous taps, i.e. they produce local error corrections.

However, humans sometimes also deviate from an isochronous rhythm in systematic ways. Systematic deviations could for example be a result of the underlying hierarchical structure of a musical piece (Palmer, 1996b), of a performer’s expressive intentions (Palmer,

1989; Palmer, 1996a), or they could reflect systematic variability due to performance constraints exhibited by an interaction partner (Aoki, Furuya & Kinoshita, 2005). Accounting for such systematic deviations is for example required when you synchronize your steps with the systematically irregular steps of a limping joint action partner, with whom you are carrying furniture. Another example would be adapting to systematic timing errors of a piano novice due to insufficient skills for more difficult passages of a piano duet. The challenge of coordinating with systematic temporal deviations leads to the question of whether and how we are able to adapt to and predict systematic timing deviations in order to improve temporal coordination.

1.3.2 Challenge 2: Extraordinary Coordination Patterns

Synchronization, which calls for the temporal alignment of behavior, i.e. matched period and phase, is only one of a multitude of possible coordination patterns. I gave examples for some of these patterns in the introduction of this chapter, such as the polyrhythms in Ligeti's piano etudes, or micro-rhythmic deviations as they are common in orchestral Viennese waltz performances. Whereas some coordination patterns that differ from 1:1 in-phase synchronization have been studied under the framework of dynamical accounts (mainly anti-phase patterns, see Schmidt & Richardson, 2008), these studies were primarily concerned with the instability and the break-down of coordination as a function of tempo increase rather than with the means necessary to maintain these coordination patterns successfully. In fact, dynamical approaches predict instability for all coordination patterns except for in-phase and anti-phase which have been called *canonical steady states* (Schmidt & Richardson, 2008).

The challenge that humans face is then to maintain *extraordinary coordination patterns*, where *extraordinary* refers to any pattern other than the canonical in-phase and anti-phase. Imagine for example three people who use hammers to drive a large pole into the

ground. If they alternate hitting the pole at regular intervals to prevent collisions, they have to maintain a relative phase of $360 / 3$ degrees, which falls between in-phase (0 degree) and anti-phase (180 degree). Additional difficulties can arise in cases where a phase shift of two action streams is required due to different latencies inherent to devices or tools we use. Imagine for example that one of three hammers is much larger than the other two hammers, such that it takes longer from action initiation until the hammer hits the pole. In order to hit the pole in regular timing intervals the person with the larger hammer has to start her action earlier than the others would. This is analogue to what double bass players in ensembles that play Viennese Waltzes are required to do. As noted above, in such cases it is custom for double bass players to play a bit too early, so that the sound reaches a considerable amount of its full volume at the right time. As response time is inversely correlated with frequency, a double bass player has to compensate more than the violinists in the ensemble (Guettler & Hallam, 2002). Hence, joint actions regularly require temporal coordination that goes beyond canonical coordination patterns. An important question originating from this challenge is whether and how temporal coordination can be achieved and maintained despite the inherent instability of such extraordinary patterns?

1.3.3 Challenge 3: Flexibility of Control

In the challenges I have listed so far, flexibility is desired to adapt to systematic deviations or extraordinary coordination patterns. However, whereas flexibility is required in some interactions, in others it might work against coordination or reduce performance quality. In the introduction to this chapter I mentioned the compositional technique of phasing in Steve Reich's music, where the same pattern is played by two hands (either intra- or interpersonally). Initially these patterns are played in in-phase coordination, in other words perfectly in time with each other. However, as a next step one of these patterns is shifted in time by slightly

increasing the tempo in one of the two hands. Hence, successful phasing requires the flexibility of one hand to break free of the synchronization, but it also requires the flexibility of control in the other hand to not adapt, to not go along with the increased tempo, but rather to remain unchanged. When phasing is achieved, it is a successful instance of controlled flexibility. However, to investigate topics related to flexibility of control in temporal coordination I will instead focus on instances of failure; instances in which control is not flexible enough to prevent counter-productive and unintended adaptation. Interpersonal coordination among mutually adaptable agents could lead to a spiral of uncontrolled adaptations that do not emerge to the same extent from extra- and intrapersonal coordination. Such undesired mutual adaptations specific to interpersonal coordination therefore require humans to systematically exert control during interpersonal coordination. The questions therefore arise of whether and how such *over-adaptations* are specific to interpersonal coordination and how their effects could be mitigated.

1.4 Joint Music-Making as Temporal Coordination

Why is joint music-making relevant for the investigation of temporal coordination? Is joint music-making not a rather small niche of human behavior? Is it not a very specialized and narrow set of human activities? How are joint music-making and flexibility in temporal coordination related and finally, does it make sense to employ experimental tasks inspired by joint music-making to study temporal coordination? As I will argue in this section, joint music-making comprises a very diverse set of practices and offers several theoretical and practical advantages for the study of temporal coordination and timing mechanisms that are of general importance for joint action research. Following the credo that “joint action is by no means just a specialty of experts in domains like sports or art” (Sebanz et al., 2006), I will argue that joint music-making is by no means just a specialty of highly skilled musicians and that the

mechanisms involved in joint music-making are in fact paradigmatic for joint action. This is in line with D'Ausilio et al.'s argument (D'Ausilio, Novembre, Fadiga & Keller, 2015) that using music as a model for joint action can help to inspire 'novel perspectives and original questions that touch on the core of human social cognition'.

1.4.1 Music-Making is Universal and the Majority of it is Joint

Music-making is assumed to be a human universal, something to be found in all human cultures (Brown, 1991; Savage, Brown, Sakai & Currie, 2015). When two or more people make music together, they perform a joint action, as joint music-making requires coordinating to bring about a change in the environment. Researchers seem to agree that music-making is inherently social (Wöllner & Keller, 2017). An article by Savage et al. (2015) can provide us with an intuition of how frequent music-making is indeed *joint* music-making. Savage and colleagues examined a set of 304 music recordings that were selected to represent a large variety of music cultures around the world. As part of their search for universal features of music, they coded the dataset for 32 musical features. Group performance (absent, present) was one of these features. The part of the dataset classified as Middle East and European indicated the lowest percentages of joint music-making (around 66% and 58 % respectively). More than 90 % of the recordings labeled as Southeast Asian (100%) and South Asian (94%) were marked as group performances. Calculated for the whole sample of recordings, the results indicate that around 75 % were the product of joint music-making. Thus, it seems that more often than not music-making is a joint endeavor, and in some cultures, exclusively so. Joint music-making is, therefore, a joint activity that humans cross-culturally engage in, which opens up the possibility for experimental designs that have cultural credibility. Even when music is produced by individuals it is usually performed for others to hear it (Wöllner & Keller, 2017).

1.4.2 Diversity of Practices and Functions of Joint Music-Making

Joint music-making practices comprise a wide range of group activities, from different make-ups of ensembles to different instruments that require different actions, from simple rhythmic games played by children to performances that require a lifetime of training, from participatory practices to practices with a strict distinction between performers and audience, and from private settings to public displays. Some practices involve multiple musicians gathered around one instrument weaving together musical patterns that are not played by any individual musician alone (Kubik, 1960; Kubik 1964). Others involve musical competitions between groups of drummers that try to uphold their group's identity by making an effort not to fall into the tempo of another simultaneously performing group (Lucas, et al., 2011; Lucas, 2002). It has been argued that the variety of ensemble types, sizes, their functions and contexts to reflect the variety of societies (Lomax & Berkowitz, 1972). Interestingly, the diversity of music-making within cultures is even larger than the variability of practices between cultures (Rzeszutek, Savage & Brown, 2011).

The cases of joint music-making that are frequently cited as examples of joint actions are usually performances of expert musicians and musical tasks used in the joint action literature often rely on expertise. Hence, it is important to point out that joint music-making is far from restricted to the interactions of professional experts. It also encompasses infant-caretaker interactions (Trehub & Trainor, 1998; Trehub, Unyk & Trainor, 1993), rhythmic childrens' games (Brodsky & Sulkin, 2011; Merrill-Mirsky, 1986), recreational music-making (Finnegan, 2007), and many others. The fact that joint music-making does not necessarily rely on expertise combined with the development of new technologies that provide various degrees of scaffolding for novice musicians led to an increase in studies that invited novice musicians to participate in musical tasks. Novembre and colleagues, for instance, developed a custom-made e-music box that participants can operate by turning a handle in order to investigate

novices' ability to coordinate their timing (Novembre, Varlet, Muawiyath, Stevens & Keller, 2015).

The term joint music-making, which refers to people making music together, already includes a rich array of interesting interactions. It should also be noted here, however, that an even broader music related category of joint actions might also be a worthwhile area of investigation, namely *music-related joint actions*. *Music-related joint action* refers to any joint action that occurs in a context in which music is present. It could, for example, be the case that there is only one person making music, but others clapping along, dancing or facilitating the performance in a technical or organizational manner. Yet other music-related joint actions might be completely void of co-present music-making, and involve recorded music, like hip-hop dance classes or silent disco events (Bamford, Burger & Toiviainen, 2016; Tarr, Launay & Dunbar, 2016).

1.4.3 Evidence for Flexibility in Joint Music-Making

The universality, ubiquity and diversity of joint music-making contribute to the cross-cultural, ecological validity of using paradigms that make use of joint music-making for the study of temporal coordination. However, more importantly, for investigating flexibility in temporal coordination among humans, the domain of joint music-making has two desirable characteristics. First, among the many types of joint music-making, there are interactions that do indeed require flexibility as spelled out in the three challenges above, some of which I have listed in the introduction to this chapter. Second, joint music-making enables one to measure quantitatively whether people are indeed able to meet the demands of coordinating in a precise and flexible manner.

With respect to systematic deviations (challenge 1), coordination in a music ensemble can for example require individual players to adapt to a whole range of systematic deviations,

which can be based on, among others, expressive intentions (Palmer, 1989; Palmer, 1996) or micro-timing conventions (Naveda, Gouyon, Guedes & Leman, 2011). However, interestingly, it has been suggested that expressive timing is dampened during ensemble performances in comparison to individual practice (Rink, 2002). With respect to extraordinary coordination patterns (challenge 2), one can identify music pieces that require such patterns and investigate whether people are able to learn to perform them. When a piece requires different co-occurring pulses for example, they are sometimes split among musicians in an ensemble (Kubik, 1960). With regards to controlling flexibility (challenge 3) one can look, for example, at instances of unintended tempo changes, whether they occur and how to prevent them (Thomson, Murphy & Lukeman, 2018; Okano, Shinya & Kudo, 2017).

1.4.4 Additional Advantages of Music-Making Paradigms

In order to investigate how individuals coordinate their actions to perform joint actions a range of tasks and paradigms have been developed to study the mechanisms underlying interpersonal coordination, including, but not limited to, passing objects (Constable et al., 2016; Gonzalez, Studenka, Glazebrook & Lyons, 2011), clinking glasses (Kourtis, Knoblich, Woźniak & Sebanz, 2014) or jumping together (Vesper, Van der Wel, Knoblich & Sebanz, 2012). Joint music-making has become more and more popular (Repp & Su, 2013), as a research domain for studying general mechanisms of joint action.

D'Ausilio, Novembre, Fadiga and Keller (2015) provide a list of advantages of studying joint action in the domain of music. On this list there are three additional advantages that have not been discussed above and which should be mentioned here. Irrespective of skill level, joint music-making can be inherently enjoyable and motivating. Musical scores provide formal descriptions of an interaction and can thus be used as a reference to which actual performance

can be compared. Furthermore, musical scores enable researchers to precisely specify various properties of an interaction that they intend to address in their research.

1.5 Research Questions: Flexible Temporal Coordination

In this section I will summarize how each of the three main chapters of this thesis corresponds to one of the challenges spelled out in Section 1.3. Each challenge and the general question posed above will lead to a specific research question, each of which is aimed at furthering our understanding of whether and how humans can overcome these challenges, hence our understanding of flexibility in temporal coordination and of temporal coordination in joint action in general.

1.5.1 From Expert-Expert to Expert-Novice Interactions (Study 1)

The first challenge raised was that successful temporal coordination sometimes comes with the requirement to adapt to and to take into account systematic timing deviations of a co-actor. This led to the general question of whether and how we are able to adapt to systematic timing deviations in order to improve temporal coordination. Systematic deviations, as I introduced them in section 1.1.3, can afford to be predicted if they are not too complex. Studies by Pecenka and Keller (2011, 2009a, 2009b) have revealed individual differences in predictive capacities across participants that were stable for several months. The importance of these capacities for interpersonal coordination was corroborated by the findings that prediction tendencies were good predictors of interpersonal coordination performance. One option for how temporal predictions are formed might come from studies that investigate how our motor system is recruited while we observe others' actions (Rizzolatti & Craighero, 2004). This overlap of perception and motor might not only facilitate action understanding, but also the prediction of action outcomes and action timing (Sebanz & Knoblich, 2009). The implications

of using our own motor system for the prediction of others' actions (Blakemore & Frith 2005; Keller 2008; Knoblich & Jordan 2003; Wilson & Knoblich 2005; Wolpert, Doya & Kawato, 2003) and how these implications relate to joint music-making will be investigated in Chapter 2.

Indeed, previous research indicates that skilled joint action partners use predictive models to achieve temporal coordination, for instance, when playing a music duet (Keller, Knoblich & Repp, 2007). These findings also show that predictions are best when the predicting and the predicted system are very similar, with the best performance for when the predicting system and the predicted system is the same (Keller et al., 2007; Knoblich & Flach, 2001). Previous research suggests that this depends on experts predicting the actions of experts. This has been tested in a musical paradigm, namely with expert pianists playing along recordings of expert pianists (Keller et al., 2007).

However, not everyone playing a piano is a professional musician. Probing the hitherto existing findings reveals that it is unclear how predictions would play out for interacting musicians of different skill levels. A common case, for example, would be teacher-student interactions. In Chapter 2, I will therefore address the following questions: How do joint action partners with different skill levels achieve coordination? How do musical experts adjust internal models to suboptimal performance? Can experts predict the suboptimal timing of novices? What kind of information allows them to predict novices' timing? To address these questions, I invited skilled pianists to perform duets with piano novices. I varied whether, prior to performing duets, experts were familiar with how novices perform their individual parts and whether experts had access to the musical scores including the novices' part of the duet. I also compared experts' coordination performance in passages that were comparably easy for the novice with their performance in passages that were more difficult for the novice.

1.5.2 Adaptation to Extraordinary Coordination Patterns (Study 2)

The second challenge that will be addressed by this dissertation is to better understand whether and how temporal coordination can be achieved and maintained in extraordinary coordination patterns despite inherent instability. Research on interlimb coordination has shown that some coordination patterns are more stable than others, and function as attractors in the space of possible phase relations between different rhythmic movements. The canonical coordination patterns, i.e., the two most stable phase relations, are in-phase (0 degree) and anti-phase (180 degrees) (Kay, Saltzman, Kelso & Schöner, 1987; Kelso, 1984). The relative stability of these canonical patterns has also been confirmed for interpersonal coordination (Schmidt et al., 1990). Yet, musicians manage to perform other coordination patterns in intrapersonal as well as in interpersonal coordination with remarkable precision such as when producing the required micro-rhythmic deviations while playing a Viennese waltz together. In the following, we will use the term ‘extraordinary coordination patterns’ to refer to all stable rhythmic coordination patterns that are neither in-phase nor anti-phase.

In Chapter 3, I will investigate two factors that could facilitate adaptation to extraordinary coordination patterns. The first factor is the inherent coupling strength between the entities (limbs, people, objects) that perform the individual components of the coordinated pattern. More specifically, I will investigate whether adapting to extraordinary coordination patterns is easier during joint than during individual performance. This seemingly counter-intuitive hypothesis is motivated by the finding that strong coupling pulls coordination into one of the canonical patterns. The weaker coupling between two limbs of two people in comparison to the neural intrapersonal coupling of limbs (Schmidt, Bienvenu, Fitzpatrick & Amazeen, 1998) might allow for better adaptation to extraordinary coordination patterns. The second factor, structure of action effects could facilitate adaptation across different coordination settings. In interpersonal coordination joint outcome representations encompassing the effects

of one's own and others' actions may help ensembles to maintain extraordinary coordination patterns.

In order to investigate these two factors, I modified a finger tapping task to study adaptation rates to an extraordinary coordination pattern. I tested adaptation under different coordination settings with different inherent coupling strengths. Participants coordinated either two limbs in intrapersonal coordination, in interpersonal coordination or one limb with outcomes produced by a computer. The structure of action effects was modulated by introducing a melodic contour to the action effects, a measure that has been shown to improve the structuring of timed action sequences (Dyer, Stapleton & Rodger, 2017). I will argue that the questions addressed in my experiments are not just relevant for basic research on interlimb coordination and joint action, but that they are also relevant for addressing questions of cultural evolution (Scott-Phillips, Blanke & Heintz, 2018).

1.5.3 Joint Rushing: The Bug Within (Study 3)

In Chapter 4 I will address the third challenge concerning flexibility of control in interpersonal coordination. I will investigate an instance in which the need for interpersonal coordination creates a systematic bias during joint performance, which can only be overcome by exerting control over one's flexibility in temporal coordination. As an instance of such a bias, I will investigate interpersonal coordination mechanisms that lead to unintended tempo increases during joint rhythmic performances.

When people engage in rhythmic joint actions, from simple clapping games to elaborate joint music-making, they tend to increase their tempo unconsciously. This phenomenon of *joint rushing* has only recently been addressed by research, despite the rich literature on rhythmic performance in humans. I will propose that joint rushing arises from the concurrent activity of two separate mechanisms: a phase advance mechanism and a period correction

mechanism. The phase advance mechanism was first proposed in research on synchronously flashing fireflies and chorusing insects. When this mechanism is combined with the human-specific period correction mechanism, the shortened periods of individual intervals are translated into a tempo increase. In Chapter 4, I present and discuss three experiments conducted to investigate whether joint rushing can reliably be observed in a synchronization-continuation task. Furthermore, I investigated whether there is evidence for a phase advance mechanism that is tuned to react more strongly to certain signals. In chorusing insects, such a signal-discriminating mechanism is thought to help in avoiding inter-species interference. I will also control for social facilitation effects - a possible alternative hypothesis. In addition, I will investigate a more specific alternative hypothesis related to mirroring effects.

2 Study 1: Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing?

2.1 Introduction

When people perform joint actions together, they need to coordinate their actions in time (Butterfill, 2016; Keller, Novembre & Hove, 2014; Pecenka & Keller, 2011; Sebanz & Knoblich, 2009). Previous research has demonstrated that internal models enable joint action partners to predict each other's timing if both are skilled in performing the individual parts of a joint action (e.g., Kourtis, Knoblich, Woźniak & Sebanz, 2014). However, an open question is how joint action partners who differ in their individual skills achieve joint action coordination. For instance, when two musicians play a piano duet together, they need to adhere to certain tempo requirements and to minimize interpersonal asynchronies, regardless of differences in their skills. How coordination is achieved despite large differences in skill is an important question because it pertains to many joint actions performed in the context of teaching (Csibra & Gergely, 2009). The aim of the present study was to investigate whether skilled performers can ensure successful interpersonal coordination despite novices' suboptimal timing, and what kind of information helps them to achieve this.

We started from the hypothesis that interpersonal temporal coordination can be achieved if the skilled joint action partner compensates for a novice's suboptimal performance. When adapting to novices' suboptimal timing experts have to go beyond using their own internal models to predict a joint action partner's performance in real time (Noy, Dekel & Alon, 2011; Wolpert, Doya & Kawato, 2003). The reason is that, according to internal model accounts, the accuracy of predictions of a joint action partner's performance should depend on a high degree of similarity of the predicting and the predicted system (Grezes, Frith &

Passingham, 2004; Knoblich & Flach, 2001). The dimensions of similarity that can affect the accuracy of coordination range from a shared preference for a general tempo (Loehr & Palmer, 2011) to fine-grained similarities in micro-timing that characterize particular individual performances (Keller, Knoblich & Repp, 2007). Although there is some evidence that, in the context of action observation, predictions can be adjusted to reflect general differences in action capabilities (Ramenzoni, Riley, Davis, Shockley & Armstrong, 2008; Welsh, Wong & Chandrashekar, 2013), it is an open question whether temporal predictions can be adjusted to the suboptimal timing that characterizes novice performance in domains such as sports, dance, and music where such joint actions are frequent.

We studied this question in the domain of musical joint action, where continuous and accurate temporal coordination is crucial to achieve successful joint performance (Keller et al., 2014). It is well established that music experts use internal models to predict the outcomes of their own and others' actions (Haueisen & Knösche, 2001) and experts' performance parameters differ systematically from novices' performance parameters (Aoki, Furuya, Kinoshita, 2005; Loehr & Palmer, 2007).

Observation of music ensembles rehearsals (Ginsborg, Chaffin & Nicholson, 2006; Ginsborg & King, 2012) suggested two important sources of information for improving temporal coordination: knowledge of the structure of a partner's part, and knowledge of a partner's interpretation of her part, including expressive timing (Repp, 1990). The former can be acquired in the absence of a partner from musical scores. In contrast, a partner's idiosyncratic interpretation will need to be experienced first-hand.

A study by Ragert, Schroeder and Keller (2013) provided support for the importance of experiencing a partner's performance. They asked highly trained pairs of piano experts to repeatedly perform duets. Temporal coordination between the pianists improved across consecutive repetitions of the same duet, supporting the claim that increasing familiarity with

a partner's playing style improved interpersonal coordination. This finding indicates that pianists were able to adjust their predictions to capture parameters of an expert partner's timing. However, it is an open question whether experts are also able to adjust to novices' suboptimal timing that does not express a certain musical style. This may be a precondition for teaching through joint actions where experts provide a timing scaffold for novices.

But are there any regularities in novices' timing that experts could pick up on to improve their predictions? In order to address this question, we varied whether experts were familiar with novices' performances of their individual parts before performing a duet with them. We hypothesized that experts' familiarity with novices' playing would improve temporal coordination during ensuing duet performance. This is only expected to occur if experts can extract from a novice's performance idiosyncratic patterns that help them to improve their real-time predictions of the novice's performance during a joint performance. One factor that is likely to produce predictable timing variability in the novice is encountering particular motor difficulties such as having to switch the hand position on an instrument. Such difficulties are often visible from the musical score because these scores can include particular instructions for the positioning of fingers and hands. In order to find out whether experts can translate such symbolic information into real-time predictions during joint performance we used music notation from the tradition of Western classical music and varied whether experts knew in advance the musical score including the novices' parts of the duets. Knowing novices' scores is only expected to help experts to improve temporal coordination during joint performance if they can identify difficulties for the novices and translate these into accurate delays in their own performance to match delays in the novice's performance. Finally, to check that timing variability in novices originate from their performance difficulties, i.e. was suboptimal, we attempted to rule out the theoretical possibility that some of their variability originates from

expressive timing based on musical intentions. This was done by checking for autocorrelations in novices' performances which are present in experts' expressive timing.

2.2 Methods

2.2.1 Participants

Twelve expert pianists (5 women, 7 men, mean age = 24 years, SD = 3 years) participated in the experiment. All experts had at least 10 years of private piano lessons ($M = 12$ years, $SD = 3$ years). Twelve non-pianists (8 women, 4 men, mean age = 25 years, $SD = 5$ years) were randomly paired with the 12 experts. None of the novices had received piano lessons, but all of them had completed minimally 5 years of private lessons on another instrument ($M = 9$ years, $SD = 3$ years). Experts and novices were recruited through flyers distributed in music schools and university campuses in Budapest. All participants gave their informed consent and received gift vouchers as compensation. This study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

2.2.2 Apparatus and Material

Participants played on two Yamaha digital pianos. Presentation of visual information (standard scores and adapted scores for novices), auditory feedback, metronome beats, and data recording were implemented using a custom Max MSP patch. The eight duets the participants performed were based on easy piano duets from the standard repertoire. Expert and novice parts were simplified, shortened, and modified to conform to a length of four bars of four quarter notes each. The novice parts were adapted so that they could be performed with the white keys from C4 to C5. A special notation allowed the novices to sight-read simple melodies

and to reproduce them on the piano after a short training (see Figure 2.1C). For novices, each melody started in a lower hand position (thumb on C4) and required two shifts of hand position, first upwards (thumb from C4 to F4) and then downwards (thumb from F4 to C4). This means that in each melody there were two shifts to be performed by the novice. These two shifts per melody will be referred to as difficult passages, whereas the rest of the melody, where novices do not have to shift their hands, will be referred to as easy passages.

2.2.3 Procedure and Design

The main experiment consisted of eight blocks. In each block, the novice performed the duet eight times. The first four repetitions were performed by the novice alone and therefore characterize individual novices' performances of their part of the duet. In the second four repetitions the expert played along with the novice. Uni-directional auditory feedback ensured that the novice did not hear the expert and thus could not adapt to the expert's performance. Experts performed their parts of the duets individually after the main experiment for two repetitions.

In the first part of each block, where novices played alone, the expert either heard the novices' four individual performances or not (Familiarity with Novice Performance, yes or no) and the expert either had access to the sheet music for the novice's part or not (Familiarity with the Score, yes or no). In the second part of the block novices never heard the experts and were instructed to simply repeat playing their parts of the duet four more times. Experts played their part of the duet along with the novice, with the instructions to synchronize their keystrokes with the novices' keystrokes as accurately as possible. An occluder prevented visual contact between the two participants of a pair (see Figure 2.1A). The data for each cell in this 2x2 within-participant design were derived from two different duets (blocks).

Novices were invited to a training session that took place 2-14 days in advance of the main experiment. The aim of the training session was to ensure that novices could produce error-free performances of their parts of the duets. Each novice played the novice part of each of the eight duets eight times. As in the main experiment each performance was preceded by four leading metronome beats at a tempo of 60 bpm. Novices were instructed to match this target tempo. All twelve novices managed to reach the criterion of three error-free performances in a row for all melodies.

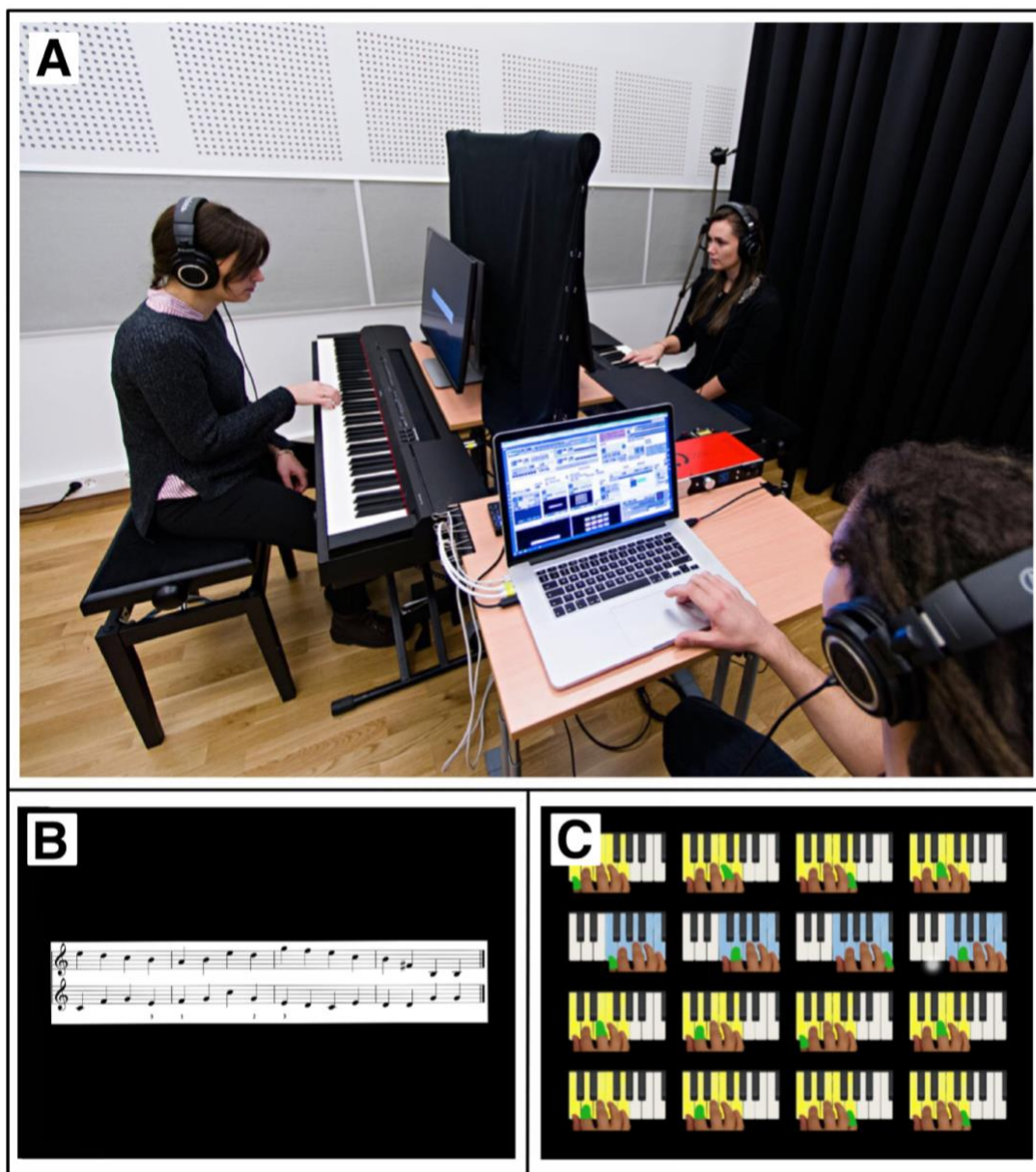


Figure 2.1: A) Experimental setup. B) Experts performed their own part based on a standard musical notation (upper row). For half of the duets experts received the full musical score including the novice's part (lower row). C) Novices performed their part based on a simple custom notation that was read line by line from top to bottom. Each picture corresponded to

one beat and showed one of two hand positions (color-coded with yellow and blue) and the finger to be used to press one of the piano keys (green).

2.3 Results

Before analyzing synchronization performance, we compared the variability of Inter-Keystroke-Intervals (IKIs) of individual performances of novices and experts. As expected, novices had larger individual variability in IKIs, i.e., higher SD of IKIs ($M = 44$ ms, $SD = 7$) than experts ($M = 35$ ms, $SD = 7$). A Welch t-test revealed that this difference was significant, $t(21.97) = 3.30$, $p = .003$ (Delacre, Lakens & Leys, 2017). Furthermore, experts' variability in IKIs during test trials was significantly higher ($M = 50$ ms, $SD = 9$), than during their individual performances ($M = 35$ ms, $SD = 7$), $t(21.08) = 4.68$, $p < .001$.

As a measure of interpersonal temporal coordination, we analyzed the asynchronies between novices' and experts' keystrokes in the test phase where the expert accompanied the novice. Of 512 asynchronies per pair, 12.5% were produced after a hand shift. A small portion of data points were excluded due to technical error (0.26%) or wrong pitches, insertions, omissions, incorrect alignment (6.24%). From the raw asynchronies, we derived the mean absolute asynchronies and the variability of absolute asynchronies. Both measures were analyzed with 2x2 repeated measures ANOVAs with the factors Familiarity with Novice Performance (yes/no) and Familiarity with Score (yes/no). Asynchronies with preceding hand shifts (Difficult Passages) and asynchronies without preceding hand shifts (Easy Passages) were analyzed separately because they reflect qualitatively different movements that imply different levels of difficulty for novices.

2.3.1 Mean Absolute Asynchronies

Panels A and B in Figure 2.2 show the results for mean absolute asynchronies in ms separately for Easy and Difficult Passages. For keystrokes from Easy Passages (Panel A) the

mean asynchrony between expert and novice was lower when the expert had heard the novice practice her melody beforehand ($M = 44$ ms, $SD = 8$) than when they had not heard the novice practice ($M = 47$ ms, $SD = 9$). Accordingly, a 2x2 repeated measures ANOVA showed a significant effect of Familiarity with Novice Performance, $F(1, 11) = 6.003$, $p = .032$ but no significant main effect of Familiarity with Score, $F(1, 11) = .005$, $p = .946$, and no significant interaction between the two factors, $F(1, 11) = .983$, $p = .343$.

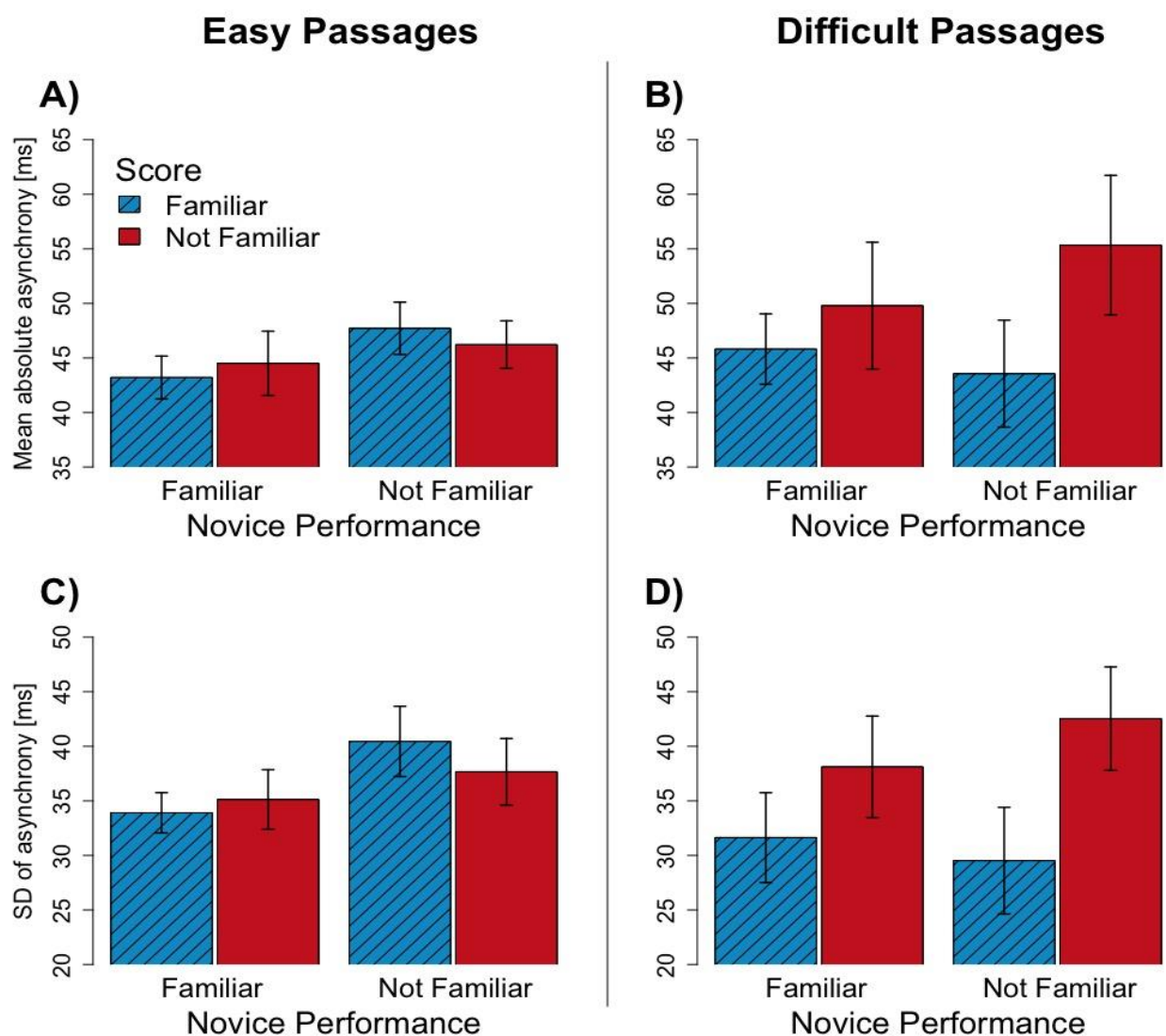


Figure 2.2: A) Mean absolute asynchronies between the keystrokes of experts and novices not preceded by a hand shift of the novice. B) Mean absolute asynchronies between the keystrokes of experts and novices preceded by a hand shift of the novice. C) Variability of asynchronies not preceded by a hand shift of the novice. D) Variability of asynchronies preceded by a hand shift of the novice. The error bars represent within-subject confidence intervals according to Cousineau (2005).

For keystrokes from Difficult Passages, i.e. keystrokes that were preceded by a novice's shift in hand position (Figure 2.2, Panel B) experts produced lower asynchronies when they were familiar with the score including the novice's part ($M = 45$ ms, $SD = 9$) than when they were not familiar with the score ($M = 53$ ms, $SD = 13$). Accordingly, the ANOVA showed a significant main effect of Familiarity with Score, $F(1, 11) = 14.918$, $p = .003$, but not of Familiarity with Novice Performance, $F(1, 11) = .187$, $p = .674$. There was no significant interaction between the two factors $F(1, 11) = 1.960$, $p = .189$. Signed asynchronies showed a similar pattern. On average signed asynchronies (expert - novice) were negative in all conditions with an overall mean of $M = -4$ ms, $SD = 59$.

2.3.2 Variability of Asynchrony (SD)

Panels C and D in Figure 2.2 show the results for the variability of asynchrony in terms of standard deviation in milliseconds separately for Easy and Difficult Passages. The results are in line with the results for mean absolute asynchrony. For keystrokes from Easy Passages (left panel) the standard deviation of asynchrony between expert and novice was lower when experts had heard the novice practice her melody beforehand ($M = 35$ ms, $SD = 6$) than when they had not heard the novice practice ($M = 39$ ms, $SD = 8$). The ANOVA revealed a significant main effect of Familiarity with Novice Performance, $F(1, 11) = 9.297$, $p = .011$ but no significant main effect of Familiarity with Score, $F(1, 11) = .164$, $p = .693$, and no significant interaction between the two factors, $F(1, 11) = 2.809$, $p = .122$.

For keystrokes from Difficult Passages, i.e. keystrokes that were preceded by a novice's shift in hand position (Figure 2.2, right panel) experts produced lower variability of asynchronies when they were familiar with the score including the novice's part ($M = 31$ ms, $SD = 8$) than when they were not familiar with the score ($M = 40$ ms, $SD = 11$). Accordingly, the ANOVA showed a significant main effect of Familiarity with Score, $F(1, 11) = 15.809$, p

= .002, but not of Familiarity with Novice Performance, $F(1, 11) = .237, p = .636$. There was no significant interaction between the two factors, $F(1, 11) = 1.199, p = .297$.

2.3.3 Cross-Correlation Analysis

To provide a further measure of adaptation based on experts' familiarity with a novice's performance, we performed a cross-correlation analysis in which we compared the timing patterns of novices with the corresponding patterns of experts at lag zero. This analysis showed that experts were higher correlated with novices after they had heard the novice performance (mean $R = 0.33$, $SD = 0.20$) than when they had not heard the novice performance (mean $R = 0.21$, $SD = 0.21$). A paired sample t-test showed a significant difference, $t(11) = 2.591, p = .025$. Therefore, the cross-correlation further corroborates the findings obtained in our analysis of absolute asynchronies.

2.3.4 Random Pairing Analysis

Furthermore, we assessed whether experts' adaptations were specific to the novice they coordinated with. We compared the correlation of IKI patterns of experts and novices for actual pairs with the correlations for random pairs. Random pairings were constructed in a way so that each expert was paired with each novice except for her actual partner, matched for melody and condition. We then calculated correlations of IKI patterns for the random pairs and computed the upper confidence intervals, a conservative estimate (see Zamm, Pfordresher & Palmer, 2015). These upper confidence intervals were then compared to the correlation values of the actual pairings. Lag 0 correlations among actual pairs was higher ($M = 0.33$, $SD = 0.22$) than the upper confidence intervals of lag 0 correlations for random pairs ($M = 0.12$, $SD = 0.10$). Using Bonferroni adjusted alpha levels of .025 per test (.05/2), the results of paired-samples t-tests showed significant differences between observed correlations and CIs of

random-pair-correlations when experts had heard the novice beforehand: $t(11) = 2.98, p = .012$ and not when experts had not heard the novice beforehand: $t(11) = 2.16, p = .054$.

2.3.5 Autocorrelation Analysis

Finally, to check for indications that novices' variability originates from expressive timing, we analyzed the autocorrelation of timing intervals in novices' performances. We calculated autocorrelation values for each novice and each melody's temporal pattern of inter-keystroke-intervals separately. Based on the melody length of 16 tones, i.e. 15 inter-keystroke-intervals, we were able to compute the correlation coefficients for lags 1 through 11. We found that none of the computed correlation coefficients reached the confidence limits of $\pm 2/\sqrt{n}$, with n being 15 IKI values per melody in our case.

2.4 Discussion

The present study examined how experts adapt their performance when playing piano duets with novices to ensure joint action coordination despite novices' suboptimal timing. The results showed that there are at least two different factors that enable experts to achieve coordination. During easy passages experts were able to improve coordination when they were familiar with novices' performances. In passages that were difficult for novices, experts were able to improve coordination when they were familiar with novices' scores.

Familiarity with novices' individual performance can only facilitate coordination if experts pick up on timing regularities in novices' performances that reflect novices' idiosyncratic timing patterns. Our results show that experts were able to identify and use such regularities quite well to predict the timing of the novices' performances. Our analysis of random pairings provides evidence that experts indeed identified idiosyncrasies. Furthermore, our finding that the average of the overall signed asynchronies was around zero, i.e., well below

the shortest possible reaction time, implies that experts predicted the timing of novice performance (Repp & Su, 2013). Thus, we propose the following mechanism for how experts achieved improved coordination based on being familiar with novices' performances: While listening to the novice, experts compare novices' timing to the timing predicted by their internal models (Repp & Knoblich, 2004). This allows them to generate an error matrix that they can use, during later joint performance, to modulate their predictions of novices' timing, enabling them to reduce asynchronies despite novices' suboptimal timing.

The second way in which experts can improve coordination with a novice is to predict performance difficulties for novices based on symbolic information about their task. In the present study, the crucial information consisted of symbols in the musical score indicating to the expert when the novice would have to change hand positions. Importantly, the marked changes in hand position did not imply any difficulty from an expert's point of view but posed difficulties selectively for novices. Our interpretation that experts relied on symbolic information is supported by the finding that the experts' prior exposure to the novices' sheet music only facilitated joint action coordination in difficult parts that required novices to shift their hand position. The symbols in the musical score might have acted as performance cues (see Ginsborg & King, 2012) that enabled experts to delay their own performance to give novices enough time to change the position of their hand, thereby departing from the constant tempo prescribed in the musical score. To achieve this, experts needed to convert the symbolic information in the musical score into a modulation of the timing of their motor commands.

It is likely that modulations of expert performance due to symbolic cues occur at a longer timescale than modulations of performance based on an error matrix derived from listening to novices' individual performances. Whereas converting symbolic information into a modulation of timing is largely explicit, continuous modulation of one's own predictions based on an error matrix occurs at the level of internal models and is likely to be implicit and

to draw on systems that enable massive parallel processing (Wolpert & Kawato, 1998). This interpretation is in line with Ragert and colleagues' (2013) finding that familiarity with playing another's part was detrimental to interpersonal coordination at the keystroke level associated with shorter timescales, but improved interpersonal coordination of bodily movements associated with longer timescales (see Davidson, 2009) and higher levels of the hierarchical structure of a musical piece (Koelsch, Rohrmeier, Torrecuso & Jentschke, 2013; Lerdahl & Jackendoff, 1993; Michael & Wolf, 2014).

Despite operating at different timescales, both postulated mechanisms enable individual musicians to prepare for interpersonal coordination before they actually engage in a joint performance and can contribute to our understanding of the impact of individual and joint rehearsals on joint action coordination (Ginsborg et al., 2006; Ginsborg & King, 2012). These offline-adaptations could reduce the amount of online adaptation and anticipation (van der Steen & Keller, 2013) required during the joint performance.

A potential alternative explanation for the lower asynchronies in easy passages after experts heard novices perform could be that experts form episodic memories of the novices' performances while listening to them. It has been shown that episodic memory can be used to correctly differentiate between different performances of the same piece with the help of prosodic cues (Palmer, Jungers & Jusczyk, 2001). However, it is unlikely that episodic memory drove expert performance in the current study. Here, experts were presented with four consecutive performances of a novice and our results suggest that some form of generalization occurred. In the episodic memory account, one would have to assume that each time interval across four highly similar performance was stored, which not only implies a high demand on memory, but would likely result in 'memory mixing' (Rijn, 2016). As musicians exhibit strong links between musical action, outcomes and representations thereof (Bangert & Altenmüller, 2003; Brodsky, Kessler, Rubinstein, Ginsborg & Henik, 2008; Haueisen & Knösche, 2001), an

explanation postulating an error matrix based on generalized asynchronies seems more parsimonious and plausible.

Considering our claim that the novices' timing was suboptimal we checked whether there were any indications that novice variability reflected musical intentions through expressive timing. In order to do so, we performed autocorrelation analyses to check for patterns of expressive timing (Desain & de Vos, 1990). We found no significant correlation coefficients, which indicates that expressive timing did not contribute substantially to the novices' variability. Thus, it seems very likely that novices' timing was indeed based on unintended performance difficulties with the task rather than some sort of musical style the novices intended to express.

With regard to implications for teaching, the present study highlights that experts can better ensure successful coordination with novices' suboptimal timing if they have sufficient advance information about novices' task and performance parameters. This may be a precondition for providing coordination scaffolds for novices when joint actions are performed for teaching purposes. The uni-directional feedback in the present study ensured that experts entirely carried the burden of ensuring joint action coordination. They were remarkably flexible in modulating their performance to adapt to novices' suboptimal timing, sacrificing the musical quality of their own performance to achieve interpersonal coordination.

An important goal for future research is to study which additional means experts use to provide a temporal scaffold for novices in teaching situations when novices also receive feedback from experts' performance, and when experts and novices can also visually perceive each other's actions. Some possible options including exaggerating movements (McEllin, Knoblich & Sebanz, in press), reducing performance variability (Vesper, van der Wel, Knoblich & Sebanz, 2011), and relying on emerging leader-follower dynamics (Konvalinka, Vuust, Roepstorff & Frith, 2010). A further interesting question is whether and when teachers

strategically avoid adjusting to novices to maximize learning opportunities for their students. This strategy could be especially useful when a teacher intends to convey to a student particular expressive timing patterns. Generally, studying how teachers adapt to their students' performance and determining when they decide not to adapt seems to be a promising way to further improve our understanding of how procedural skills are learned and taught through social interaction.

2.5 Follow-Up Study: What is the Contribution of Pitch Information?

2.5.1 Introduction

In the study presented in the previous chapter we found that expert pianists can adapt their predictions to the timing of a novice pianist, on the basis of perceiving a novice's performance. We suggested that experts do so by simulating their part while listening to the novice and thereby create an error matrix that captures the timing differences between internally simulated keypresses and the novice's actual timing. One of the questions that remain regarding this mechanism is whether the pitch information contributes to how accurately experts can predict a novice's timing. For instance, they could use the melodic contour, determined by the sequence of successive pitches, as a mnemonic device to remember where deviations occurred. Melodies can be used as mnemonic devices (James, 2013) and internally replaying them is sometimes hard to inhibit (Jakubowski, Finkel, Stewart & Müllensiefen, 2017). Alternatively, predicting a novice's timing may not depend on pitch information because the timing information itself exclusively drives experts' generation of an error matrix while listening to a novice's performance. Several effects that are thought to be based on action simulation and internal models have been shown to depend on timing information rather than acoustic properties (Flach, Knoblich & Prinz, 2004).

To address the question of whether pitch significantly contributes to the adaptation of an expert to the suboptimal timing of a novice based on familiarity with that novice's playing style, we retained the factor of Familiarity with Novice Performance from the original study and crossed it with whether experts had access to the pitch of the novice's part or not. Expert pianists either heard novices' performances as they were recorded during the original study, including pitch information, or they heard performances that retained the timing of the novice

performance but were realized with a percussive sound and thereby stripped of their pitch information.

2.5.2 Methods

Twelve expert pianists (8 women, 4 men, mean age = 26 years, SD = 4 years) participated in the experiment. All experts had at least 10 years of private piano lessons (M = 14 years, SD = 2 years). Seven of the experts had also participated in the original study (see 2.1.-2.4). The 5 new experts were recruited through flyers distributed in music schools and university campuses in Budapest. All participants gave their informed consent and received gift vouchers as compensation. This study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

Apparatus and Material

Participants played on two Yamaha digital pianos. Presentation of scores, auditory feedback, metronome beats, and data recording were implemented using a custom Max MSP patch. For this study we used novice performances recorded in midi format during the previous study (see 2.1-2.4). For each melody we choose recordings from five different novices so that we could obtain four recordings for each melody even for experts who had already participated in the previous study. In this way we could avoid that these experts simply adapted to recordings of the novice they had done the first experiment with. Recordings were chosen to minimize the amount of errors (particular omissions of notes) in the novice performances. Since the recordings were obtained from the original study, the characteristics of each performance is the same as in the original study. Each performance started in a lower hand position (thumb on C4) and required two shifts of hand position, first upwards (thumb from C4

to F4) and then downwards (thumb from F4 to C4). For more information on the novice parts please refer to section 2.2.2.

Procedure and Design

In this experiment we crossed the factors Familiarity with Novice Performance (yes, no) and Pitch (yes, no) in a 2x2 within-subjects design. As in the original study, Familiarity with Novice Performance determined whether the experts heard multiple performances of the novice before the coordination phase or not. The factor Pitch determined whether experts were informed about the pitch of the notes the novice was playing. In the Pitch – yes condition experts saw a sheet music representation of the novice’s part, while in the Pitch – no condition they did not. Additionally, when experts heard the novice’s performances, the performances were either replayed from a midi file with pitched piano sounds (Pitch – yes), i.e. as it was performed by the novice, or they were replayed with an unpitched percussive sound, instead (Pitch – no). Thus, timing information was retained in both kind of replays, while the pitch information was only retained in the Pitch – yes condition. As in the original study the manipulations were realized in a first phase that preceded the coordination phase.

The experiment consisted of four blocks, one for each condition. The order of conditions was counter-balanced across experts. In each block experts encountered each melody recorded by one novice per block. This means that experts had to adapt to all eight melodies under all four conditions recorded from different novices.

2.5.3 Results

As dependent variables we looked at the absolute asynchrony between the expert’s and the novice’s tone onsets, as well as the variability of these asynchronies. As we were mainly interested in the interaction of Pitch and Familiarity with Novice Performance, we restricted

the analysis to asynchronies from easy passages, i.e. passages in which the novice did not need to perform a hand position change (see section 2.2.2). This was justified because the previous study had shown that effects of familiarity with the novice's performance are only visible in easy passages.

A 2x2 repeated measures ANOVA with the factors Familiarity with Novice Performance and Pitch on the absolute asynchrony values showed a significant main effect of Familiarity with Novice Performance, $F(1, 11) = 7.889, p = .017, \eta^2 = .060$, but no significant main effect of Pitch, $F(1, 11) = 1.236, p = .290$. There was no significant interaction between the two factors, $F(1, 11) = 3.717, p = .080$ (see Figure 2.3A). Accordingly, the mean of absolute asynchrony was lower when experts had heard the novice's performance ($M = 42$ ms, $SD = 5$ ms) than when they had not heard the novice's performance ($M = 45$ ms, $SD = 6$ ms).

We had similar results for a 2x2 repeated measures ANOVA on the variability of absolute asynchrony values with the same factors. The ANOVA showed a significant main effect of Familiarity with Novice Performance, $F(1, 11) = 12.518, p = .005, \eta^2 = .092$, but no significant main effect of Pitch, $F(1, 11) = 3.656, p = .082$, and no significant interaction, $F(1, 11) = 0.028, p = .870$ (see Figure 2.3B). Variability was significantly lower when experts had heard the novice's performance ($M = 33$ ms, $SD = 5$ ms), than when they had not heard the novice's performance ($M = 36$ ms, $SD = 6$ ms).

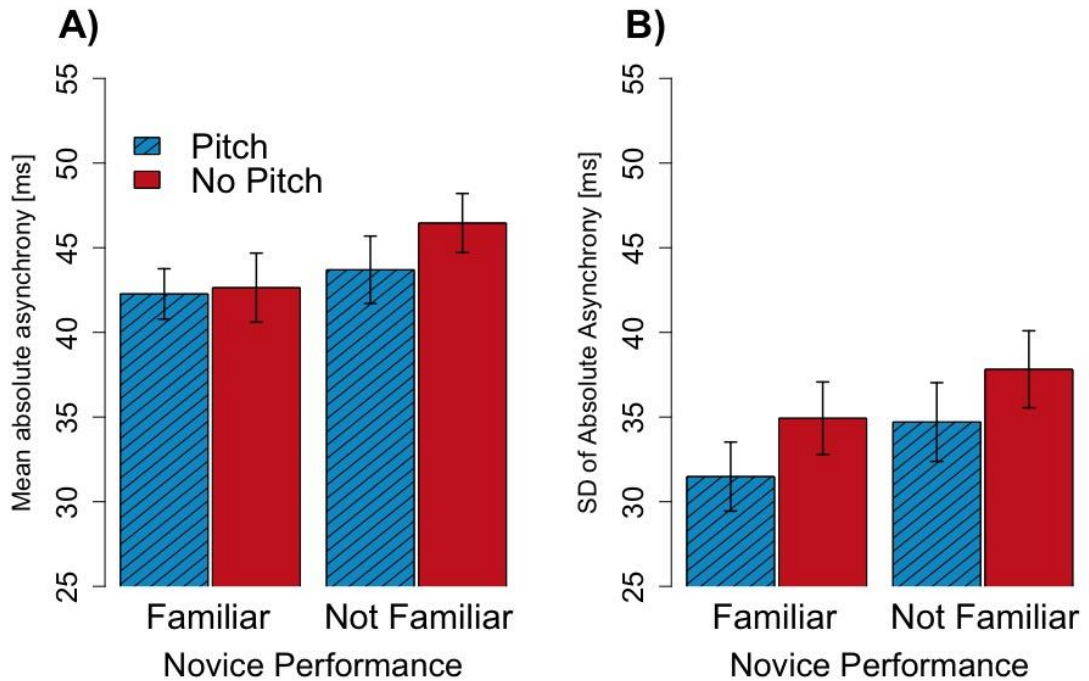


Figure 2.2: A) Mean absolute asynchronies between the keystrokes of playing experts and recordings of novices not preceded by a hand shift of the novice. B) Variability of asynchronies not preceded by a hand shift of the novice. The error bars represent within-subject confidence intervals according to Cousineau (2005).

2.5.4 Discussion

In the original study we found that familiarity with a novice's performance improved the expert's coordination with the novice's suboptimal timing. In the present study we investigated whether pitch information is a precondition for the improved expert coordination with a novice. As in the original study we did find a significant effect of familiarity with the novice's performance. Thus, we replicated the effect of familiarity with a novice's performance, where familiarity lead to better synchronization performance. However, we did not find a significant effect of pitch information and no significant interaction between pitch and familiarity with the novice's performance.

Based on the findings of the original study, we proposed that while listening to the novice's performance experts stored deviations between predictions of the novice's timing and the novice's actual timing in an error matrix. In the coordination phase experts can then use

this matrix to alter their timing to be better coordinated with novices. The findings of the current study support this interpretation and add to it in the following way. The fact that we did not find a significant effect of pitch in the current study indicates that improvements due to the familiarity with a novice's performance were not based on pitch information, but rather on the timing information that was retained across the two levels of the factor Pitch. This is consistent with the proposed mechanism of an error matrix in which temporal deviations are stored. It is also consistent with the literature on self-recognition tasks in which people's ability to recognize their past action despite sparse information is also attributed to internal models. Experiments in the auditory domain came to the conclusion that timing information was more important than acoustic features (Flach et al., 2004).

3 Study 2: Adaptation to Extraordinary Coordination Patterns in Joint and Individual Actions

3.1 Introduction

When humans engage in rhythmic joint actions, the underlying rhythm can act as a coordination smoother (Vesper, Butterfill, Knoblich & Sebanz, 2010) and allows for especially tight temporal coordination. In rhythmic interactions, such as joint music-making, temporal coordination can reach a precision of a 100th of a second (Keller, 2008). Temporal coordination, however, is easier for certain coordination patterns than for others. In-phase coordination has been shown to be the most stable coordination pattern for intra- as well as for interpersonal coordination, followed by anti-phase coordination which is less stable (Kay, Saltzman, Kelso & Schöner, 1987; Kelso, 1984; Richardson, Marsh, Isenhower, Goodman & Schmidt, 2007; Schmidt, Carello & Turvey, 1990; Schmidt & Turvey, 1994). These two coordination patterns have been called ‘canonical steady states’ (Schmidt & Richardson, 2008). This means that *extraordinary coordination patterns*, i.e., patterns other than in-phase or anti-phase, are especially challenging for temporal coordination, both during joint actions as well as during intrapersonal coordination of different limbs. Here, I will address two factors that could facilitate temporal coordination of extraordinary patterns. First, I will argue that despite extra effort being required during interpersonal coordination, adaptation to extraordinary patterns is actually more efficient during joint actions than during individual performance. Second, I will investigate whether the structure of action outcomes can provide a scaffold for achieving extraordinary patterns.

The requirement to produce an extraordinary coordination pattern, emerges for example, when three people stand around a large tent pole and use three hammers to drive the pole into the ground. The pole is too narrow to allow for in-phase coordination. To avoid collisions their timing has to be coordinated in a way that maximizes the time between each person's stroke and the preceding and succeeding strokes. Since there are three people involved, the coordination is organized around a phase shift of $360 \text{ degrees} / 3$, which corresponds to 120 degrees and falls in the instable area between in-phase (0 degree) and anti-phase (180 degree). Whereas this may seem like a rare example of an interaction that requires an extraordinary coordination pattern, in the domain of music-making patterns like this are common. Musicians regularly master temporal coordination despite difficult coordination patterns both in bimanual solo performance as well as, interpersonally, in joint music-making.

Polyrhythms, for example, combine rhythms at non-integer multiples of each other such as 3:2 and 4:3 (Deutsch, 1983; Klapp, 1979; Summers, Davis & Byblow, 2002), and are used in a wide variety of music genres (see e.g. Arom, 1991; Copland, 2004; Folio, 1945; Locke, 1982; Pieslak, 2007; Poudrier & Repp, 2013; Pressing, 2002; Taylor, 2003). Furthermore, some instrumentalists need to produce phase shifted movements with different limbs due to the physical setup of their instrument. Small phase shifts between actions, for example, have to be produced on bimanual instruments, where pitch, or string selection has to precede the initiation of the sound, e.g., on the violin where a string has to be pushed down with the left hand before the bow is moved or the string is plugged with the right hand (Schoonderwaldt & Altenmüller, 2014). Larger phase shifts can occur due to the mechanics of some instruments that result in considerable delays between instrumental movements and tone onsets. Pipe organs, for example, can exhibit delays of up to 150 ms (Pollard, 1968) that vary according to pitch and may vary for different manuals and pedals.

Coordinating tone onsets during music-making therefore entails compensating for these delays by introducing offsets between instrumental actions. In mixed ensembles different instruments can exhibit distinct delays between movement initiation and sound onset (Eronen & Klapuri, 2000; Herrera-Boyer, Peeters & Dubnov, 2003). Nevertheless, musicians have to coordinate their tone onsets by compensating for various delays.

3.1.1 Previous literature

Early experiments on interlimb coordination found specific patterns of break-down suggesting that coordination of rhythmic limb movements is governed by the same laws as coupled oscillators (Kay et al., 1987; Kelso, 1984). When Haken, Kelso and Bunz (1985) modelled interlimb coordination in terms of two coupled oscillators, coupling strength was taken to be an important parameter that governs how the coordination unfolds. The stronger the coupling the easier it is to maintain simple relations like in-phase coordination and the harder it is to maintain extraordinary coordination patterns.

Coupling strength influences both intrapersonal as well as interpersonal coordination. This has been shown for example in an interpersonal leg oscillating task by R. Schmidt and colleagues (1990). In a subsequent paper, Schmidt, Bienvenu, Fitzpatrick & Amazeen (1998) found coupling to be significantly lower during interpersonal limb coordination than during intrapersonal limb coordination. In a more fine-grained visual coupling manipulation, Richardson et al. (2007) compared peripheral visual coupling with direct visual coupling and found significantly more un-instructed in-phase coordination in the direct vision condition than in the peripheral vision condition, presumably due to the stronger coupling in the former, as weaker coupling reduces the tendency to fall into in-phase coordination.

As the coupling between two limbs of two individuals has been shown to be weaker than the coupling between two limbs of the same person (R. Schmidt et al., 1998), I argue that

it is easier to distribute unstable coordination patterns among the limbs of multiple musicians. Ugandan xylophone music is a perfect example of how this can be exploited by composers and musicians. Kubik studied *Amadinda* music (1960, 1964), which is traditionally performed by three musicians on one instrument. He describes the emergence of intricate *inherent rhythms*, which are “played at an incredible speed” (Kubik, 1964). These inherent rhythms however are not played by any individual musician, but instead shared across the first and second *Amadinda* player in such a way that parts of the pattern that are in-phase are produced intrapersonally, whereas more complex phase relations are distributed interpersonally.

Whereas it might be easier to adapt to extraordinary coordination patterns in joint performances, musicians are able to perform extraordinary patterns both in joint and individual coordination settings. I therefore investigated a second factor that could facilitate adaptation across joint and individual performances. On a physiological level, the co-activation of homologous muscles has been proposed to account for the difficulties of producing certain bimanual patterns (Kelso, 1984). However, there is evidence for the claim that the structure of perceptual action effects plays an important role in enabling intra- and interpersonal rhythmic coordination (Mechsner, Kerzel, Knoblich & Prinz, 2001).

Mechsner and colleagues developed a task that isolated the effects of homologous muscles and perceptual symmetry of movements on rhythmic movement coordination. They found that perceptual symmetry was a better predictor for the stability of interlimb coordination. In another experiment Mechsner et al. (2001) expanded this finding to a polyrhythm task, where musically untrained participants managed to produce circular motions in a 4:3 frequency ratio, when their goal was to perceptually align two rotating flags. Due to differences in the transmission from cranks to flags a successful alignment of flags meant a 4:3 ratio between the movements of the two. Mechsner et al. took this as another piece of evidence that hand coordination is governed by perceptual features of action outcomes.

Researchers have studied similar facilitation effects in the auditory domain. Sonification, for example, can facilitate bimanual skill acquisition in the context of extraordinary coordination patterns (Dyer, Stapleton & Rodger, 2017a). Dyer and colleagues argued that these facilitating effects are caused by the perceptual unification of complex coordination patterns. The structure of auditory action effects seems to be an important factor for the unification. The same authors have shown the superiority of melodic over rhythmic sonification and argue that melodic content in addition to temporal content helps to better structure complex target patterns (Dyer, Stapleton & Rodger, 2017b).

The role and strategies of the third player in Amadinda music is a good example of how musicians use auditory action effects to overcome motoric difficulties. The third Amadinda musician starts to play last and often has to execute especially complicated patterns. Kubik writes “it would be impossible to play this pattern [...] by referring it metrically to one of the basic parts [...]. The only chance to come in is by ‘thinking’ this pattern as a gestalt in its own right” (Kubik, 1964). Forming this Gestalt is facilitated by the two previous players jointly providing the pattern in their lowest two notes. “When the pattern he listens to has become very firm in his imagination and the rest of the composition sounds only in the background he simply repeats two octaves higher what he hears” (Kubik, 1964).

Whereas all of the experimental studies reviewed so far used bimanual coordination in individuals, the Amadinda example and some studies of joint action suggest that there may be similar effects of action effect structure for joint performances. In particular, there is evidence that joint action is often based on representing joint action outcomes that combine individual action effects into a pattern that is more than the sum of its parts. Loehr, Kourtis, Vesper, Sebanz and Knoblich (2013) invited pianists to duet with each other, while a computer introduced errors from time to time. An analysis of the performers’ EEG signals showed differences in ERPs that depended on whether the artificial error affected only the individual

outcomes, i.e. wrong note, but same harmony, or the joint outcomes as well, i.e. wrong note, different harmony. Aucouturier and Canonne (2017) asked duos of improvising musicians to convey various social intentions and audio recorded their performances. Possible auditory markers of these social intentions were retrievable only from their combined audio signals, i.e. the joint outcome and not from any individual audio signal.

3.1.2 Current Study

I investigated two potential factors that may affect how able people are to produce extraordinary temporal patterns during rhythmic performances. First, I considered the inherent coupling strength characterizing different coordination settings. While the coupling between two limbs of the same person is strong, coupling between the two limbs of two people is weaker and the coupling between the limb of a person and a computer produced sounds is basically absent (unidirectional). Coupling strength is one of the factors that determines how difficult it is to maintain unstable phase relations (Large & Kolen, 1994). To investigate effects of coupling strength as mediated by coordination contexts such as intrapersonal and interpersonal coordination, I devised an adaptation paradigm, where participants needed to align tone onsets triggered by finger tapping. In order to align tone onsets participants needed to adapt to an artificially introduced constant delay and thereby to an extraordinary coordination pattern. This coordination pattern required an interlimb phase relation of about 26.7 degrees.

If coupling strength is indeed an important factor for the adaptation to extraordinary coordination patterns, we should see better adaptation in coordination settings with weaker inherent coupling than in coordination settings with stronger inherent coupling. In Experiment 1b, I aimed to replicate Experiment 1, under conditions that I expected to improve participants' baseline performance during interpersonal coordination by asking them to perform the task

bimanually in all conditions which has been shown to reduce timing variability (Drewing & Aschersleben, 2003; Helmuth & Ivry, 1996).

The second factor, structure of action effects, was tested in Experiment 2. Based on the previous literature I examined the role of melodic structure of action effects. Experiments 1 and 1b provided participants with concurrent auditory feedback realized in percussive sounds to highlight the rhythmic pattern. In Experiment 2, I tested whether adding a harmonic pitch dimension to the action effects helps to structure and retain complex target patterns (Dyer et al., 2017b). To investigate whether similar effects occur independently of coupling strength, as the findings by Mechsner et al. (2001) suggest, I tested this in various coordination contexts in Experiment 2.

3.2 Experiment 1 - Methods

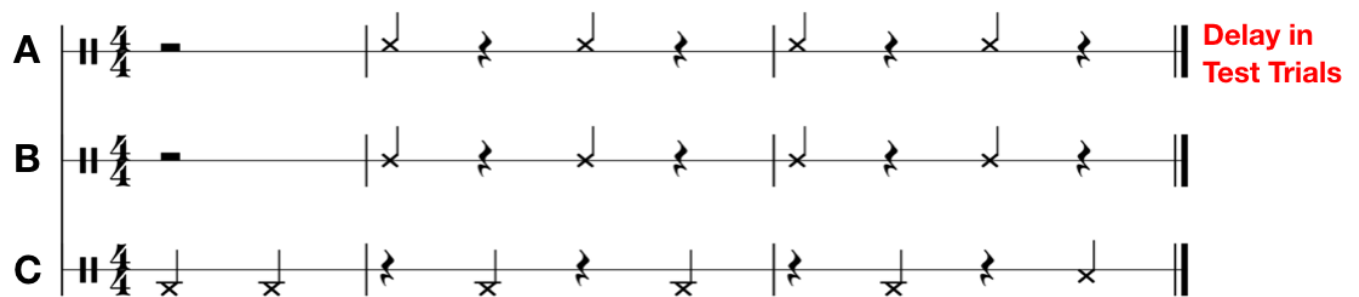
3.2.1 Participants

Participants were recruited through flyers distributed in music schools in Budapest, on Facebook sites related to music-making in Budapest and on the CEU campus. Informed consent was obtained from all participants. Participants received gift vouchers as compensation. This study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary. I invited 16 participants, of which 4 participants were not able to fulfill the minimal task requirements, i.e. they completed less than 80% of all trials successfully. The remaining 12 participants (4 women, 8 men, mean age = 25 years, SD = 4 years) were included in the data analysis. All participants in Experiment 1 had completed at least 5 years of private lessons on a musical instrument ($M = 11$ years, $SD = 4$ years).

3.2.2 Apparatus and Material

Participants were tapping on iPads Pro that were connected via the iOS app MIRA 1.2.2 to a custom patch created in Max MSP 7.3.1 running on a Mac computer. The Max patch created tap contingent sounds that were played back to the participants via headphones. Dependent on the current condition, the patch sometimes introduced a delay of 89 ms between tap registration and sound playback. An occluder prevented both participants from seeing their own iPad and their partner's iPad in the Interpersonal condition.

Experiment 1 and 1b



Experiment 2



Figure 3.1: Sheet music representation of the target rhythm. In the Intrapersonal condition, staff A was always played by a participant's right hand, whereas staff B was played by the same participant using her left hand, while the computer filled in staff C. In the Interpersonal condition staff A was played by the participant on the right, while staff B was played by the participant on the left, both using their right hands. Staff C was again filled in by the computer. In the Computer condition, staff B and C were filled in by the computer, while the participant used her right hand to play staff A. This was the same for Experiment 1 and 2. In Experiment 1b, however, participants used both hands in all conditions. In the Interpersonal and the Computer condition the Max patch randomly selected which hand triggered the sound. During test trials the tone onsets of notes in staff A were artificially delayed by a Max patch.

3.2.3 Procedure and Design

Participants were instructed to tap out the rhythm depicted in Figure 3.1 at an inter-tap-interval (ITI) of 1200 ms. Two metronome clicks were followed by four finger taps of the

participant. Four additional metronome clicks subdivided participants' taps to assure a steady tempo. The last metronome click was realized as a bell-like sound to signal the end of the trial. In the Intrapersonal condition (high coupling), participants produced the rhythm alone using both hands, with one hand producing one percussion sound and the other hand producing a different percussion sound. Participants were asked to synchronize the two sounds produced by their two hands. In the Interpersonal condition, the two sounds needed to be synchronized across the two right hands of two participants, with the computer fulfilling the same metronomic function. In the Computer condition, the computer provided the metronomic structure, but also produced sounds to synchronize with (like one's other hand in the Intrapersonal condition and like one's partner in the Interpersonal condition). In this case the participant produced one percussion sound just as in the Interpersonal condition but with the computer as a perfectly timed partner.

In test trials, I introduced an artificial delay of 89 ms to the production of the sounds notated in staff A (see Figure 3.1). Hence, in order to align tone onsets, participants needed to compensate for this delay. The delay corresponded to 26.7 degrees of the 1200ms interval in phase space, a non-standard phase relation between two sound-producing movements.

Each experiment consisted of four segments (see Figure 3.2 – Panel A): Two segments involved joint performances (Interpersonal condition) and two segments were performed individually (Intrapersonal and Computer condition). Each pair of participants started either with the two segments in the Interpersonal condition or the two segments in the Intrapersonal and Computer condition. Who played which staff in the Interpersonal condition and thereby also who experienced the delay was switched after the first Interpersonal segment, so that each participant performed the Interpersonal condition playing staff A in one segment, which included the delay in some trials, and staff B in the other segment.

Each segment consisted of 24 baseline trials and 60 delay trials in the following arrangement (see Figure 3.2 – Panel B): 6 non-delay baseline trials, 20 delay trials (Block 1), 6 non-delay baseline trials, 20 delay trials (Block 2), 6 non-delay baseline trials, 20 delay trials (Block 3), 6 non-delay baseline trials. This trial arrangement allowed us to collect baseline data in non-delay trials and to assess changes in performance over the time of 60 delay-trials. Participants were told in advance how many trials were coming up and whether or not they would include the delay.

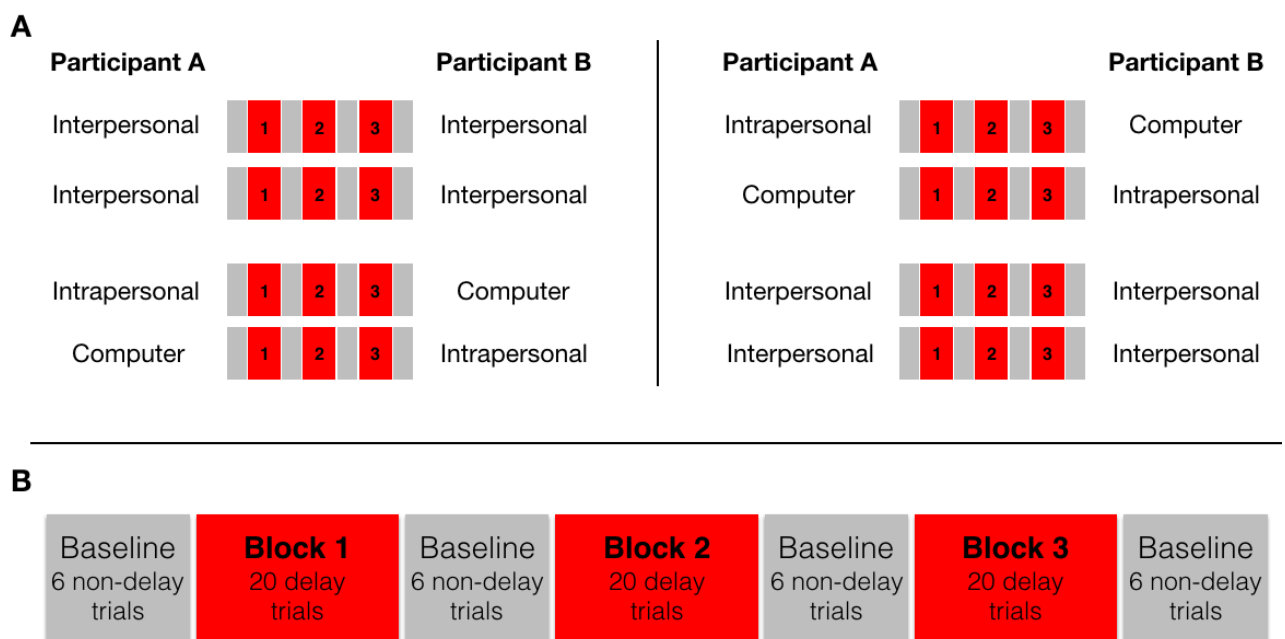


Figure 3.2: Panel A) Each dyad completed four segments. Participants either started with two segments of the Interpersonal condition and then proceeded to individually complete one segment each of the Intrapersonal and the Computer condition (Panel A - left), or they started with the Intrapersonal and the Computer condition, before completing two segments of the Interpersonal condition (Panel A - right). In one of the Interpersonal segments participant A experienced the delay and in the other segment participant B experienced the delay. Panel B shows the trial arrangement within one segment, which was the same in all conditions/throughout all segments.

3.2.4 Analysis

Performance measures were computed from the timestamped midi data. First, I checked whether participants followed the task of producing the required rhythm. The interspersed metronome clicks allowed us to calculate target times for the tone onsets produced by the participants. I computed the temporal deviation from these target times. In the Intrapersonal condition this measure was computed for each hand, in the Interpersonal condition, for each

participant and in the Computer condition only for the participant's right hand. If a tap fell outside a target window of ± 300 ms around the target time, it was considered an error. A missing tap or too many taps between two metronome clicks were also considered to be errors.

The main dependent variables were unsigned asynchrony and its variability. Unsigned asynchrony between tone onsets in milliseconds is a direct measure of how well participants managed to align two sounds in time. Variability of unsigned asynchrony, calculated as the standard deviation of unsigned asynchrony also in milliseconds, describes how stable this temporal alignment is. To assess how well participants performed in delay trials with respect to their baseline performance in non-delay trials, I computed baseline corrected values for all asynchrony measures. For this calculation I took the asynchrony values and subtracted the condition-specific baseline performance for each participant. A baseline corrected value of around zero milliseconds therefor implied that the performance in the delay trials was equal to the performance in the non-delay trials.

To investigate whether participants' group average improved over time, I computed 3 x 3 ANOVAs, with the two within-subjects factors Condition (Intrapersonal, Interpersonal and Computer) and Block (One, Two and Three). Significant results for Block imply a significant change of performance over the course of three blocks of delay trials. To assess whether individual participants managed to reach their baseline performance, I split each of the three blocks of delay trials in each condition into two halves which results in six bins spanning ten trials each. I then calculated 95 % confidence intervals for mean unsigned asynchrony and its variability for each of the 6 bins. Using both measures, asynchrony and variability, I determined in how many bins the non-delay baseline was encompassed by the confidence interval, i.e. in how many bins the performance on delay trials was not significantly different from baseline performance.

3.3 Experiment 1 - Results

3.3.1 Accuracy

Trials that included missing, additional or misplaced taps were marked as errors. These trials were excluded from further analysis. The mean accuracy for the participants in Experiment 1 was 93.49 % (SD = 4.32 %). A Greenhouse-Geiser corrected one-way ANOVA with the factor Condition ($M_{\text{Intrapersonal}} = 93.35\%$, $SD_{\text{Intrapersonal}} = 4.85\%$, $M_{\text{Interpersonal}} = 90.58\%$, $SD_{\text{Interpersonal}} = 8.24\%$, $M_{\text{Computer}} = 96.53\%$, $SD_{\text{Computer}} = 4.69\%$) revealed that the main effect of condition fell just above a significance level of .05, $F(1.531, 16.845) = 3.707$, $p = .056$, $\eta^2 = .146$.

3.3.2 Asynchrony in Baseline Trials

For the performance in non-delay baseline trials I observed the lowest unsigned mean asynchrony for the Intrapersonal condition ($M = 14$ ms, $SD = 3$ ms), with Joint ($M = 38$ ms, $SD = 12$ ms) and Computer ($M = 33$ ms, $SD = 13$ ms) showing comparable levels of asynchrony (see Figure 3.3 - Panel A). I computed a one-way ANOVA for the factor Condition (Intrapersonal, Interpersonal and Computer). The Greenhouse-Geiser corrected ANOVA revealed a significant effect for Condition, $F(1.669, 18.356) = 18.963$, $p < .001$, $\eta^2 = .520$. Post-hoc paired-samples t-tests showed that this was mainly due to the lower unsigned asynchrony in the Intrapersonal condition, as compared to the Joint, $t(11) = 6.613$, $p < .001$, $d = 1.909$, and the Computer condition, $t(11) = 5.277$, $p < .001$, $d = 1.523$. There was no significant difference between the Joint and the Computer condition $t(11) = .955$, $p = .360$, $d = .276$.

3.3.3 Asynchrony in Delay Trials

To assess participants' improvements across delay trials, I computed a 3 x 3 ANOVA with the factors Condition (Individual, Joint and Computer) and Block (One, Two and Three).

The corresponding plot for unsigned asynchronies can be seen in Figure 3.4 – Panel A. Degrees of freedom were corrected using Greenhouse-Geiser estimates of sphericity. The ANOVA revealed both a main effect of Condition, $F(1.277, 14.050) = 5.531, p = .027, \eta^2 = .159$, and Block, $F(1.435, 15.780) = 18.124, p < .001, \eta^2 = .062$, but no significant interaction $F(2.044, 22.483) = 2.516, p = .102, \eta^2 = .018$. Post-hoc comparisons revealed that the main effect of condition was driven by the lower unsigned asynchrony in the Computer condition. There was no significant difference between the Intra- and the Interpersonal condition, $t(11) = .500, p = .627, d = .144$, nor between the Intrapersonal and the Computer condition, $t(11) = 2.013, p = .069, d = .581$. Compared to the Computer condition the asynchrony was higher in the Interpersonal condition, $t(11) = 5.043, p < .001, d = 1.456$. Post-hoc comparisons further revealed that the performance in the Intrapersonal condition in the first block ($M = 55$ ms, $SD = 26$ ms) and in the last block ($M = 47$ ms, $SD = 28$ ms) did not differ significantly, $t(11) = 1.557, p = .148, d = .449$, whereas the performance in the Interpersonal condition ($M_{\text{First}} = 64$ ms, $SD_{\text{First}} = 20$ ms, $M_{\text{Last}} = 44$ ms, $SD_{\text{Last}} = 12$ ms) improved significantly, $t(11) = 4.747, p < .001, d = 1.370$. There was also a significant improvement in the Computer condition ($M_{\text{First}} = 38$ ms, $SD_{\text{First}} = 10$ ms, $M_{\text{Last}} = 32$ ms, $SD_{\text{Last}} = 11$ ms), $t(11) = 3.148, p = .009, d = .909$.

A similar ANOVA was computed for baseline-corrected unsigned asynchronies (see Figure 3.4 – Panel B). Degrees of freedom were corrected using Greenhouse-Geiser estimates of sphericity. The ANOVA revealed both a main effect of Condition, $F(1.430, 15.725) = 8.386, p = .006, \eta^2 = .289$, and Block, $F(1.434, 15.780) = 18.124, p < .001, \eta^2 = .043$, but no significant interaction $F(2.044, 22.483) = 2.516, p = .102, \eta^2 = .013$. In the first block, the baseline-corrected performance in the Interpersonal condition ($M = 26$ ms, $SD = 27$ ms) was statistically not different from the performance in the Intrapersonal condition ($M = 40$ ms, $SD = 26$ ms), $t(11) = 2.074, p = .062, d = .599$, but significantly higher than the performance in the Computer condition ($M = 5$ ms, $SD = 17$ ms), $t(11) = 2.402, p = .035, d = .693$. Due to the improvement

of the performance in the Interpersonal condition, this pattern was reversed for the last block, where the baseline-corrected performance in the Interpersonal condition ($M = 6$ ms, $SD = 20$ ms) was significantly lower than the performance in the Intrapersonal condition ($M = 33$ ms, $SD = 29$ ms), $t(11) = 3.626$, $p = .004$, $d = 1.047$, but not statistically different from the performance in the Computer condition ($M = -1$ ms, $SD = 18$ ms), $t(11) = 1.022$, $p = .329$, $d = .295$. When tested against 0 (here: equal to baseline), a one-sample t-test revealed that in the Computer condition participants performance was not significantly different from baseline performance already in the first block, $t(11) = .983$, $p = .347$, $d = .284$. In the Interpersonal condition, participants' performance was significantly different from baseline in the first block, $t(11) = 3.306$, $p = .007$, $d = .954$, but not in the last block, $t(11) = 1.008$, $p = .335$, $d = .291$.

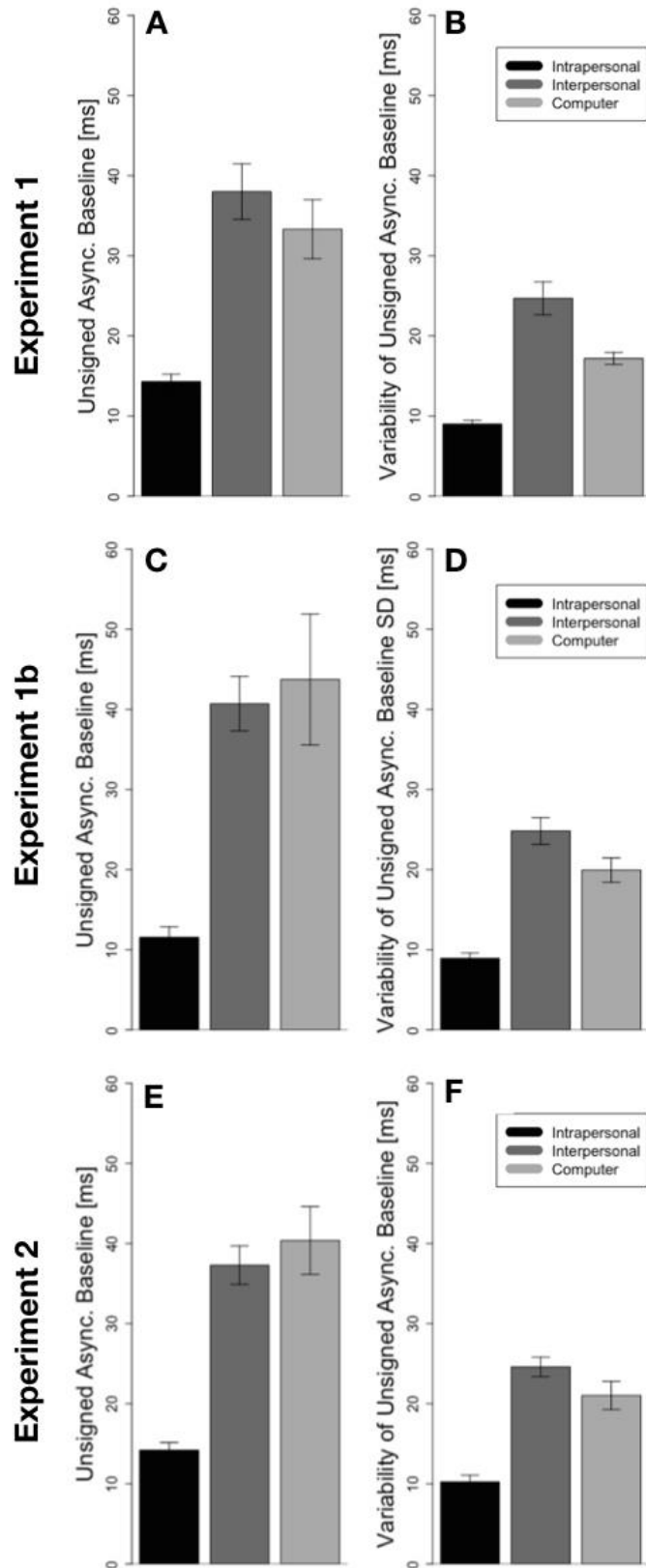


Figure 3.3: Baseline performance from non-delay trials for experiments 1, 1b and 2. The Panels on the left (A, C, and E) show unsigned asynchronies and the panels on the right show standard deviation of unsigned asynchronies (Panels B, D and F). Error bars display standard errors in each condition.

3.3.4 Variability in Baseline Trials

As we turn to the variability of unsigned asynchrony, we see a slightly different picture. In non-delay baseline trials I observed the lowest variability in the Intrapersonal condition ($M = 3$ ms, $SD = 2$ ms), with Interpersonal ($M = 12$ ms, $SD = 10$ ms) showing the highest variability and Computer ($M = 6$ ms, $SD = 4$ ms) falling in between (see Figure 3.3 – Panel B). I computed a one-way ANOVA for the factor Condition (Intrapersonal, Interpersonal and Computer). The Greenhouse-Geiser corrected ANOVA revealed a significant effect for Condition, $F(1.168, 12.845) = 38.107, p < .001, \eta^2 = .690$. Post-hoc paired-samples t-tests showed that variability in the Intrapersonal condition was significantly lower than in the Computer condition, $t(11) = 11.416, p < .001, d = 3.296$ and variability in the Computer condition was significantly lower than in the Interpersonal condition, $t(11) = 3.584, p = .004, d = 1.035$.

3.3.5 Variability in Delay Trials

To assess participants' reduction of variability across delay trials, I computed a 3 x 3 ANOVA with the factors Condition (Intrapersonal, Interpersonal and Computer) and Block (1, 2, and 3). The corresponding plot for variability can be seen in Figure 3.4 – Panel C. Degrees of freedom were corrected using Greenhouse-Geiser estimates of sphericity. The ANOVA revealed both a main effect of Condition, $F(1.478, 16.259) = 35.871, p < .001, \eta^2 = .544$, and Block, $F(1.483, 16.313) = 11.853, p = .001, \eta^2 = .104$, but no significant interaction, $F(1.908, 20.99) = 1.810, p = .189, \eta^2 = .028$. Post-hoc comparisons showed that variability in the Intrapersonal condition was significantly lower than in the Computer condition, $t(11) = 6.617, p < .001, d = 1.910$, which was in turn significantly lower than the variability in the Interpersonal condition, $t(11) = 6.401, p < .001, d = 1.848$. Post-hoc comparisons furthermore revealed that the performance in the Intrapersonal condition in the first block ($M = 18$ ms, $SD = 6$ ms) and in the last block ($M = 16$ ms, $SD = 4$ ms) differed significantly, $t(11) = 2.333, p =$

.04, $d = .674$, as did the performance in the Interpersonal condition ($M_{\text{First}} = 34$ ms, $SD_{\text{First}} = 7$ ms, $M_{\text{Last}} = 27$ ms, $SD_{\text{Last}} = 6$ ms), $t(11) = 4.350$, $p = .001$, $d = 1.256$. There was also a significant reduction of variability in the Computer condition ($M_{\text{First}} = 23$ ms, $SD_{\text{First}} = 4$ ms, $M_{\text{Last}} = 21$ ms, $SD_{\text{Last}} = 5$ ms), $t(11) = 3.130$, $p = .010$, $d = .904$.

A similar ANOVA was computed for baseline-corrected variability (see Figure 3.4 – Panel D). The Greenhouse-Geiser corrected ANOVA only revealed a main effect of Block, $F(1.483, 16.313) = 11.853$, $p = .001$, $\eta^2 = .070$. Neither the main effect of Condition, $F(1.258, 13.838) = 1.130$, $p = .323$, $\eta^2 = .041$, nor the interaction, $F(1.908, 20.99) = 1.810$, $p = .189$, $\eta^2 = .018$, were significant. Comparing the baseline-corrected variability measures to zero in one-sample t-tests, showed that in the Intrapersonal condition even in the last block participants did not reach the baseline performance, $t(11) = 7.146$, $p < .001$, $d = 2.063$. There was also a significant difference between the performance in delay trials and the baseline in the Computer condition, $t(11) = 2.423$, $p = 0.034$, $d = .699$. However, the performance in the last block of the Interpersonal condition was not significantly different from baseline, $t(11) = .783$, $p = 0.450$, $d = .226$, indicating that baseline performance was reached in the Interpersonal condition.

3.3.6 Unsigned Asynchrony and its Variability Combined

Table 3.1 offers a detailed overview of individual differences among participants. Three participants managed to reach their baseline performance in at least one bin in each condition. Five more participants reached baseline performance in at least one bin in the Interpersonal and the Computer condition. Three participants managed to reach baseline only in the Computer condition and one participant did not reach their baseline in any of the three conditions.

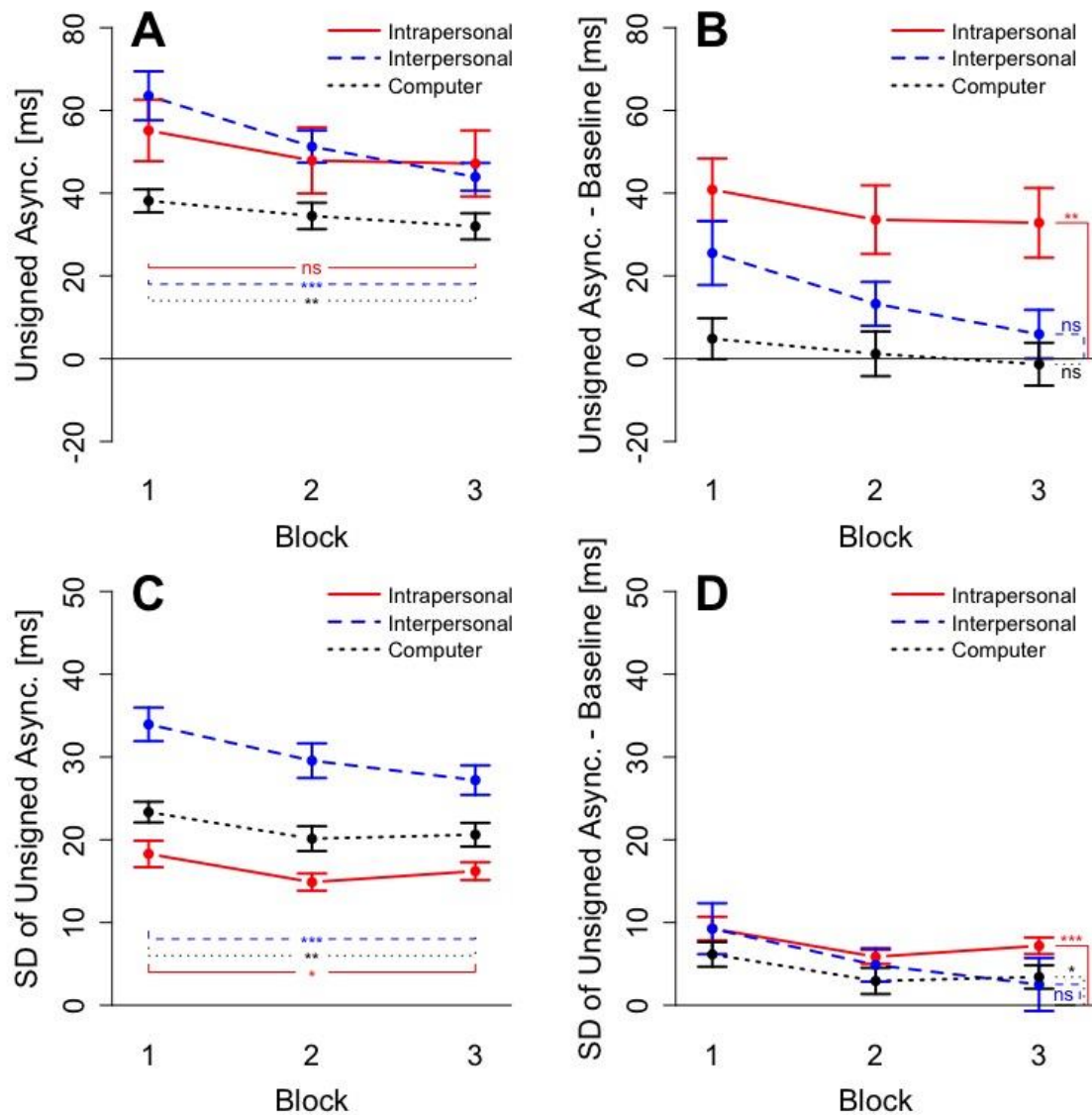


Figure 3.4: Results for delay trials of Experiment 1. All panels show Block on the x axis. Panel A) shows the results for unsigned asynchrony. Panel B) also shows unsigned asynchronies but corrected for the baseline from non-delay trials in each condition. Panel C) shows performance in terms of variability (SD) of unsigned asynchrony. Panel D) shows baseline-corrected variability. Error bars in all four panels show standard errors.

3.4 Experiment 1b – Methods

3.4.1 Participants

For Experiment 1b, I invited 14 participants of which 2 were not able to produce more than 80 % of all trials successfully. The remaining 12 participants (9 women, 3 men, mean age = 26 years, SD = 4 years) were included in the analysis. All participants in Experiment 1b had

completed at least 5 years of private lessons on a musical instrument ($M = 13$ years, $SD = 6$ years).

3.4.2 Procedure and Design

Experiment 1b was almost identical to Experiment 1, with the following crucial difference. In an attempt to reduce participants' baseline performance in the Computer and the Interpersonal condition, I introduced bimanual actions to all conditions (Drewing & Aschersleben, 2003; Helmuth & Ivry, 1996). Hence, in Experiment 1b, participants used two hands to control their iPad not just in the Intrapersonal condition, but also in the Interpersonal and in the Computer condition, in which participants in Experiment 1 only used their right hand. Two buttons had to be pressed simultaneously. Which button was used to trigger one piano sound per iPad was chosen randomly by the Max patch. This means that in all conditions of Experiment 1b participants performed their taps bimanually. The produced action effects, however, were identical to those in Experiment 1.

3.5 Experiment 1b - Results

3.5.1 Accuracy

The mean accuracy for the participants in Experiment 1b was 92.76 % ($SD = 5.55$ %). A Greenhouse-Geiser corrected one-way ANOVA with the 3-level factor Condition returned no significant differences ($M_{\text{Intrapersonal}} = 92.66$ %, $SD_{\text{Intrapersonal}} = 8.80$ %, $M_{\text{Interpersonal}} = 88.59$ %, $SD_{\text{Interpersonal}} = 12.21$ %, $M_{\text{Computer}} = 97.02$ %, $SD_{\text{Computer}} = 2.71$ %), $F(1.673, 18.407) = 3.013$, $p = .081$, $\eta^2 = .142$.

3.5.2 Asynchrony in Baseline Trials

As in Experiment 1, for the performance in non-delay baseline trials I observed the lowest unsigned mean asynchrony for the Intrapersonal condition ($M = 12$ ms, $SD = 5$ ms), with Interpersonal ($M = 41$ ms, $SD = 12$ ms) and Computer ($M = 44$ ms, $SD = 28$ ms) showing comparable levels of asynchrony (see Figure 3.3 – Panel C). A Greenhouse-Geiser corrected one-way ANOVA revealed a significant effect for Condition, $F(1.232, 13.556) = 13.503$, $p = .002$, $\eta^2 = .418$. Post-hoc paired-samples t -tests showed that this was due to the lower asynchrony in the Intrapersonal condition, as compared to the Interpersonal $t(11) = 9.133$, $p < .001$, $d = 2.636$, and the Computer condition, $t(11) = 4.128$, $p = .002$, $d = 1.192$, whereas there was no significant difference between the Interpersonal and the Computer condition $t(11) = .362$, $p = .724$, $d = .105$.

3.5.3 Asynchrony in Delay Trials

The corresponding plot for the 3x3 ANOVA with the factors Condition (Intrapersonal, Interpersonal and Computer) and Block (One, Two and Three) for unsigned asynchronies can be seen in Figure 3.5 – Panel A. The Greenhouse-Geiser corrected ANOVA revealed a main effect of Condition, $F(1.455, 16.007) = 21.936$, $p < .001$, $\eta^2 = .409$, but no effect of Block, $F(1.300, 14.295) = 2.640$, $p = .120$, $\eta^2 = .026$, nor a significant interaction, $F(3.135, 34.481) = .233$, $p = .880$, $\eta^2 = .003$. The main effect of Condition was driven by the overall higher asynchrony in the Interpersonal condition, which was significantly different from the Intrapersonal, $t(11) = 10.142$, $p < .001$, $d = 2.928$, as well as the Computer condition, $t(11) = 4.133$, $p = .002$, $d = 1.193$. There was no difference between Intrapersonal and Computer condition, $t(11) = 1.366$, $p = .199$, $d = .394$. Post-hoc comparisons furthermore showed that the performance in the Intrapersonal condition in the first block ($M = 34$ ms, $SD = 12$ ms) and in the last block ($M = 29$ ms, $SD = 9$ ms) did not differ significantly, $t(11) = 1.234$, $p = .243$, $d =$

.356, nor has the performance in the Interpersonal condition ($M_{\text{First}} = 63$ ms, $SD_{\text{First}} = 18$ ms, $M_{\text{Last}} = 55$ ms, $SD_{\text{Last}} = 17$ ms) changed significantly, $t(11) = 1.260$, $p = .234$, $d = .364$. There was however a significant improvement in the Computer condition ($M_{\text{First}} = 40$ ms, $SD_{\text{First}} = 15$ ms, $M_{\text{Last}} = 35$ ms, $SD_{\text{Last}} = 13$ ms), $t(11) = 2.475$, $p = .031$, $d = .714$.

A similar ANOVA was computed for baseline-corrected unsigned asynchronies (see Figure 3.5 – Panel B). The Greenhouse-Geiser corrected ANOVA revealed a main effect of Condition, $F(1.170, 12.868) = 6.171$, $p = 0.024$, $\eta^2 = 0.180$, but no effect of Block, $F(1.300, 14.295) = 2.640$, $p = 0.120$, $\eta^2 = 0.009$, nor of interaction $F(3.135, 34.481) = 0.233$, $p = 0.880$, $\eta^2 = 0.001$. The baseline corrected values in the Computer condition were significantly lower than the ones in the Intrapersonal condition, $t(11) = 2.539$, $p = .028$, $d = .733$, and in the Interpersonal condition, $t(11) = 2.570$, $p = .026$, $d = .742$. There was no difference between the Intra- and the Interpersonal condition, $t(11) = .402$, $p = .695$, $d = .116$. Comparing the baseline-corrected asynchrony measures to zero in one-sample t-tests, showed that even in the last block participants did not reach the baseline performance, neither in the Intrapersonal condition $t(11) = 5.007$, $p < .001$, $d = 1.445$, nor in the Interpersonal condition $t(11) = 2.682$, $p = .021$, $d = .774$. In the Computer condition however, baseline was already reached in Block 1, as indicated by a one-sample t-test, $t(11) = .351$, $p = .732$, $d = .101$.

3.5.4 Variability in Baseline Trials

In non-delay baseline trials I observed the lowest variability in the Intrapersonal condition ($M = 9$ ms, $SD = 2$ ms), with higher variability in the Interpersonal condition ($M = 25$ ms, $SD = 6$ ms) and in the Computer condition ($M = 20$ ms, $SD = 5$ ms) (see Figure 3.3 – Panel D). I computed a one-way ANOVA with the three levels Intrapersonal, Interpersonal and Computer. The Greenhouse-Geiser corrected ANOVA revealed a significant effect for Condition, $F(1.783, 19.614) = 45.864$, $p < .001$, $\eta^2 = .686$. Post-hoc paired-samples t-tests

showed that this was due to the significant difference between the Intrapersonal condition and both the Interpersonal condition, $t(11) = 9.637, p < .001, d = 2.782$, and the Computer condition, $t(11) = 7.575, p < .001, d = 2.187$. The difference between the Interpersonal condition and the Computer condition was also significant, $t(11) = 2.489, p = .03, d = .718$.

3.5.5 Variability in Delay Trials

The results of a 3 x 3 ANOVA with the factors Condition (Intrapersonal, Interpersonal and Computer) and Block (One, Two, and Three) for variability of unsigned asynchrony can be seen in Figure 3.5 – Panel C. The Greenhouse-Geiser corrected ANOVA revealed only a significant main effect of Condition, $F(1.738, 19.118) = 23.467, p < .001, \eta^2 = .480$, but no significant main effect of Block, $F(1.290, 14.187) = .478, p = .548, \eta^2 = .005$, and no significant interaction $F(2.813, 30.942) = .798, p = .497, \eta^2 = .009$. Variability in the Intrapersonal condition was significantly less than variability in the Computer condition, $t(11) = 2.973, p = .013, d = .858$, which was, in turn, less variable than the Interpersonal condition, $t(11) = 3.542, p = .005, d = 1.023$. Post-hoc comparisons furthermore showed that the performance in the Intrapersonal condition in the first block ($M = 17$ ms, $SD = 3$ ms) and in the last block ($M = 17$ ms, $SD = 4$ ms) did not differ significantly, $t(11) = .552, p = .592, d = .159$. There was also no reduction of variability in the Interpersonal condition ($M_{\text{First}} = 37$ ms, $SD_{\text{First}} = 13$ ms, $M_{\text{Last}} = 37$ ms, $SD_{\text{Last}} = 10$ ms), $t(11) = .187, p = .855, d = .054$. Only variability in the Computer condition improved significantly ($M_{\text{First}} = 26$ ms, $SD_{\text{First}} = 8$ ms, $M_{\text{Last}} = 22$ ms, $SD_{\text{Last}} = 9$ ms), $t(11) = 2.806, p = .017, d = .810$.

A similar ANOVA was computed for baseline-corrected variability (see Figure 3.5 – Panel D). In contrast to Experiment 1, where we saw a main effect of block, the Greenhouse-Geiser corrected ANOVA for the second experiment revealed no significant results. Neither the main effect of Condition, $F(1.815, 19.967) = 2.426, p = .118, \eta^2 = .081$, nor the main effect

of Block, $F(1.29, 14.187) = .478$, $p = .548$, $\eta^2 = .003$, and the interaction, $F(2.813, 30.942) = .798$, $p = .497$, $\eta^2 = .006$, were significant. Comparing the baseline-corrected variability measures to zero in one-sample t-tests, showed that only in the last block of the Computer condition participants reached a level of performance that was not significantly different from baseline, $t(11) = .651$, $p = .528$, $d = .188$. Participants' performance in the last block was significantly different from baseline in the Intrapersonal condition, $t(11) = 4.751$, $p = .001$, $d = 1.371$, and in the Interpersonal condition, $t(11) = 4.140$, $p = 0.002$, $d = 1.195$.

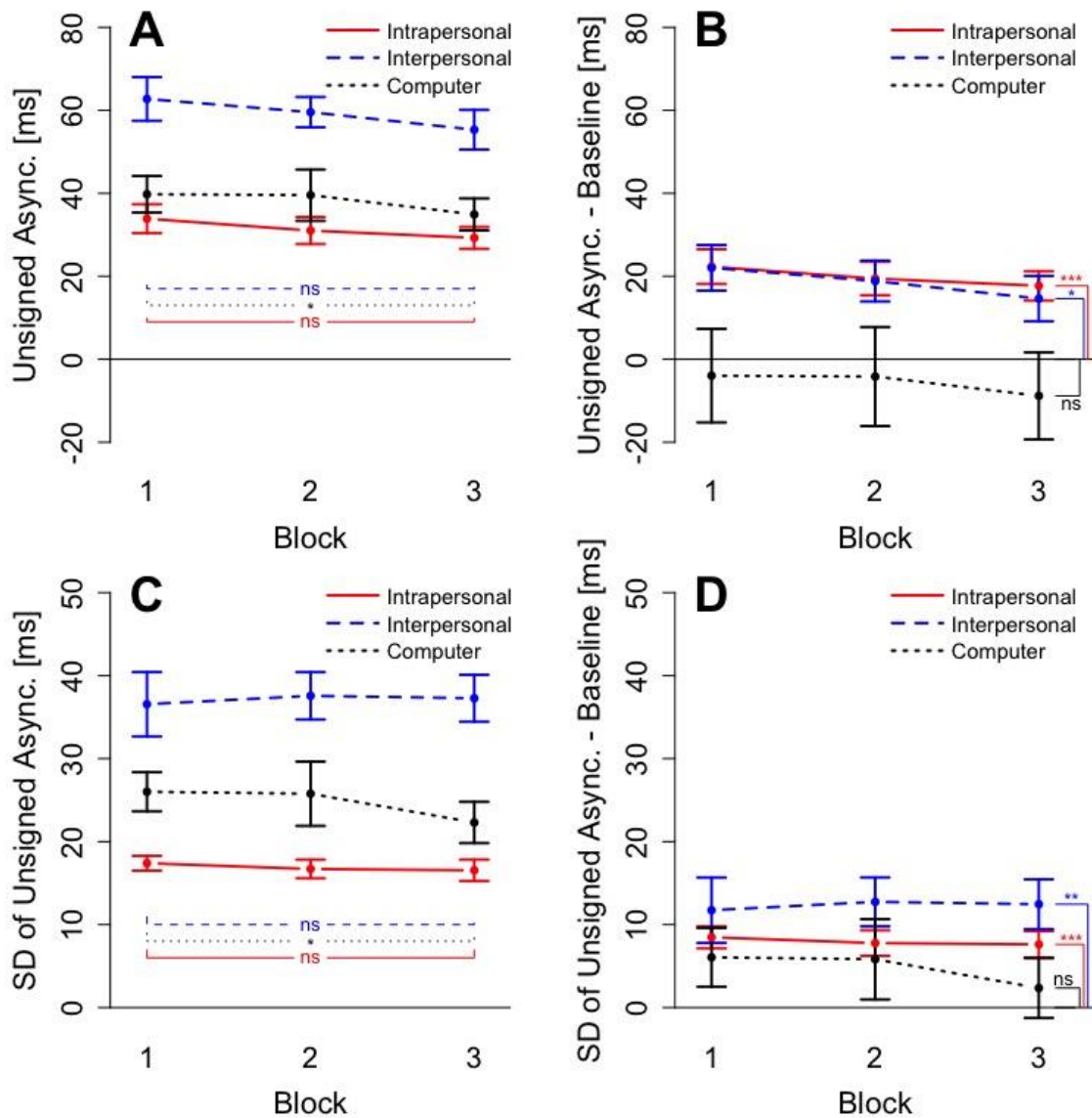


Figure 3.5: Results for delay trials of Experiment 1b. Panel A) shows the results for unsigned asynchrony. Panel B) also shows unsigned asynchronies but corrected for the baseline from non-delay trials. Panel C) shows performance in terms of variability

(SD) of unsigned asynchrony. Panel D) shows baseline-corrected variability. Error bars in all four panels show standard errors.

3.5.6 Unsigned Asynchrony and its Variability Combined

Table 3.1 offers a detailed overview over individual participants' performance. Three participants reached baseline performance in all three conditions. Five participants did so in two of the three conditions and four participants reached their baseline performance only in one condition.

3.6 Experiment 2 – Methods

3.6.1 Participants

For the third experiment, I invited another 12 musicians (5 women, 7 men, mean age = 27 years, SD = 6 years) to participate in the experiment. All 12 participants managed to produce 80% of the trials successfully. All participants in Experiment 2 had completed at least 5 years of private lessons on a musical instrument (M = 10 years, SD = 4 years).

3.6.2 Procedure and Design

Procedure and design were almost identical to Experiment 1. In Experiment 2, the actions participants had to perform were exactly the same as in Experiment 1. Only the outcomes that were produced by their actions were different. In Experiment 2 all participants' taps as well as the computer metronome produced piano sounds of different pitches that taken together generated a short polyphonic melody of ten beats (see Figure 3.1). Hence, action outcomes exhibited an additional structure of harmonic pitch. In short, in Experiment 2 the

actions were identical to those in Experiment 1, but the produced outcomes differed from those in Experiment 1 and 1b.

3.7 Experiment 2 – Results

3.7.1 Accuracy

The mean accuracy for the participants in Experiment 2 was 94.91 % (SD = 3.50 %). A Greenhouse-Geiser corrected one-way ANOVA with the Factor Condition revealed a significant main effect, $F(1.208, 13.289) = 4.753, p = .042, \eta^2 = .22$. Post-hoc comparisons revealed that this effect was due to accuracy being highest in the Computer condition (M = 98.31 %, SD = 2.51 %), and significantly different from both the accuracy in the Intrapersonal condition (M = 90.97 %, SD = 9.52 %), $t(11) = 2.623, p = .024, d = .757$, and the accuracy in the Interpersonal condition (M = 95.44 %, SD = 3.00 %), $t(11) = 2.737, p = .019, d = .79$. However, there was no significant difference between Intrapersonal and Interpersonal condition, $t(11) = 1.546, p = .15, d = .446$.

3.7.2 Asynchrony in Baseline Trials

In non-delay baseline trials in Experiment 2, I observed the lowest unsigned mean asynchrony for the Intrapersonal condition (M = 14 ms, SD = 3 ms), with comparable levels of asynchrony for Interpersonal (M = 37 ms, SD = 8 ms) and Computer condition (M = 40 ms, SD = 15 ms) (see Figure 3.3 – Panel E). I computed a one-way ANOVA with the factor condition (Intrapersonal, Interpersonal and Computer). The Greenhouse-Geiser corrected ANOVA revealed a significant effect for Condition, $F(1.772, 19.496) = 29.94, p < .001, \eta^2 = .603$. Post-hoc paired-samples t-tests showed that this was mainly due to the lower asynchrony in the Intrapersonal condition, as compared to the Interpersonal $t(11) = 7.563, p < .001, d =$

2.183, and the Computer condition, $t(11) = 6.205, p < .001, d = 1.791$. There was no significant difference between the Interpersonal and the Computer condition, $t(11) = .828, p = .425, d = .239$.

3.7.3 Asynchrony in Delay Trials

Unsigned asynchronies are displayed in Figure 3.6 – Panel A. A Greenhouse-Geiser corrected 3 x 3 ANOVA with the factors Condition (Intrapersonal, Interpersonal and Computer) and Block (1, 2, and 3) revealed only a main effect of Block, $F(1.242, 13.666) = 6.71, p = .017, \eta^2 = .046$, but no effect of Condition, $F(1.712, 18.827) = 2.894, p = .087, \eta^2 = .12$, and no significant interaction effect, $F(2.784, 30.626) = 2.165, p = .116, \eta^2 = .018$. Post-hoc comparisons revealed that in contrast to the results in Experiment 1 and 1b, in Experiment 2 the performance in the Intrapersonal condition improved significantly between the first block ($M = 49$ ms, $SD = 17$ ms) and the last block ($M = 37$ ms, $SD = 17$ ms), $t(11) = 2.325, p = .040, d = .671$. The performance in the Interpersonal condition also improved significantly ($M_{\text{First}} = 57$ ms, $SD_{\text{First}} = 18$ ms, $M_{\text{Last}} = 46$ ms, $SD_{\text{Last}} = 15$ ms), $t(11) = 2.679, p = .021, d = .773$. There was however no significant improvement in the Computer condition ($M_{\text{First}} = 40$ ms, $SD_{\text{First}} = 21$ ms, $M_{\text{Last}} = 37$ ms, $SD_{\text{Last}} = 16$ ms), $t(11) = 1.448, p = .176, d = .418$.

A similar ANOVA was computed for baseline-corrected unsigned asynchronies (see Figure 3.6 – Panel B). The Greenhouse-Geiser corrected ANOVA revealed a significant main effect of Condition, $F(1.652, 18.169) = 6.998, p = .008, \eta^2 = .252$, and a significant main effect of Block, $F(1.242, 13.666) = 6.71, p = .017, \eta^2 = .030$. The interaction, however, was not significant, $F(2.784, 30.626) = 2.165, p = .116, \eta^2 = .012$. Post-hoc t-tests showed that the baseline-corrected performance in the Interpersonal condition was neither significantly different from the Intrapersonal condition, $t(11) = 2.043, p = .066, d = .59$, nor from the Computer condition, $t(11) = 2.083, p = .061, d = .601$. Asynchronies in the Intrapersonal

condition, however, were significantly higher than in the Computer condition, $t(11) = 3.224$, $p = .008$, $d = .931$. Comparing the baseline-corrected values to zero in one-sample t-tests, showed that, as in Experiment 1, participants reached baseline performance in the last block of the Interpersonal condition, $t(11) = 1.643$, $p = 0.129$, $d = .474$, and already in the first block in the Computer condition, $t(11) = .071$, $p = 0.945$, $d = .020$. However, in the last block of the Intrapersonal condition, performance was still significantly higher than baseline, $t(11) = 5.232$, $p < .001$, $d = 1.51$.

3.7.4 Variability in Baseline Trials

In non-delay baseline trials I observed the lowest variability in the Intrapersonal condition ($M = 10$ ms, $SD = 3$ ms), with higher variability in the Interpersonal condition ($M = 25$ ms, $SD = 4$ ms) and in the Computer condition ($M = 21$ ms, $SD = 6$ ms) (see Figure 3.3 – Panel F). I computed a one-way ANOVA with the factor Condition (Intrapersonal, Interpersonal and Computer). The Greenhouse-Geiser corrected ANOVA revealed a significant effect for Condition, $F(1.937, 21.304) = 37.596$, $p < .001$, $\eta^2 = .658$. Post-hoc paired-samples t-tests showed that this was due to the significant difference between the Intrapersonal condition and both the Interpersonal condition, $t(11) = 8.145$, $p < .001$, $d = 2.351$, and the Computer condition, $t(11) = 5.892$, $p < .001$, $d = 1.701$, and a significant difference between the Interpersonal condition and the Computer condition, $t(11) = 2.273$, $p = .044$, $d = .656$.

3.7.5 Variability in Delay Trials

I computed a 3 x 3 ANOVA with the factors Condition (Intrapersonal, Interpersonal and Computer) and Block (One, Two, and Three). The corresponding plot can be seen in Figure 3.6 – Panel C. The Greenhouse-Geiser corrected ANOVA revealed a significant main effect of condition, $F(1.818, 20.000) = 21.282$, $p < .001$, $\eta^2 = .450$, and a significant main effect of block,

$F(1.950, 21.451) = 5.737, p = .011, \eta^2 = .038$. The interaction was not significant, $F(2.193, 24.120) = .757, p = .491, \eta^2 = .012$. The main effect of condition was driven by the variability in the Intrapersonal condition being significantly lower than the variability in the Computer condition, $t(11) = 2.601, p = .025, d = .751$, which in turn was significantly lower than the variability in the Interpersonal condition, $t(11) = 7.18, p < .001, d = 2.073$. Post-hoc comparisons revealed furthermore that the performance in the Intrapersonal condition in the first block ($M = 20$ ms, $SD = 4$ ms) and in the last block ($M = 18$ ms, $SD = 5$ ms) did not differ significantly, $t(11) = 2.020, p = .068, d = .583$. However, there was significant reduction of variability in the Interpersonal condition ($M_{\text{First}} = 33$ ms, $SD_{\text{First}} = 6$ ms, $M_{\text{Last}} = 28$ ms, $SD_{\text{Last}} = 6$ ms), $t(11) = 2.480, p = .031, d = .716$. There was no significant reduction in the Computer condition ($M_{\text{First}} = 24$ ms, $SD_{\text{First}} = 6$ ms, $M_{\text{Last}} = 23$ ms, $SD_{\text{Last}} = 6$ ms), $t(11) = .619, p = .548, d = .179$.

A similar ANOVA was computed for baseline-corrected variability (see Figure 3.6 – Panel D). The Greenhouse-Geiser corrected ANOVA for the third experiment revealed a significant main effect of block, $F(1.950, 21.451) = 5.737, p = .011, \eta^2 = .024$, but not of condition, $F(1.746, 19.202) = 2.810, p = .091, \eta^2 = .099$. The interaction was also not significant, $F(2.193, 24.120) = .757, p = .491, \eta^2 = .008$. The main effect of Block was driven by the difference between the Block One and Block Three in the Interpersonal condition, $t(11) = 2.48, p = .031, d = .716$. The same comparison between Block One and Block Three, however, showed no differences in the Intrapersonal condition, $t(11) = 2.020, p = .068, d = .583$, nor in the Computer condition, $t(11) = .619, p = .548, d = .179$. Comparing the baseline-corrected variability measures to zero in one-sample t-tests, showed that even in the last block participants had a performance significantly worse than baseline, in the Intrapersonal condition, $t(11) = 5.461, p < .001, d = 1.576$. However, the performance in the last block of the

Interpersonal condition, $t(11) = 2.179$, $p = 0.052$, $d = .629$ and of the Computer condition, $t(11) = 1.190$, $p = 0.259$, $d = .344$, was not significantly different from baseline performance.

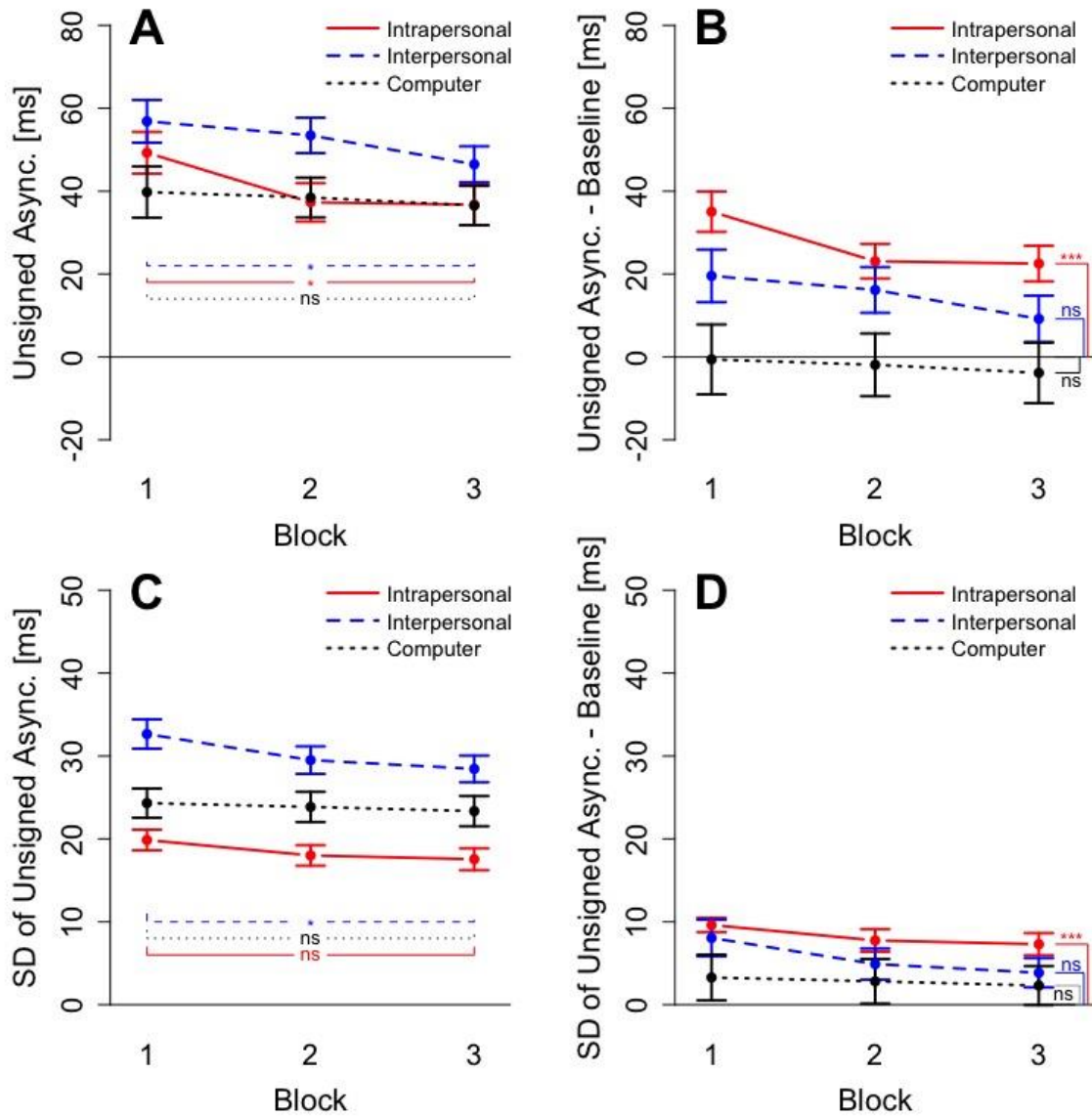


Figure 3.6: Results for delay trials of Experiment 2. Panel A) shows the results for unsigned asynchrony. Panel B) also shows unsigned asynchronies but corrected for the baseline from non-delay trials. Panel C) shows performance in terms of variability (SD) of unsigned asynchrony. Panel D) shows baseline-corrected variability. Error bars in all four panels show standard errors.

3.7.6 Unsigned Asynchrony and its Variability Combined

In Experiment 2, there was no participant who reached their baseline performance in all three conditions (see Table 3.1). Nine participants reached their baseline in two conditions and three participants only in one condition.

3.8 Cross-Experiment Comparison

3.8.1 Adaptation performance

In Experiment 1 participants reached their baseline performance on average in 2.5 of 6 bins in the Interpersonal condition and on average 3.6 of 6 bins in the Computer condition. In Experiment 2, these numbers are 2.9 of 6 bins and 3.8 of 6 bins respectively. A Greenhouse-Geiser corrected 3 x 2 ANOVA on this data with the within-subjects factor Condition (Intrapersonal, Interpersonal, Computer) and the between-subjects factor Experiment (1 and 2) revealed a main effect for Condition, $F(1.918, 42.186) = 24.507, p < .001, \eta^2 = .357$, but no effects for Experiment, $F(1, 22) = .020, p = .888, \eta^2 = .000$, and no significant interaction, $F(1.918, 42.186) = .288, p = 0.742, \eta^2 = .006$.

Experiment 1				Experiment 1b				Experiment 2			
ID	Intra-personal	Inter-personal	Computer	ID	Intra-personal	Inter-personal	Computer	ID	Intra-personal	Inter-personal	Computer
1	0	4	6	1	0	1	2	1	0	4	0
2	0	0	4	2	6	1	0	2	0	5	6
3	0	3	4	3	0	3	0	3	0	6	6
4	0	1	6	4	0	2	6	4	1	1	0
5	0	0	1	5	1	2	6	5	0	6	6
6	0	0	0	6	3	6	6	6	0	0	5
7	1	6	4	7	0	3	3	7	0	2	6
8	1	6	2	8	0	2	4	8	0	6	5
9	3	6	6	9	0	0	4	9	0	3	4
10	0	0	5	10	1	0	0	10	0	1	0
11	0	1	1	11	0	0	2	11	0	0	1
12	0	3	4	12	2	6	6	12	0	1	6
	5	30	43		13	26	39		1	35	45

Table 3.1: Shown are the number of bins (each bin was 10 trials / half a block) in which individual participants' performance was comparable to their non-delay baseline performance, both, in terms of absolute asynchronies and its variability. There were 6 bins in each condition. The last row shows the column sums. Note that whereas in the group statistics participants never reached the baseline performance in the Intrapersonal condition, this table shows that there were six participants who managed to do so in all conditions, including the Intrapersonal condition (green fill color/dashed frame lines).

3.8.2 Questionnaire Data

All participants provided us with some biographic data in a post-experiment questionnaire. Besides standard items such as age and handedness, I asked which instruments they played, for how long they had training on them, whether they had experience playing in ensembles or teaching their instrument. To see whether any of these items are good predictors for their performance in the adaptation task, I entered them in a linear regression model as predictors for how many delayed bins were performed at baseline levels. I started with a model that contained the following four predictors: *age*, *main training* (the number of years they had received training on their main instrument), *total training* (the number of years they had received training on any instrument aggregated over all instruments they had provided us with data, e.g. for a participant that had 19 years of training on the piano and 9 years on the flute, *main training* was coded as 19, whereas *total training* was coded as 28) and *practice per week* (in hours). I systematically removed the factor with the highest p-value until only significant predictors were left (Field, Miles & Field, 2012). This left us with the two significant predictors *age* and *main training*. Whereas an increase in *main training* predicted better performance, an increase in *age* predicted worse performance (see Table 3.2 for further details). As assumptions of normality were not met, I validated this model with a boot strapping procedure (Field, Miles & Field, 2012) that confirmed the significant effects of *age* and *main training*.

	Model Stats	Estimate	Standard Error	p-value	95 % Confidence Intervals		95 % Confidence Intervals (Boot strapping)	
<i>intercept</i>	R ² = 0.194 (adjusted)	9.657	3.573	0.011	2.388	16.925	2.425	16.020
<i>age</i>		-0.310	0.152	0.050	-0.620	0.000	-0.531	-0.017
<i>main training</i>	p = .011	0.439	0.139	0.003	0.156	0.722	0.151	0.678

Table 2.2: Shown are detailed results for the linear model of the following formula: successful bins ~ age + main training + error. For the bootstrapping procedure I used 2000 iterations. Age was coded as the age of participants at the time of the experiment. Main training was coded as the amount of years each participant had on their main instrument.

3.9 Discussion

In this study I investigated the effects of coordination setting and structure of action outcomes on musicians' ability to adapt to extraordinary coordination patterns. First, I tested conditions implying various degrees of coupling strength between two movements. For non-delay baseline performance, I found a clear pattern that repeated across all three experiments. Unsigned asynchrony was significantly lower and more stable in the Intrapersonal condition than in the Interpersonal condition. This is in line with findings from studies that compare intrapersonal and interpersonal coordination (Knoblich & Jordan, 2003) and probably due to the lack of shared internal processes (Knoblich & Jordan, 2003) and the beneficial effects of stronger coupling on in-phase coordination (Kay et al., 1987; Kelso, 1984, Richardson et al., 2007). In the unidirectional coupling condition where participants performed with a computer they exhibited unsigned asynchronies to the same degree as in the interpersonal condition. Konvalinka and colleagues (2010) argued on the basis of similar results that a less predictable, but responsive partner facilitates synchronization just as much as a perfectly predictable, but un-responsive computer. In the current study I found however, that variability of asynchrony was consistently better in the Computer condition than in the Interpersonal condition. This suggests that the asynchrony in the Computer condition arose to some extent from a more stable type of coordination, most likely the well-established negative mean asynchrony (see Repp, 2005).

Whereas participants' baseline performance was best in the Intrapersonal condition, I predicted that the weaker coupling in the Interpersonal condition and the Computer condition should be advantageous for the performance of extraordinary coordination patterns. In terms of unsigned asynchrony, for which we saw a clear advantage of stronger coupling during in-phase coordination, we find that this advantage disappeared during the performance of an extraordinary coordination pattern. In Experiment 1, asynchrony was lowest in the Computer

condition and equal in the Intrapersonal and Interpersonal condition. In Experiment 2, I found no differences between the three coordination conditions. In Experiment 1, the improvement of asynchrony performance from the first to the last block was not significant in the Intrapersonal condition, whereas it was significant for the Interpersonal condition. For asynchrony in the Interpersonal condition, on the other hand, we see significant improvement across blocks in Experiment 1 and Experiment 2. The effects of an extraordinary coordination pattern in the different coordination conditions is clearly visible in comparison to the baseline. In the Computer condition, where the coupling is just unidirectional, we saw that participants managed to reach the same performance as in their baseline already in Block 1 in all three experiments. In the Interpersonal condition, participants initially performed significantly worse than baseline. However, thanks to continuous adaptation across the three blocks, participants eventually reached their baseline performance in Experiment 1 and 2. In the Intrapersonal condition participants never reached their baseline performance. Hence, in terms of unsigned asynchrony, I conclude that stronger intrapersonal coupling impedes coordination of and adaptation to extraordinary patterns.

While in absolute terms the variability of asynchrony in delay trials was lowest in the Intrapersonal condition, baseline corrected variability showed that participants were on average not able to reach their baseline performance. In contrast, in the Interpersonal condition baseline performance was reached in Experiment 1 and Experiment 2. Taking together asynchrony and its variability, I found that during the Intrapersonal Condition, the condition with the strongest coupling, the group average never reached the baseline performance, even though we saw significant reduction in variability for Experiment 1 and significant reduction in unsigned asynchrony for Experiment 2. It is interesting to note that the Intrapersonal condition was also the condition with the consistently lowest variability. Individual motor learning studies have shown that variability is advantageous for motor learning (Ölveczky, Andalman & Fee, 2005;

Wu, Miyamoto, Castro, Ölveczky & Smith, 2014). This leads to the question of whether adaptation in the Interpersonal condition was facilitated by the higher variability in that condition. Whether high variability in joint tasks leads to better motor learning is an important question for research on coordination in joint action and is currently under investigation by Sabu and colleagues (Sabu, Curioni, Vesper, Sebanz & Knoblich, in prep.).

In Experiment 1b, I introduced bimanual actions to the Interpersonal and the Computer condition in an effort to reduce the variability and therefore also improve the corresponding baseline performance based on findings in the literature that bimanual actions are more stable in time than unimanual ones (Drewing & Aschersleben, 2003; Helmuth & Ivry, 1996). Ideally, this would have led to more similar baseline performances across the coordination conditions. However, our results showed that baselines were not reduced in Experiment 1b, neither in the Interpersonal nor in the Computer condition (see Figure 3.3). Hence this attempt failed. Interestingly, though, instead of reducing variability in the baseline performance we see reduced adaptation to the extraordinary coordination pattern. In fact, in the Interpersonal condition in Experiment 1b we see no significant adaptation, to the effect that participants' performance in Block 3 was still significantly worse than their baseline performance. This somewhat surprising finding might be due to the way bimanual actions were performed in the Interpersonal and the Computer condition. Since a Max patch randomly allocated the triggering of the sounds to one of the two hands, it might not have been obvious which hand produced the action effect. While this apparently disrupted adaptation in the Interpersonal condition, it did not seem to affect performance in the Computer condition to the same extent.

Taken together, the results discussed so far show that reducing the coupling strength between interacting limbs by the means of splitting a task across two people might facilitate the production of otherwise unstable coordination patterns. This principle is reminding of Ugandan xylophone players who manage to produce extraordinary coordination patterns at

incredibly high tempi by distributing them across two musicians while keeping in-phase actions within musicians.

In cultural evolution where intending patterns, retaining them and reproducing them is limited by cognitive constraints (Miton, Claidière & Mercier, 2015; Miton & Mercier, 2015) and possibly also by physical constraints (Savage, Tierney & Patel, 2017), copying errors should lead to a convergence towards in-phase patterns over time (Scott-Phillips, Blanke & Heintz, 2018). To use the terminology of the cultural attraction theory, in-phase dominance (see e.g. Repp, London & Keller, 2012) could be a factor of attraction and if so, it would be a global factor, not limited to a certain population or time (Scott-Phillips et al., 2018). However, instead of a convergence, we see a variety of musical contexts that deviate from in-phase coordination patterns. Recent transmission chain experiments showed that one-part rhythmic patterns converged to exhibit structural organization that was based on small-integer ratios (Ravignani, Delgado & Kirby, 2016). Whether two-part rhythmic patterns would also converge on small-integer ratios seems likely but remains to be tested empirically. This is also the case for possible differences between transmission chains in which two-part patterns are produced bimanually and chains in which patterns are produced jointly. Hence, the question of whether intricate rhythms in the evolution of music initially emerged out of joint music-making rather than solo music-making remains to be answered by further research.

As a second factor I investigated the effects of structure realized along the harmonic pitch dimension on the adaptation to extraordinary coordination patterns. I predicted better adaptation in Experiment 2, where the auditory outcome was melodically enhanced, than in Experiment 1. In terms of how often participants accomplished performance in delay trials that was comparable to their baseline, the data exhibited the effect of coordination setting, but no differences between the results of the two experiments. Hence, the melodic enhancement of the auditory outcome did not lead to significantly better adaptation.

Given that musicians frequently encounter extraordinary coordination patterns, I had expected the invited musicians to more readily adapt to the pattern than the data suggested. As mentioned above, our subjects had on average around 11 years of musical training. Maybe this is not enough training to build the necessary skills to deploy similar strategies like the third player in Amadinda music, who has to be able to use structured action effects, on which to map the necessary motor commands. However, the data show that at least some participants were in fact able to adapt to the strange phase relation in all conditions, especially in the Computer and the Interpersonal condition, but also in the Intrapersonal condition, where coupling was strongest.

In the Intrapersonal condition of Experiment 1, three participants reached a delay performance that was not significantly different from their non-delay baseline, both in terms of unsigned asynchrony as well as its variability. In the Intrapersonal condition in Experiment 1b, there were 5 participants who reached a delay performance that was not significantly different from their baseline performance. In the Intrapersonal condition in Experiment 2, only one participant managed to reach their baseline performance. Across all three experiments, there were six participants who reached the baseline performance within the 3 blocks in all three conditions, including the Intrapersonal condition for which the group average never reached the baseline. One participant did not manage to reach their baseline performance in any of the coordination conditions.

To identify possible predictors of success in terms of reaching one's baseline performance, I used a linear model procedure on collected questionnaire data, from which two significant predictors emerged: age and years of training on the main instrument. *Years of training* was positively correlated with success, which suggests that experience as a musician improves the ability to adapt to extraordinary phase relations. Age, however, was negatively correlated with success. The fact that the model accounted for less than 20 % of the variability

suggests that there are other factors that contribute to the ability to adapt to extraordinary coordination patterns. The effect of years of training could for example be mediated by instrument type and type of training. Some instruments might require more temporal flexibility than other instruments, such as drums and piano, on which playing different rhythmic patterns at the same time is frequent. Experience in mixed ensembles with instruments that exhibit different rise times might also train the ability to adapt to extraordinary coordination patterns.

It might also be noteworthy that participant 9 in Experiment 1, who adapted well to all three conditions, had eight years of experience as an organ player. Gould and Helder (1970) reported that the only person that was able to speak coherently in a speech delay experiment that they had conducted was a professional organist. As mentioned in the introduction, pipe organs can exhibit pitch-dependent delays of up to 150 ms. This could mean that some organists are especially trained in flexibly mapping different delays to different finger actions in order to align tone onsets.

For further research on extraordinary coordination patterns in joint actions it could be interesting to identify the skills or strategies that allowed some participants to quickly adapt to the phase shift. Following up on this it would be interesting to investigate whether these skills or strategies are transferable to other extraordinary coordination patterns and also whether non-musicians would be able to learn these skills and/or strategies isolated from musical training. Another possible avenue for further research in this direction could lead towards an understanding of the influence of joint music-making as a tool to avoid certain attractor states, such as in-phase coordination in the transmission of culture.

4 Study 3: The Bug Within: Why Do People Rush during Rhythmic Interactions?

4.1 Introduction

When humans engage in synchronized, rhythmic joint activities, they tend to increase their pace unconsciously. Even though this phenomenon appears to be ubiquitous and well known among musicians, dancers and their audiences (Barnes, 2008; bvdd, 2001; Colson, 2012; Dillard, 2013; McNamara, 2017), it has hardly been addressed in research on timing mechanisms in humans (for two recent exceptions see Okano, Shinya & Kudo, 2017; and Thomson, Murphy & Lukeman, 2018). Thus, despite the rich literature on the cognitive and neural bases of rhythmic performance in humans (for reviews see Collier & Ogden, 2004; Repp & Su, 2013), the psychological mechanisms underlying *joint rushing* are still unknown.

We argue that if joint rushing indeed emerges from human interaction in contrast to purely individual processes, then predominant models of inter-subjective sensorimotor synchronization are incomplete. Specifically, they do not consider how sounds produced during synchronous, rhythmic joint actions are integrated to result in specific tempo drifts. Thus, the systematic study of joint rushing has the potential to lead to more adequate models of inter-personal coordination by identifying missing components. In this chapter we first present evidence that joint rushing is a robust phenomenon that emerges in interpersonal sensorimotor synchronization. As an explanation for joint rushing we then propose a model that combines a period correction mechanism, which is a standard component of models of human sensorimotor synchronization (Van der Steen, Jacoby, Fairhurst & Keller, 2015), with a phase advance mechanism, an evolutionarily ancient mechanism that has been proposed to

regulate rhythmic synchronicity in some firefly species and chorusing insects (Greenfield, 2005).

4.1.1 Previous research

Timing mechanisms in humans have been studied extensively with sensorimotor synchronization paradigms, yet tempo drift in time series has usually been considered to be a “methodological inconvenience” (Madison, 2001) necessitating techniques for minimizing or eliminating it (Collier & Ogden, 2004). It is not surprising, then, that current models of interpersonal sensorimotor synchronization in humans do not account for tempo drift and can thus not explain why groups engaging in rhythmic joint activities tend to increase tempo over time. Two indications that this is a robust phenomenon come from recent studies by Okano et al. (2017) and Thomson et al. (2018). These studies rule out several potential explanations for joint rushing. Okano et al. (2017) provided evidence against the hypothesis that joint rushing simply emerges from the faster tapper in a pair acting as a leader and thereby setting a faster tempo for the interaction. Thomson et al. (2018) investigated the role of negative mean asynchrony (Repp & Su, 2013), i.e. the tendency to tap too early when synchronizing with an external timekeeper, without being aware of it. They concluded that the evidence speaks against negative mean asynchrony as the cause of joint rushing.

A further explanation in musicians’ discussions of joint speeding is that increased arousal causes the tempo increase in group performance. Thus, joint rushing may be an instance of social facilitation (see Aiello & Douthitt, 2001). Zajonc’s (1965) arousal-based theory of social facilitation states that the mere presence of a conspecific increases the level of arousal and thereby facilitates dominant responses, leading to an increase in performance speed (i.e. audience effects). Such arousal-based theories of social facilitation appear to correspond well with the experience of musicians and dancers that joint rushing is more pronounced during an

exciting performance than during rehearsal (bvdd, 2001; Colson, 2012). Another possible explanation comes from the literature on mirroring accounts, in which facilitating effects of mirrored actions are reported (Brass, Bekkering & Prinz, 2001). These effects could lead to joint rushing, when co-actors mirror each other's actions.

4.1.2 Present study

We propose that joint rushing arises from the concurrent activity of two separate mechanisms: 1) The phase advance mechanism, an evolutionarily old mechanism that shortens single intervals and thereby brings about synchrony, and 2) a human-specific adaptive period correction mechanism that translates the shortened period of individual intervals into a tempo change by adjusting internal timekeepers.

The phase advance mechanism was first proposed in research on synchronously flashing fireflies and chorusing insects (Blair, 1915; Buck 1988; Greenfield, 1994a; Greenfield, 1994b; Greenfield, 2005; Richmond, 1930). It assumes an oscillating timekeeper that initiates a signal whenever a signaling threshold is reached. Its eponymous characteristic however, is that it corrects phase differences between neighboring signalers by reducing the time it takes the lagging signaler to reach its signaling threshold (see Figure 4.1), i.e., by advancing the phase of the lagging signaler. This advancement leads to a single shortened period. For the advancement to occur, a conspecific's signal has to fall within a *sensitive temporal window*. This window is partially defined by the fact that some time passes between the oscillator reaching its threshold and the signal being broadcasted (see Figure 4.1). Perfectly aligned signals would therefore never fall into the sensitive window and therefore never trigger the mechanism. Furthermore, experiments in several species have shown that mechanisms like the phase-advance mechanism are tuned to particular frequencies or rhythms, dependent on the agents' own signals to avoid interference from signals emitted by other species (Alexander &

Moore, 1958; Walker, 1957). Even though this mechanism is biased towards shortening periods, the single, shortened periods it produces cannot account for a continuous tempo increase.

In order to explain how the phase advance mechanism can lead to a continuous tempo increase and thereby can result in joint rushing a second component is needed. We assume that this role is played by a concurrently active period correcting mechanism (Mates, 1992; Repp & Keller, 2004; Schulze, Cordes & Vorberg, 2005). This mechanism picks up on the temporal differences coming from one-shot phase advancements and then leads to a shortening of all consequent intervals. If a period correction mechanism changes the cycle duration of internal timekeepers in interacting partners, i.e. the base-rate of the underlying oscillations, this would ensure that each time a sufficiently extensive shortening occurs as a result of phase advancement, period correction results in a permanent increase in tempo. The adjustments have to be based on a partner's rate changes. Possible candidates for period corrections of this kind are human-specific adaptation and anticipation mechanisms that govern interpersonal sensorimotor synchronization (Van der Steen & Keller, 2013; Van der Steen, Jacoby, Fairhurst & Keller, 2015). We propose that this combination of an evolutionarily ancient mechanism with a more sophisticated and human-specific mechanism could account for *joint rushing*. In order to test this proposal, we conducted three experiments using a simple synchronization-continuation task performed by small groups of participants.

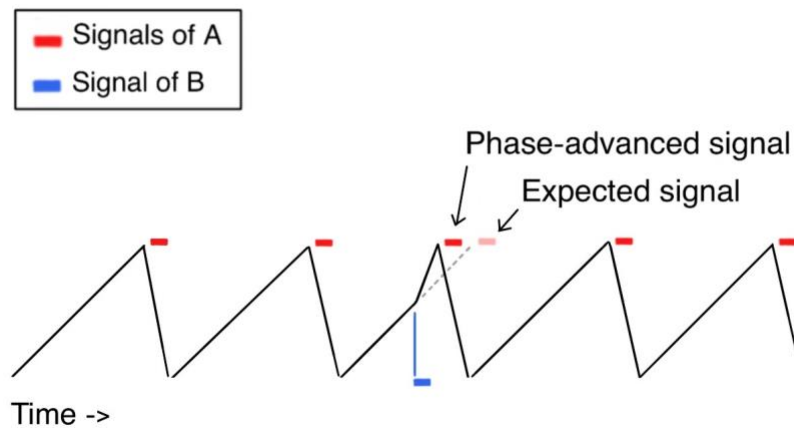


Figure 4.1: The phase advance mechanism is based on an oscillating timekeeper. Whenever a certain level of propensity to signal is reached a signal is initiated, the oscillator is reset, and the signal is broadcasted. If a conspecific's signal falls into a certain temporal window before the threshold is reached the current oscillation is accelerated (see Greenfield, 2005).

4.2 Experiment 1

4.2.1 Methods

In Experiment 1 we investigated whether rushing occurs specifically in joint performance, and if so, whether the presence of another (passive) person is enough to elicit rushing (an explanation based on social facilitation), or whether rushing is indeed contingent on the interaction with a co-actor. To do so, we asked participants to hit a drum pad to perform a synchronization-continuation task in which a leading metronome fades away and participants try to maintain the tempo either in an Individual or a Joint setting. Furthermore, we tested two potential factors that should foster joint rushing by increasing the chance of a signal falling into the sensitive window. First, we hypothesized that acting in larger groups (groups of three people) would elicit more rushing than acting in smaller groups (dyads) due to an increased chance of any co-actors' signal falling into a partner's sensitive window. Second, if a phase advance mechanism is involved in causing joint rushing, increasing individuals' variability

should also increase the chance of a signal falling into a partner's sensitive window and should thereby result in a larger tempo increase. We used a manipulation of target force as lower force should lead to higher variability (Sternad, Dean & Newell, 2000).

4.2.2 Analysis

We calculated synchronization indices to investigate whether participants were able to follow the instructions of synchronizing with each other. A one-sample t-test showed that participants' synchronization indices were significantly higher than a synchronization threshold of 0.73 (see Tognoli, Lagarde, DeGuzman & Kelso, 2007), $t(23) = 7.685$, $p < 0.001$, $d = 1.569$, with a mean of 0.86 and a standard deviation of 0.08. A further manipulation check showed that participants indeed hit the drum pads with a low velocity in the low force condition ($M = 25$, $SD = 3$) and a high velocity in the high force condition ($M = 113$, $SD = 12$). A Welch t-test showed that this difference was statistically significant: $t(23) = 41.092$, $p < .001$, $d = 10.384$ (Delacre, Lakens & Leys, 2017).

Tempo change was calculated in such a way that negative tempo change stands for shorter inter-response intervals (IRIs) in the end than in the beginning, which in turn stands for a tempo increase. In line with our prediction that the tempo should increase to a greater extent in joint performances, i.e. IRIs should become shorter, participants showed a more pronounced, negative tempo change in the Individual condition ($M = -18$ ms, $SD = 22$ ms) than in the Individual condition ($M = -2$ ms, $SD = 35$ ms), see Figure 4.2A. A Welch t-test revealed that this difference was significant, $t(23) = 3.170$, $p = .004$, $d = .59$ (Delacre et al., 2017). One sample t-tests showed that the tempo increase in the Individual condition was not significantly different from zero, $t(23) = .277$, $p = .784$, whereas the tempo change in the joint condition was different from zero, $t(23) = 3.972$, $p < .001$, $d = .811$. Thus, as expected, rushing occurred to a larger extent in the joint condition and was, in fact, absent during the Individual condition. As

there was always an experimenter present, we can conclude that the mere presence of another person was not sufficient to cause rushing. Figure 4.2B shows the data segmented into bins of 10 seconds to depict the average development of the tempo over the course of a trial split for joint and Individual condition.

4.2.3 Results

Group Size had no significant effect on tempo change (see Figure 4.2C). A two-by-two mixed ANOVA with the within factor Task (Individual or Joint) and the between factor Group Size (Two or Three) revealed only a main effect for Task, $F(1, 22) = 8.487, p = .008, \eta^2 = .084$, but no significant effect for Group Size, $F(1, 22) = .630, p = .436$, and no significant interaction, $F(1, 22) < .001, p = .998$. These results indicate that joint rushing occurred to the same extent in groups of two and groups of three.

In order to check whether higher target force led to higher temporal variability as intended by our manipulation, we calculated variability as squared residuals from a fitted linear model to account for potential drift. We replicated the findings of Sternad and colleagues (2000) in the Individual condition. When participants were drumming individually, lower target force indeed resulted in higher amounts of residuals ($M = 24.64, SD = 6.61$) than higher target force ($M = 20.69, SD = 5.21$), $t(23) = 3.281, p = .003, d = 0.664$. However, in the joint condition there was no significant difference in temporal variability between lower target force ($M = 26.11, SD = 6.77$) and higher target force ($M = 26.11, SD = 11.02$), $t(23) = .001, p = .999$. In both joint conditions variability was higher than in the Individual condition with the higher variability. Therefore, there were no significant effects of the force manipulation on joint rushing. A 2 x 2 ANOVA with the within subjects factors Task (Individual / Joint) and Target Force (high / low) showed a significant main effect of Task, $F(1, 23) = 8.873, p = .007, \eta^2 = .062$, but no significant main effect of Target Force, $F(1, 23) = .381, p = .543$, and no

significant interaction, $F(1, 23) = 2.45$, $p = .131$, (see Figure 4.2D). The lack of significant effects of Force are most likely due to the fact that the force manipulation did not affect temporal variability as intended in the joint condition.

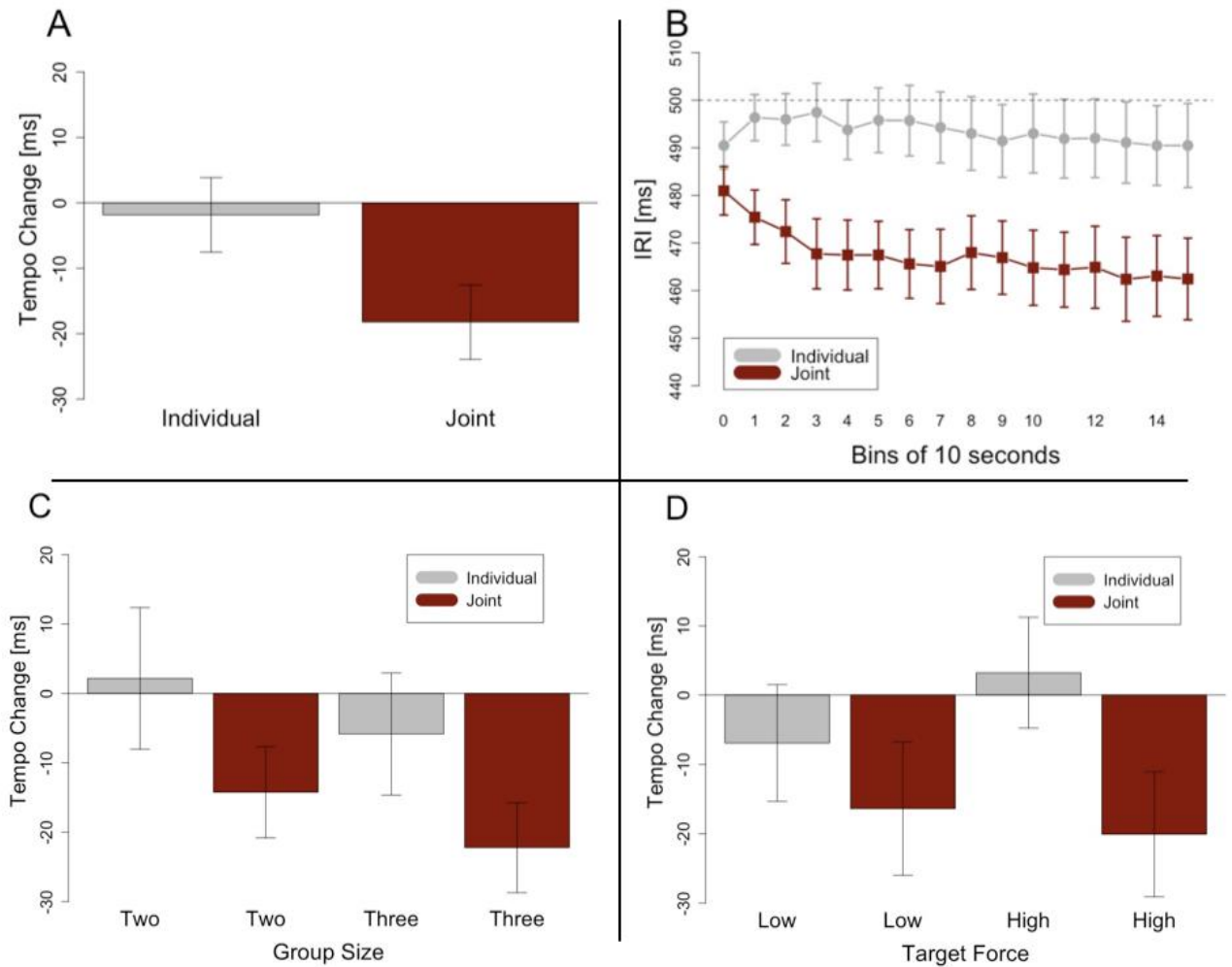


Figure 4.2: Results of Experiment 1. Tempo change indicates the difference between inter-response intervals at the end of a trial and at the beginning of a trial. Negative tempo change stands for a tempo increase. Error bars in A and D are calculated following the procedure recommended for within-subjects designs by Cousineau (2005). Error bars in B and C show standard errors. A) Tempo change results for the individual and the joint condition. The higher negative tempo change in the joint condition is significantly different from zero. B) Shown is the development of inter-response intervals over the course of trials, averaged over all participants and segmented into bins of 10 seconds each. C) The panel shows tempo change for Task and Group Size. Only the main effect for Task was significant. D) The panel shows tempo change for Task and Target Force. Only the main effect of Task was significant.

4.3 Experiment 2

4.3.1 Methods

In Experiment 2, we tested whether producing same outcomes (i.e., the same pitched tones) leads to more rushing than when participants produce different outcomes (tones with different pitches). This is predicted by a phase advance mechanism that is tuned to respond better, i.e. more strongly, to same outcomes than to different outcomes. In chorusing insects this phenomenon is believed to reduce cross-species interference (Alexander & Moore, 1958; Walker, 1958). Furthermore, we aimed to replicate the general finding of Experiment 1 that participants rushed more when they acted jointly with others than when they acted alone.

4.3.2 Results

As in Experiment 1, a one-sample t-test showed that participants' synchronization indices were significantly higher than the threshold of 0.73, $t(23) = 32.928$, $p < 0.001$, $d = 6.721$, with a mean of 0.94 and a standard deviation of 0.03. As can be seen in Figure 4.3A, participants showed a more negative tempo change in the Joint condition ($M = -22$ ms, $SD = 25$ ms) than in the Individual condition ($M = 3$ ms, $SD = 19$ ms), thereby replicating the overall joint rushing effect. A Welch t-test revealed that this difference was significant, $t(23) = 3.673$, $p = .001$, $d = 1.125$. Furthermore, tempo changes in joint trials were significantly different from zero, $t(23) = 4.306$, $p < .001$, $d = .879$, whereas tempo changes in the Individual condition were not, $t(23) = .727$, $p = .475$. Figure 4.3B shows the data segmented into bins of 10 seconds to depict the average development of the tempo over the course of a trial.

A 2x2 ANOVA with the factors Task (Individual / Joint) and Action Effects (Same / Different) confirmed these results with a main effect of Task, $F(1, 23) = 13.487, p = .001, \eta^2 = .212$ (see Figure 4.3C). There was no significant main effect for Action Effects, $F(1, 23) = .941, p = .342$, but a significant interaction between Task and Action Effects, $F(1, 23) = 15.278, p < .001, \eta^2 = .046$. Post-hoc t-tests showed that in the joint condition, there was a larger tempo increase for Same Action Effects ($M = -29$ ms, $SD = 32$ ms) than for Different Action Effects ($M = 15$ ms, $SD = 21$ ms), $t(23) = 3.421, p = .002, d = .511$. The difference between those two in the Individual condition was not significant, $t(23) = 1.563, p = .132$.

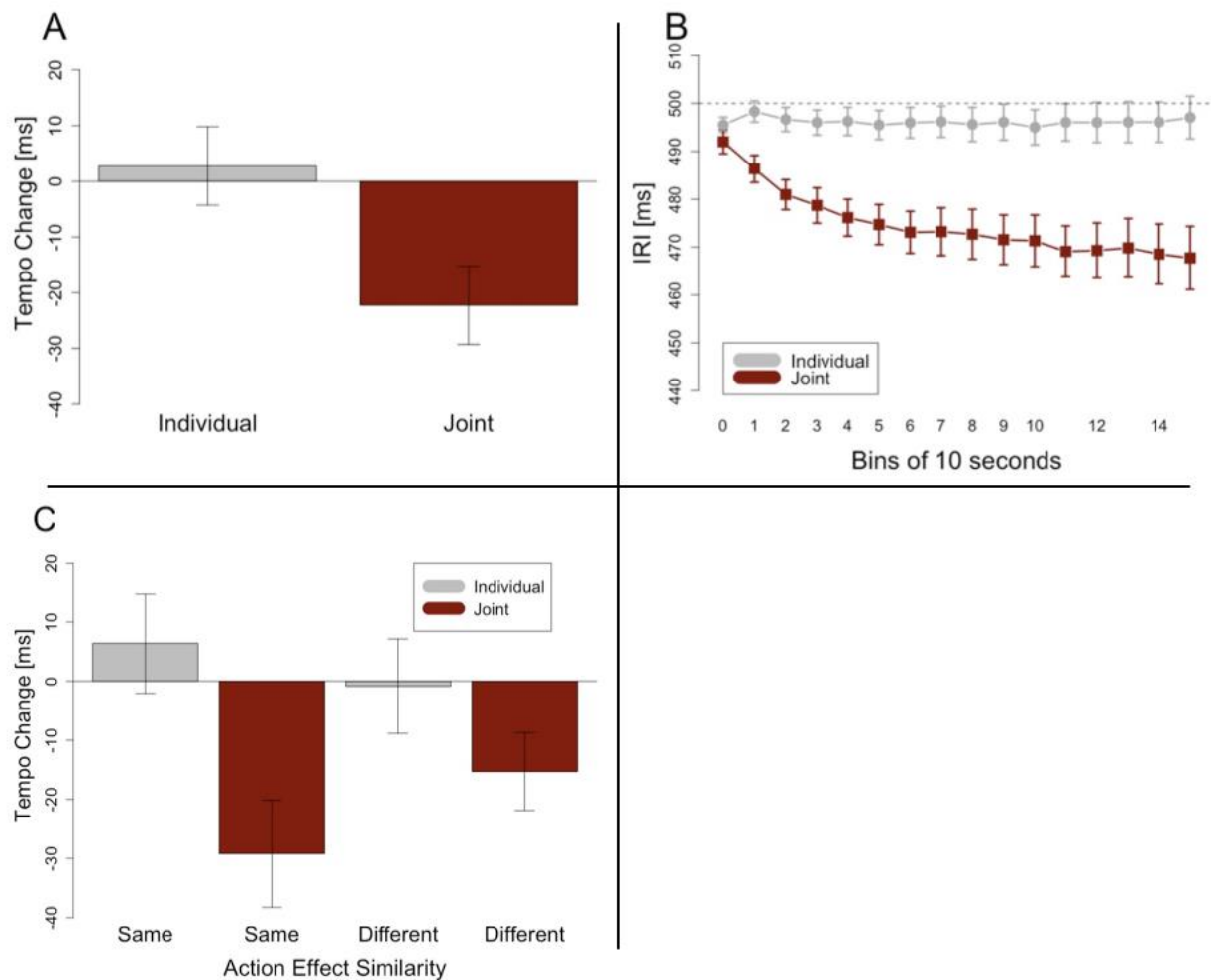


Figure 4.3: Results of Experiment 2. Tempo change indicates the difference between inter-response intervals at the end of a trial and at the beginning of a trial. Negative tempo change stands for a tempo increase. Error bars in A and C are calculated following the procedure recommended for within-subjects designs by Cousineau (2005). Error bars in B show standard errors. A) This panel shows significantly more tempo change in the joint condition

than in the Individual condition. B) Shown is the development of inter-response intervals over the course of trials, averaged over all participants and segmented into bins of 10 seconds each. C) Shows tempo change for Task and Action Effect Similarity. Post-hoc comparisons following a significant interaction revealed significantly more rushing in the joint condition with the same action effect than in the joint condition with a different action effect.

In Experiment 2, we tested whether producing the same sound versus producing a different sound than your partner has an effect on joint rushing. When the produced sounds were exactly the same joint rushing was more pronounced than when the sounds differed in pitch. This finding supports the hypothesis that a phase advance mechanism is at work in joint rushing, because this hypothesis predicts that the higher similarity between own signal and other signal will increase the likelihood of corrections occurring.

4.4 Experiment 3

4.4.1 Methods

In Experiment 3, we investigated whether performing the same actions—that is, movements with matching trajectories—has a similar effect on joint rushing as the similarity of action effects. This would be predicted by mirroring accounts (Brass et al., 2001) that highlight facilitating effects of performing the same action. Such a facilitation could cause joint rushing. Alternatively, if a phase advance mechanism is involved in causing joint rushing one would only expect this mechanism to be sensitive to signals that produce similar action effects (as demonstrated in Experiment 2) but not sensitive to another agent performing the same actions. The manipulation was implemented by asking participants to either hit the drum pads with the same trajectory, e.g. both hit a horizontally mounted drum pad with a vertical motion, or with different trajectories, i.e. one participant hitting a drum pad with a vertical motion,

while the other participant hits a vertically mounted drum pad with a horizontal trajectory. Additionally, we aimed to replicate the general findings of Experiment 1 and 2 that participants rush more when they act jointly with others.

4.4.2 Results

Synchronization indices were significantly higher than the threshold of 0.73, $t(23) = 71.339$, $p < 0.001$, $d = 14.562$, with a mean of 0.94 and a standard deviation of 0.01. As in Experiments 1 and 2, we predicted the tempo change from trials in group settings to be negative and significantly different from / more negative than the tempo change from trials in solo settings. As shown in Figure 4.4A, participants showed a more negative tempo change in the joint condition ($M = -17$ ms, $SD = 21$ ms) than in the Individual condition ($M = 5$ ms, $SD = 22$ ms). A Welch t-test revealed that this difference was significant, $t(23) = 3.964$, $p < .001$, $d = 1.042$. As in experiments 1 and 2, the tempo change for group trials was significantly different from zero, $t(23) = 4.041$, $p < .001$, $d = .825$, whereas the tempo change for trials in which participants acted alone was not significantly different from zero, $t(23) = 1.154$, $p = .260$. Figure 4.4B shows the data segmented into bins of 10 seconds to depict the average development of the tempo over the course of a trial.

Furthermore, a 2 x 2 ANOVA with the factors Task (Individual / Joint) and Actions (Same / Different), revealed a main effect for Task, $F(1, 23) = 15.713$, $p < .001$, $\eta^2 = .166$, but no main effect for Actions, $F(1, 23) = .007$, $p = .934$ (see Figure 4.4C). Surprisingly, the difference between Individual and Joint was larger in the Different Action condition than in the Same Action condition, $F(1, 23) = 5.176$, $p = .033$, $\eta^2 = .042$. Post-hoc comparisons revealed that the only significant difference was between Individual Different and Joint Different, $t(23) = 4.565$, $p < .001$, $d = .932$.

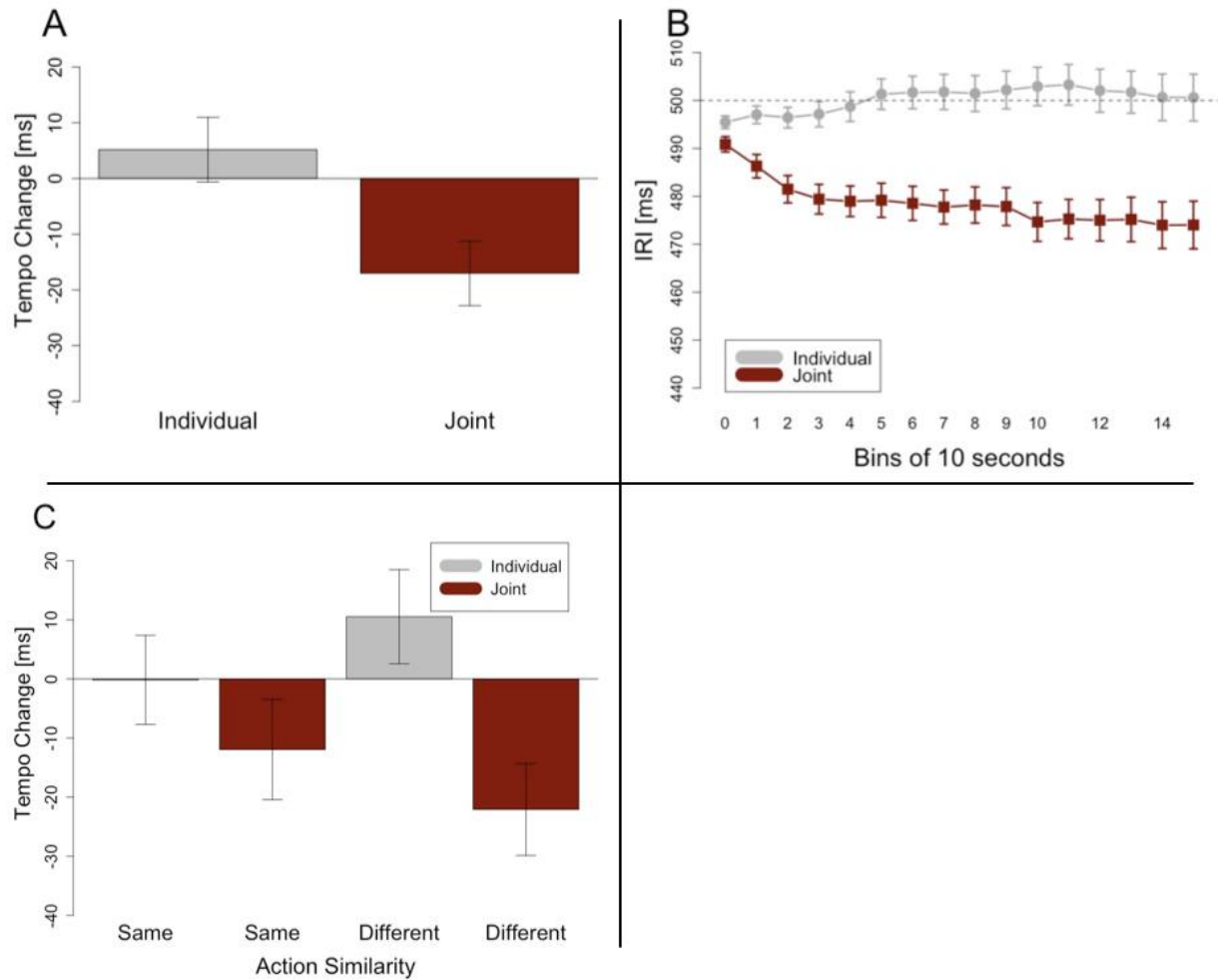


Figure 4.4: Results of Experiment 3. Tempo change indicates the difference between inter-response intervals at the end of a trial and at the beginning of a trial. Negative tempo change stands for a tempo increase. Error bars in A and C are calculated following the procedure recommended for within-subjects designs by Cousineau (2005). Error bars in B show standard errors. A) This panel shows significantly more tempo change in the joint condition than in the Individual condition. B) Shown is the development of inter-response intervals over the course of trials, averaged over all participants and segmented into bins of 10 seconds each. C) This panel shows tempo change for Task and Action Similarity. Besides a significant main effect for Task, and a significant interaction, post-hoc comparisons revealed that the difference between the two Joint conditions (Same and Different) is not significant.

The finding that performing similar actions does not enhance joint rushing demonstrates that joint rushing is unlikely to be a result of action mirroring. Rather, it indicates that joint rushing is caused by a phase advance mechanism that is tuned to specific auditory signals. An unexpected finding was that joint rushing was significantly larger when performing different actions than when performing same actions.

4.5 General Discussion

Three experiments provided evidence that joint rhythmic performance leads to joint rushing. There was no evidence for rushing in the individual control conditions. The findings are in line with our hypothesis that rushing is due to the interaction with other people. We also conclude from our three experiments that joint rushing results from a phase advance mechanism that exhibits stronger effects when the two auditory signals share the same pitch. Such a phase advance mechanism is expected to introduce a bias towards interval shortening when different individuals produce auditory signals in a similar frequency range. Whereas in chorusing insects the shortening biases are local (Greenfield, 2005), humans possess additional anticipation and adaptation mechanisms, which are likely to pick up local biases and to transform them into period corrections resulting in a global tempo change (Repp & Keller, 2008).

We controlled for a number of alternative explanations for joint rushing. Social facilitation – i.e. speeding up because others are present in the same environment – cannot explain the joint rushing effects observed in the present study. In the current experiments, an experimenter was always present in the Individual condition. According to the social facilitation hypothesis, this should have led to some degree of rushing in the Individual condition. However, across three experiments there was no indication of rushing in the Individual condition. It is also unlikely that action mirroring causes joint rushing: the results Experiment 3 provided no support for the prediction of action mirroring theories that the extent of rushing should be higher when participants produce the same movements rather than different movements. Surprisingly, rushing effects were actually larger for different movements than for same movements. One potential explanation for this unexpected finding is that participants try to consciously counteract tempo changes they notice -- a process that

may require cognitive control, which may be more difficult when different movements are performed.

The present findings are in line with two recent studies that have reported tempo increases during rhythmic activities in groups of different sizes. Okano, Shinya and Kudo (2017) analyzed the data of 24 adults finger-tapping in a synchronization-continuation paradigm either in a paired or a solo condition and found that rushing occurred to a larger extent in the Joint condition. Okano et al. concluded that an interpersonal adaptation mechanism related to tap asynchrony underlies joint rushing. We add to this conclusion the proposal that an evolutionary ancient phase advance mechanism may be a central component of joint rushing.

Thomson, Murphy and Lukeman (2018) investigated synchronous clapping in groups of varying size. They found an asymmetrical period response curve with stronger corrections being made when a shortening of the period is required than when the period has to be prolonged. The proposed phase advance mechanism predicts stronger corrections for preceding signals than for following signals. Contrary to the results of Experiment 1 that showed no significant difference in joint rushing between groups of 2 and groups of 3, there was a positive correlation between group size and joint rushing in the Thomson et al. (2018) study where group size varied from 7 to 220. The proposed phase advance mechanism predicts such effects of group size because an increase in neighboring signalers translates to an increased chance of a signal falling into the sensitive window where a shortening of intervals occurs.

Given that robust effects of joint rushing have been established for different kinds of rhythmic performance at different tempi, the question arises how it is possible for humans to avoid joint rushing. According to our model, one way to reduce joint rushing would be to produce constantly low asynchronies between performers to prevent the occurrence of signals in the sensitive window which provoke interval shortenings. Thomson et al. (2018) reported

incidental evidence that rhythmically trained individuals show a reduced effect of joint rushing. Musical training is both known to lead to a reduction in motor/movement variability (Aoki, Furuya & Kinoshita, 2005; Loehr & Palmer, 2007) and to an improvement of tempo change detection abilities (Ellis, 1991; Garner, 2015). This would explain how (some) musical experts can actually keep the tempo during joint music-making.

Whereas joint rushing can be a problem when the aim is a constant tempo during a joint music performance – according to the internet, the Rolling Stones and other rock bands seem to have struggled with this (bvdd, 2001) – joint rushing may have advantageous effects for other forms of joint action that put less constraints on tempo. People that have to coordinate their actions in time have been found to make themselves more predictable by increasing their movement speed (Vesper, Van der Wel, Knoblich & Sebanz, 2011; Vesper, Schmitz, Sebanz & Knoblich, 2013). If more variability leads to more rushing, as Thomson et al. (2018) suggest, and the increased tempo after rushing leads to a reduction in variability, joint rushing could be part of a self-regulating mechanism that ensures smooth interactions in groups by increasing tempo as long as this increase implies a reduction of variability.

Next steps in research on joint rushing could be to examine the influence of musical training on joint rushing, and to tease apart which components of musical expertise allow interacting partners to reduce the effects of joint rushing if any. Furthermore, it would seem important to study the relationship between individual and group variability on one side and joint rushing on the other side. Such studies could answer the question whether joint rushing functions as a coordination smoother (Vesper, Butterfill, Knoblich & Sebanz, 2010) for rhythmic joint actions. Such a function could lead to a selective advantage and this would explain why humans and chorusing insects share a similar phase advance mechanism.

4.6 Methods

For Experiment 1, we invited 24 participants (15 women, 8 men, 1 unspecified, mean age = 24.7 years, SD = 4.0 years), with little to no musical training ($M = 0.4$ years, $SD = 1.3$ years). The sample size was determined through a power calculation based on our expectancy to see large effect sizes. This was motivated by the fact that the joint rushing effect has been reported to be perceivable without formal measurements. Post-hoc power analyses with the obtained effect sizes confirmed our intuitions with power values ranging from 0.86 to 0.97 for the main comparisons across the three experiments. We kept the number of participants constant across all three experiments. Participants in all three experiments gave their informed consent and received gift vouchers as compensation. All experiments in this study were conducted in accordance with the Declaration of Helsinki and approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

We used three Millenium MPS-400 Tom pads connected to a ddrum DDTi trigger interface to record responses, which participants produced with one wooden drum stick each. Auditory feedback, metronome beats, and data recording was handled with a custom Max MSP patch. Each participant produced a different piano pitch, with all pitches being more than an octave apart (15 semitones), centered around F#4.

Participants performed eight synchronization-continuation trials in four blocks. Each block consisted either of two trials where participants tapped alone (individual trials) or two trials of tapping in a group (group trials). The order of blocks was counterbalanced. During individual trials, a participant's partner(s) waited in another room. In both, individual and group trials, an experimenter was present and sat in close proximity to the participant(s) to control the experimental program.

At the beginning of each trial participants heard a metronome with an Inter-Onset-Interval of 500 ms for 10 seconds. Participants were asked to synchronize with the metronome

in the metronome phase and to then continue in the same tempo after the metronome faded out. After 150 seconds had elapsed and a minimum of 270 taps had been produced, participants heard a percussion sound informing them that the trial was over. After each trial participants indicated how difficult they found the trial on a scale from 1 (“not at all difficult”) to 7 (“very difficult”). This was included to see whether participants would perceive the task in the Joint condition to be more or less difficult than in the Individual condition. Participants were asked to fill in another questionnaire after the experiment, from which we calculated the years of training on a musical instrument.

In Experiment 1, the main within manipulation was Task, which determined whether participants performed the task alone (Individual) or in a group (Joint). Group Size was manipulated as a between factor, with half the participants performing the joint task in groups of two (Two) and half of the participants in groups of three (Three). Furthermore, we asked participants to strike the pads with two different target forces, one low and one high. The low target force was introduced in an attempt to increase participants’ variability. Participants were provided with visual feedback that indicated to them whether they had hit the drum pads with the right force.

For Experiment 2 we invited 24 new participants (14 women, 10 men, mean age = 25.7 years, SD = 4.7 years), with little to no musical training (M = 0.0 years, SD = 0.1 years). The procedure was the same as in Experiment 1. Participants acted alone (Task: Individual) or in a group of two (Task: Joint). In Experiment 2 we also manipulated whether participants produced the same pitch (Same) or pitches that were 15 semitones apart (Different). This was implemented in such a way that for each participant it sounded as if the other’s pitch changed in the Same condition, i.e. for one participant both drum pads were heard as producing a D#₃, whereas for the other participant it sounded as if both drum pads produced an F#₄. This was made possible by using different channels for each participant’s auditory feedback. To sum up

the design of Experiment 2, we used a 2 x 2 within-subjects design with the factors Task (Individual vs. Joint) and Action Effect (Same vs. Different).

For Experiment 3 we invited 24 new participants (16 women, 8 men, mean age = 25.0 years, SD = 3.8 years), with little to no musical training (M = 0.0 years, SD = 0.1 years). We kept the procedure the same as in Experiment 2 but exchanged the factor Action Effects with the factor Action. For this experiment, we used two drum pads on each stand, one mounted horizontally and one mounted vertically. Instead of hearing the same or different pitches as in Experiment 2, participants struck the drum pad either in the same way (e.g. both hit the pad with a vertical movement) or in different ways (e.g. participant A hit the drum pad vertically, while participant B hit the drum pad horizontally). In both conditions both drum pads produced a piano sound on C₄. To summarize, in Experiment 3, we used a 2 x 2 within-subjects design with the factors Task (Individual vs. Joint) and Actions (Same vs. Different).

To determine whether participants followed the instruction to synchronize in the Joint condition, we computed synchronization indices based on the circular variance of relative phase (Mardia & Jupp, 2000). This unitless index reaches from 0, absence of synchronization, to 1, perfect synchrony. In line with the convention of previous studies, we considered indices > 0.73 to be indicative of the occurrence of synchronization (Konvalinka, Vuust, Roepstorff & Frith, 2010; Tognoli et al., 2007). To assess the tempo change over the course of a whole trial we compute the difference between the mean of each participant's inter-response intervals (IRIs) in the beginning, i.e. during the metronome phase, and the mean of the last twenty IRIs for any given trial in milliseconds. Positive values of tempo change indicate that participants had larger IRIs in the end than in the beginning, i.e. slowed down, whereas negative values of tempo change indicate a shortening of IRIs, i.e. an increase in tempo.

5 General Discussion

The aim of the present thesis was to contribute to the scientific endeavor of understanding how joint action partners manage to achieve tight temporal coordination while allowing for remarkable flexibility. Previous research suggests that the best condition for temporal coordination is met when two expert musicians (Keller, Knoblich & Repp, 2007) coordinate their actions in an in-phase coordination pattern (Schmidt, Carello & Turvey, 1990) at a comfortable speed (Zamm, Wang & Palmer, 2018) close to their shared preferred tempo (Zamm, Pfordresher & Palmer, 2015). Humans that engage in joint music-making rarely find themselves in such ideal circumstances. Instead they face a multitude of challenges originating from their and their partners abilities, complex coordination requirements and systematic biases emerging in rhythmic joint actions. Yet, during joint music-making temporal coordination can be achieved at an astounding level. In this thesis I focused on three challenges musicians may face when coordinating with each other: coping with systematic temporal deviations originating from differences in skill level, producing extraordinary coordination patterns, which are commonly encountered during joint music-making, and keeping the tempo. In the following sections I will summarize the findings presented in this thesis and discuss their theoretical implications as well as open questions for further research.

5.1 Joint Action Coordination in Expert-Novice Pairs

In Chapter 2, I investigated whether and how predictive mechanisms are at work in expert-novice interactions. Whereas previous research has focused on expert-expert pairs, joint music-making involves many practices in which interaction partners of different skill levels engage in joint action. These practices include for example playing in non-professional

ensembles and a wide range of teaching situations. Therefore, in Chapter 2, I investigated how joint action partners with different skill levels achieve coordination, whether experts can adjust their predictions to the suboptimal timing of novices, whether they can use different kinds of information about what and how the novice is going to perform, and whether experts' predictive abilities depend on the difficulty of the novice's performance. To do so, I used a coordination paradigm with expert and novice pianists. The results revealed that familiarity with the score led to better coordination when the score implied a difficult passage. Familiarity with novices' idiosyncratic performance styles led to better joint action coordination for the remaining parts of the duet. Together, these results indicated that experts are quite flexible in predicting novices' suboptimal timing. As the information manipulations were incorporated before the coordination task began, we could also conclude that experts managed to adjust their predictions offline, i.e., before the interaction with the novice started.

These findings relate to and expand on previous research that suggests that skilled performers use internal models to predict the timing of other expert performers (Keller, Knoblich & Repp, 2007; Ragert, Schroeder & Keller, 2013). Previous results suggest that one's own motor system is used in making these predictions. This, however, is problematic when the other's motor system is very different, as in the case of expert-novice pairs. The results in Chapter 2 imply that experts can adjust predictions generated in their own motor system to adapt to co-actors with different skill levels and, therefore, different accuracy with regard to a prescribed timing. Furthermore, the fact that the manipulations were introduced before the coordination phase allows us to infer that these adaptations can be implemented offline. I propose that this is achieved with the help of internal models that simulate one's own playing while hearing the novice's playing. The timing discrepancies between the imagined own outcomes and the novice's outcomes could be used to generate a performance-specific error matrix. Such a matrix may be used to directly adapt the internal model that is used to derive

the predictions, or it may be added on top of the predictions to modify their timing. In Chapter 2.5, I investigated whether these adaptations are influenced, e.g. enhanced, by the pitch information of the novice's part. The results showed no significant influence of pitch, indicating that these adaptations are mainly driven by temporal information and that pitch does not contribute to efficiency of how an error matrix is generated and used in subsequent trials.

5.1.1 Future Research

In this paradigm, I intentionally restricted auditory feedback to be unidirectional, so that the novice never heard the expert play throughout the whole experiment. Though this was appropriate for the purpose of focusing on the expert's ability to adapt to a novice's suboptimal timing, it leaves questions open for future research. Due to the asymmetry in skill level, experts could be forced to assume the role of a teacher when there is bidirectional auditory feedback. One of the objectives of the expert could then be to teach stable timing by providing a stable performance that the novice is supposed to adapt to over time. This would require a reduced adaptation of the expert to the novice's timing and would be an instance of a strategic non-adaptation of the expert to provide scaffolding for the novice. Whether this is indeed a strategy that experts engage in and whether it would accelerate the pace in which the novice improves her performance are open questions for future research.

The random-pairing analysis described in Chapter 2 supported the conclusion that experts adapt to the timing of a specific novice and not to any unstable timing in general. The fact that the significant main effect of familiarity with novice performance comes from a within-subjects design (i.e., significant differences are obtained from different melodies of the same expert-novice pair), leads me to suggest that the expert's adaptation was not just novice-specific, but to some extent also specific to a specific piece or even a certain performance. Further research could investigate how specific the adaptations are exactly, and whether they can

be retained for multiple different performances with different partners. To illustrate, imagine a musician playing the same piece with two different partners. Is she able to retain her adaptations to the two partners that exhibit different timing schemes due to different skill levels or different goals of expressive timing? How flexibly can she switch from one to the other? The results from more difficult passages in my study suggest that explicit performance cues enhance coordination on larger timescales. A further question derived from this result is whether such cues also play a role when a musician needs to switch between different partners.

5.2 Adaptation to Unstable Phase Relations

In Chapter 3, I investigated the question of how well musicians can adapt to extraordinary coordination patterns and whether there is an advantage of distributing these patterns across musicians in comparison to intrapersonal bimanual execution. I also examined whether added structure to the perceptual outcome facilitates the performance of difficult phase relations. The results showed that for musicians in general, there was an advantage of joint performance. Participants reached their baseline performance more often in joint than in individual performance and adaptation was more effective for joint performance. I interpreted these effects to be mediated by coupling strength which is reduced between the moving limbs of two joint action partners compared to two limbs of the same person.

It is important to note that individual participants did manage to overcome the challenges posed by the extraordinary coordination pattern under all conditions, crucially also during individual performance. Hence, musicians are able to overcome entrainment effects and adapt to difficult phase relations in joint as well as in solo performance. One possible way in which they accomplished this during solo performances could be using the harmonic pitch

structure of action outcomes inherent in music-making. In my study, however, adding harmonic pitch structure to the action outcome did not improve participants' adaptation rates.

These results are in line with previous research on interlimb coordination. Extraordinary coordination patterns, in this case non-harmonically related phase-shifts, appear to be more difficult to perform under stronger coupling than under weaker coupling. However, the particular set-up of this study adds to previous findings in several ways. First, I provide evidence that the benefits of weaker coupling for interlimb coordination are not initially apparent but show their effects in adaptation rates over time. This led me to conclude that joint performance provides advantages for adapting different movements to produce extraordinary coordination patterns. These findings could inform future research on the interplay between coordination attractors in phase space and the notion of cultural attractors in the cultural evolution of music, and more specifically, for the diversification of music. Joint music-making may have acted as an innovation device for musical patterns because it enabled individuals to explore a larger space of rhythmic possibilities while acting together with others.

Second, previous research (Dyer, Stapleton & Rodger, 2017) as well as musical practices suggest that adding structure to patterns makes them easier to parse. Some piano students use special sentences to internalize extraordinary coordination patterns. Some examples of which are “warm turtle soup”, which, if said out loud exhibits the rhythmic structure of a 2 against 3 polyrhythm (Dow, 2013), or “not difficult, not difficult”, which does the same for a 4 against 6 polyrhythm (Bradshaw, 2010). However, I did not find evidence for this in the current study. It is important to note that the way I implemented the added structure is more closely related to the task of Mechsner et al. (2001), where the action outcome provided the participant with a simplified pattern. This is slightly different from structure that embeds the complexity as it is the case in the polyrhythm-sentences mentioned above.

5.2.1 Future Research

It remains to be investigated what enabled some musicians to adapt well to the extraordinary coordination pattern, while some other musicians barely did so at all. When I compared linear regression models using biographic data from post-experiment questionnaires to predict performance, I determined the best predictors to be years of training and age, where years of training show a positive effect and age a negative effect on adaptation to extraordinary coordination patterns. But even taken together these two factors only accounted for less than 20 % of the variance. Hence, further research is required to investigate what enables individuals to achieve quick adaptation to situation that require production of extraordinary movement patterns. It also remains to be seen whether musicians are able to adapt to multiple delays and to switch between them. For organists, for example, this seems to be a requirement when switching between instruments with different delay profiles. This also raises the question of how detailed such mappings can be. For instance, organs exhibit pitch and register dependent delay times (Pollard, 1968). It remains to be seen whether this requires organists to use a detailed delay matrix for pitch, manual, and register to achieve the desired degree of coordination or whether they use shortcuts or approximations.

Transmission chain experiments in which solo performances are reproduced and transmitted from generation have shown that the results converge to exhibit what has been listed as universals of musical rhythm (Ravignani, Delgado & Kirby, 2016). Further research is needed to determine whether transmission via individual actions converges on different features across generations than transmission via joint performances. One hypothesis could be that transmission via joint performances leads to more diversification than transmission through individual performances.

5.3 Why do People Rush during Rhythmic Interactions?

In Chapter 4, I investigated the phenomenon that people in rhythmic joint actions unintentionally increase their tempo. This phenomenon is an instance of a systematic bias which emerges from rhythmic interactions of multiple individuals, and is not present during individual performance. Such biases can pose challenges for successful coordination and explaining how they emerge and how interaction partners overcome them can help us to refine our understanding of specifics of temporal coordination in joint actions.

When people engage in rhythmic interactions, they often speed up unintentionally. Whereas this seems to be common knowledge among music teachers it had not been investigated scientifically. The results provided evidence that this *joint rushing* is indeed a product of interpersonal coordination and does not appear to the same extent in solo performances. I found medium to strong effects for this difference between joint and solo conditions across three experiments. As an explanation for this phenomenon I suggested the combination of an evolutionarily ancient mechanism that is similar to phase advance mechanisms in chorusing insects with human period correction mechanisms. Possible alternative explanations that I tested were based on the literature on social facilitation effects and on effects of action mirroring. The results showed that joint rushing is a robust phenomenon occurring in groups of two as well as in groups of three, but not during solo performance. Joint rushing was more pronounced when the action effects produced by different individuals were perceptually similar, which is in line with coordination mechanisms in chorusing insects in general, which are tuned to conspecifics to avoid inter-species crosstalk. Further control conditions ruled out the alternative hypotheses that rushing during rhythmic interactions is based on social facilitation or action mirroring effects. Whereas joint rushing usually needs to be contained during joint music-making, in other joint actions, where keeping a pre-defined tempo is not required, joint rushing might actually be beneficial for the

interaction. By increasing the tempo of rhythmic interactions, it could bring interaction partners to their optimal tempo at which they exhibit the least amount of variability and thereby facilitate smooth interactions.

5.3.1 Future Research

Whether joint rushing indeed serves a function or not, sometimes it is not desired, especially when a musical ensemble tries to perform a piece in a steady tempo. Anecdotal evidence (Thomson et al., 2018) indicates that musicians are able to reduce joint rushing to a considerable extent. Systematic testing is required to decide whether expert musicians are indeed able to do so, and if yes, how they are able to do so. The reduction in joint rushing might simply be a by-product of the reduced timing variability that musicians exhibit compared to non-musicians (Repp, 2010) or the results of conscious and constant efforts to counteract the effects of joint rushing.

Further research is also needed to investigate whether joint rushing can indeed function as a coordination smoother by causing the tempo of a rhythmic interaction to accelerate precisely until an optimal tempo is reached. If this is the case, it could have far-reaching implications for our understanding of temporal coordination in joint action. It could also lead to the development of strategies to improve temporal coordination in human-human interaction, but especially in the realm of human-robot interactions.

5.4 Joint Music-Making and Temporal Coordination in Joint Action

In this thesis, I raised questions about temporal coordination that were inspired by practices from the domain of joint music-making. In this domain joint actions require precise temporal coordination under occasionally exceptionally challenging circumstances. The three studies focus on three of these challenges, but the findings have wider implications. They offer

insights into timing mechanisms that apply to joint actions in general and open up new areas of research in joint action. As joint music-making has been proposed to have served the function of an efficient and scalable social bonding device, joint music-making might also prove to be a valuable tool to probe temporal coordination in joint actions that go beyond the dyad.

As joint music-making unfolds in time, it lends itself quite naturally to the study of temporal coordination in joint action. Indeed, all of the studies in this thesis investigated mechanisms and phenomena in the temporal dimension. However, it should be noted that joint music-making should also be considered as an option for the study of other dimensions. Interacting musicians need to coordinate their outcomes for example also along dimensions of pitch and harmony which on most instruments corresponds to coordinating actions in space, like on the keyboard of a piano or on the fretboard of a guitar. Future research could build on scalability of joint music-making and the possibility to map pitch space onto physical space to study spatial coordination in large scale joint actions. It could also be of interest to use joint music-making to study coordination on more abstract dimensions, such as coordination in pitch space, timbre or musical expressiveness.

6 References

- Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, 5(3), 163–180. <https://doi.org/10.1037//1089-2699.5.3.163>
- Alexander, R. D. , & Moore, T. E. (1958). Studies on the Acoustical Behavior of Seventeen-Year Cicadas (Homoptera: Cicadidae: Magicicada). *The Ohio Journal of Science*. *Ohio Journal of Science**Ohio Academy of Science*) *Ohio Journal of Science*, 58(2), 2–107.
- Aoki, T., Furuya, S., & Kinoshita, H. (2005). Finger-tapping ability in male and female pianists and nonmusician controls. *Motor Control*, 9(1), 23–39. <https://doi.org/15784948>
- Arom, S. (1991). *African Polyphony and Polyrhythm: Musical Structure and Methodology*. <https://doi.org/10.1017/CBO9780511518317>
- Aucouturier, J. J., & Canonne, C. (2017). Musical friends and foes: The social cognition of affiliation and control in improvised interactions. *Cognition*. <https://doi.org/10.1016/j.cognition.2017.01.019>
- Bamford, J., Burger, B., & Toiviainen, P. (2016). Are We Dancing to the Same Beat? Empathy and Interpersonal Synchronisation in the Silent Disco. *9th International Conference of Students of Systematic Musicology (SysMus 16)*, Jyväskylä, Finland, 7.
- Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience*, 26(4), 1–14. <https://doi.org/10.1186/1471-2202-4-26>
- Barnes, G. V. (2008). Challenging the Middle School Orchestra. *American String Teacher*, 58(3), 38–41. <https://doi.org/10.1177/000313130805800307>
- Blair, K. G. (1915). Luminous Insects 1. *Nature*, 96(2406), 411–415. <https://doi.org/10.1038/096411a0>
- Blakemore, S.-J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, 43(2), 260–267.
- Bradshaw, T. (2010, September 2). Phrases for Grouplets or Polyrhythms: Retrieved March 19, 2019, from Tony Bradshaw's Music Education website: <https://mymusiceducation.wordpress.com/2010/09/02/phrases-for-grouplets-polyrhythm/>

- Braff, M. (2015a). General Theory of Rhythm: Basic Principles. Retrieved February 6, 2019, from <http://general-theory-of-rhythm.org/basic-principles/>
- Braff, M. (2015b). General Theory of Rhythm: Diversity of grooves: it's all about da weight ! Retrieved February 6, 2019, from <http://general-theory-of-rhythm.org/diversity-of-grooves-its-all-about-da-weight/>
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 20.
- Brodsky, W., Kessler, Y., Rubinstein, B. S., Ginsborg, J., & Henik, A. (2008). The Mental Representation of Music Notation: Notational Audiation. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/0096-1523.34.2.427>
- Brodsky, W., & Sulkin, I. (2011). Handclapping songs: a spontaneous platform for child development among 5–10-year-old children. *Early Child Development and Care*, 181(8), 1111–1136. <https://doi.org/10.1080/03004430.2010.517837>
- Brown, D. E. (1991). *Human universals*. Boston, Mass.: McGraw-Hill.
- Buck, J. (1988). Synchronous rhythmic flashing of fireflies. II. *Quarterly Review of Biology*, 63(3), 265–289. <https://doi.org/10.1086/415929>
- Butterfill, S. (2016). Ch 20: Joint Action - A Minimalist Approach. In J. Kiverstein (Ed.), *The Routledge Handbook of Philosophy of the Social Mind* (pp. 357–369).
- bvdd. (2001). Unintentional tempo changes: “Adrenaline Rush” - MusicPlayer Forums. Retrieved June 15, 2018, from http://forums.musicplayer.com/ubbthreads.php/topics/291854/Re_Unintentional_temp_o_changes
- Carl A. Richmond. (1930). Fireflies Flashing in Unison. *Science, New Series*, 71(1847), 537–538.
- Collier, G. L., & Ogden, R. T. (2004). Adding Drift to the Decomposition of Simple Isochronous Tapping: An Extension of the Wing-Kristofferson Model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(5), 853–872. <https://doi.org/10.1037/0096-1523.30.5.853>
- Colson, J. F. (2012). *Conducting and Rehearsing the Instrumental Music Ensemble*.
- Constable, M. D., Bayliss, A. P., Tipper, S. P., Spaniol, A. P., Pratt, J., & Welsh, T. N. (2016). Ownership Status Influences the Degree of Joint Facilitatory Behavior.

- Psychological Science*, 27(10), 1371–1378.
<https://doi.org/10.1177/0956797616661544>
- Copland, A. (2004). Jazz Structure and Influence. In R. Kostelanetz (Ed.), *Aaron Copland: A Reader - Selected Writings 1923 - 1972*. Routledge.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 2–5. <https://doi.org/10.20982/tqmp.01.1.p042>
- Cross, I. (2008). Musicality and the human capacity for culture. *Musicae Scientiae*, 12(1_suppl), 147–167. <https://doi.org/10.1177/1029864908012001071>
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153. <https://doi.org/10.1016/j.tics.2009.01.005>
- DAusilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, 19(3), 1–4. <https://doi.org/10.1016/j.tics.2015.01.005>
- Davidson, J. W. (2009). Movement and collaboration in musical performance. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 364–376).
- Davies, M., Madison, G., Silva, P., & Gouyon, F. (2013). The Effect of Microtiming Deviations on the Perception of Groove in Short Rhythms. *Music Perception: An Interdisciplinary Journal*, 30(5), 497–510. <https://doi.org/10.1525/mp.2013.30.5.497>
- Delacre, M., Lakens, D., & Leys, C. (2017). Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test. *International Review of Social Psychology*, 30(1), 92. <https://doi.org/10.5334/irsp.82>
- Desain, P., & de Vos, S. (1990). Autocorrelation and the study of musical expression. *Proceedings of the International Computer Music Conference*, 357–360.
- Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception & Psychophysics*, 34(4), 331–337. <https://doi.org/10.3758/BF03203045>
- Dillard, S. (2013). Why does a clapping audience automatically increase the clap frequency? - Quora. Retrieved June 15, 2018, from <https://www.quora.com/Why-does-a-clapping-audience-automatically-increase-the-clap-frequency>
- Dow, A. (2013, June 3). Avoid Polyrhythms on the Piano Like the Plague... Or, Just Read This! Retrieved March 19, 2019, from <http://www.teachpianotoday.com/> website:

<https://www.teachpianotoday.com/2013/06/03/avoid-polyrhythms-on-the-piano-like-the-plague-or-just-read-this/>

- Drewing, K., & Aschersleben, G. (2003). Reduced Timing Variability during Bimanual Coupling: A Role for Sensory Information. *The Quarterly Journal of Experimental Psychology Section A*, 56(2), 329–350. <https://doi.org/10.1080/02724980244000396>
- Dyer, J. F., Stapleton, P., & Rodger, M. (2017a). Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning. *Experimental Brain Research*. <https://doi.org/10.1007/s00221-017-5047-8>
- Dyer, J. F., Stapleton, P., & Rodger, M. (2017b). Transposing musical skill: sonification of movement as concurrent augmented feedback enhances learning in a bimanual task. *Psychological Research*. <https://doi.org/10.1007/s00426-016-0775-0>
- Ellis, M. C. (1991). Thresholds for detecting tempo change. *Psychology of Music*, 19(2), 164–169.
- Eronen, A., & Klapuri, A. (2000). Musical instrument recognition using cepstral coefficients and temporal features. *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, 2, II753–II756. <https://doi.org/10.1109/ICASSP.2000.859069>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R* (Reprint). Los Angeles, Calif.: Sage.
- Finnegan, R. H. (2007). *The hidden musicians: music-making in an English town* (1. ed). Middletown, Conn: Wesleyan University Press.
- Flach, R., Knoblich, G., & Prinz, W. (2004). Recognizing one's own clapping: The role of temporal cues. *Psychological Research Psychologische Forschung*, 69(1–2), 147–156. <https://doi.org/10.1007/s00426-003-0165-2>
- Folio, C. (1945). An analysis of polyrhythm in selected improvised jazz solos. *Concert Music, Rock, and Jazz Since*, 103–134.
- Fuchs, F. (2019). Rhythmus - Online-Volksmusikschule. Retrieved February 6, 2019, from <https://www.volksmusikschule.at/rhythmus.htm>
- Gabler, F. J. (1998). Wiener Walzer - Vergleiche der traditionellen lokalen und der internationalen Walzerbegleitung. *Reden & Schriften / Universität Für Musik Und Darstellende Kunst Wien*, 4.

- Garner, D. K. (2015). That Driving Sound: Use of Tempo in Traditional Cape Breton Fiddle Performance. *Musicultures - Journal of the Canadian Society for Traditional Music*, 42(2).
- Ginsborg, J., Chaffin, R., & Nicholson, G. (2006). Shared performance cues in singing and conducting: a content analysis of talk during practice. *Psychology of Music*, 34(2), 167–194. <https://doi.org/10.1177/0305735606061851>
- Ginsborg, J., & King, E. (2012). Rehearsal talk: Familiarity and expertise in singer-pianist duos. *Musicae Scientiae*, 16(2), 148–167. <https://doi.org/10.1177/1029864911435733>
- Gludovatz, P. (2009). *Aktuelle Bestandsaufnahme des Wiener Walzers in Wien* (Diplomarbeit). University of Vienna.
- Gonzalez, D. A., Studenka, B. E., Glazebrook, C. M., & Lyons, J. L. (2011). Extending end-state comfort effect: Do we consider the beginning state comfort of another? *Acta Psychologica*, 136(3), 347–353. <https://doi.org/10.1016/j.actpsy.2010.12.009>
- Gould, R. G., & Helder, G. K. (1970). Transmission delay and echo suppression. *IEEE Spectrum*, 7(4), 47–54. <https://doi.org/10.1109/MSPEC.1970.5213327>
- Green, A., McEllin, L., Felber, A., & Michael, J. (under review). *Prosocial Effects of Coordination – What, How and Why?*
- Greenfield, M. D. (1994a). Cooperation and Conflict in the Evolution of Signal Interactions. *Source: Annual Review of Ecology and Systematics*, 25, 97–126.
- Greenfield, M. D. (1994b). Synchronous and Alternating Choruses in Insects and Anurans: Common Mechanisms and Diverse Functions. *American Zoologist*, 34(6), 605–615. <https://doi.org/10.1093/icb/34.6.605>
- Greenfield, M. D. (2005). Mechanisms and Evolution of Communal Sexual Displays in Arthropods and Anurans. *Advances in the Study of Behavior*. [https://doi.org/10.1016/S0065-3454\(05\)35001-7](https://doi.org/10.1016/S0065-3454(05)35001-7)
- Grèzes, J., Frith, C. ., & Passingham, R. . (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *NeuroImage*, 21(2), 744–750. [https://doi.org/10.1016/S1053-8119\(03\)00665-7](https://doi.org/10.1016/S1053-8119(03)00665-7)
- Guettler, K., & Hallam, S. (2002). String Instruments. In R. Parncutt & G. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 303–318).

- Haken, H., Kelso, J. a. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51(5), 347–356.
<https://doi.org/10.1007/BF01068748>
- Haueisen, J., & Knösche, T. R. (2001). Involuntary Motor Activity in Pianists Evoked by Music Perception. *Journal of Cognitive Neuroscience*.
<https://doi.org/10.1162/08989290152541449>
- Helmuth, L. L., & Ivry, R. B. (1996). When two hands are better than one: Reduced timing variability during bimanual movements. *Journal of Experimental Psychology: Human Perception and Performance*, 22(2), 278–293. <https://doi.org/10.1037/0096-1523.22.2.278>
- Herrera-Boyer, P., Peeters, G., & Dubnov, S. (2003). Automatic Classification of Musical Instrument Sounds. *Journal of New Music Research*.
<https://doi.org/10.1076/jnmr.32.1.3.16798>
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1664), 20140088–20140088.
<https://doi.org/10.1098/rstb.2014.0088>
- Jakubowski, K., Finkel, S., Stewart, L., & Müllensiefen, D. (2017). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, 11(2), 122–135.
<https://doi.org/10.1037/aca0000090>
- James, D. (2013). Signposted by song: Cultural routes of the Australian desert. *Historic Environment*, 25(3), 30–42.
- Jones, M. R. (2009). Chapter 8: Musical time. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 81–92). Oxford University Press.
- Kay, B. A., Saltzman, E. L., Kelso, J. A. S., & Schöner, G. (1987). Space-Time Behavior of Single and Bimanual Rhythmical Movements: Data and Limit Cycle Model. *Journal of Experimental Psychology Human Perception and Performance*, 009687(2), 178–192.
- Keller, P. E. (2008). Joint Action in Music Performance. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions* (pp. 205–221). Amsterdam: IOS Press.
- Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization.

- Consciousness and Cognition*, 16(1), 102–111.
<https://doi.org/10.1016/j.concog.2005.12.004>
- Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. <https://doi.org/10.1098/rstb.2013.0394>
- Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *The American Journal of Physiology*, 246(6 Pt 2), R1000–4.
- Klapp, S. T. (1979). Doing two things at once: The role of temporal compatibility. *Memory & Cognition*, 7(5), 375–381. <https://doi.org/10.3758/BF03196942>
- Knoblich, G., & Flach, R. (2001). Predicting the Effects of Actions: Interactions of Perception and Action. *Psychological Science*, 12(6), 467–472.
<https://doi.org/10.1111/1467-9280.00387>
- Knoblich, Günther, Butterfill, S. A., & Sebanz, N. (2011). *Psychological research on joint action : theory and data*. <https://doi.org/10.1016/B978-0-12-385527-5.00003-6>
- Knoblich, Günther, & 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>
- Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, 110(38), 15443–15448. <https://doi.org/10.1073/pnas.1300272110>
- Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology (2006)*, 63(January 2015), 2220–2230.
<https://doi.org/10.1080/17470218.2010.497843>
- 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
- Kubik, G. (1960). The structure of Kiganda xylophone music. *African Music: Journal of the African Music Society*, 2(3), 6–30. <https://doi.org/10.21504/amj.v2i3.606>
- Kubik, G. (1964). Xylophone playing in southern Uganda. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 94(2), 138–159.

- Large, E. W., & Kolen, J. F. (1994). Resonance and the Perception of Musical Meter. *Connection Science*, 6(2–3), 177–208. <https://doi.org/10.1080/09540099408915723>
- Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*.
- Lerdahl, F., & Jackendoff, R. (1993). *Generative Theory of Tonal Music*.
- Locke, D. (1982). Principles of Offbeat Timing and Cross-Rhythm in Southern Eve Dance Drumming. *Ethnomusicology*. <https://doi.org/10.2307/851524>
- 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., Large, E. W., & Palmer, C. (2011). Temporal coordination and adaptation to rate change in music performance. *Journal of Experimental Psychology: Human Perception and Performance*, 37(4), 1292–1309. <https://doi.org/10.1037/a0023102>
- Loehr, J. D., & Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research*. <https://doi.org/10.1007/s00221-006-0760-8>
- Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *The Quarterly Journal of Experimental Psychology*. <https://doi.org/10.1080/17470218.2011.603427>
- Lomax, A., & Berkowitz, N. (1972). The Evolutionary Taxonomy of Culture. *Science*, 177(4045), 228–239. <https://doi.org/10.1126/science.177.4045.228>
- Lucas, G. (2002). Musical Rituals of Afro-Brazilian Religious Groups within the Ceremonies of Congado. *Yearbook for Traditional Music*, 34, 115. <https://doi.org/10.2307/3649191>
- Lucas, G., Clayton, M., & Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musicology Review*, 6(2), 29.
- Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology: Human Perception and Performance*, 27(2), 411–422. <https://doi.org/10.1037/0096-1523.27.2.411>
- Mardia, K. V., & Jupp, P. E. (2000). *Directional Statistics*.

- Mates, J., Radil, T., & Pöppel, E. (1992). Cooperative tapping: Time control under different feedback conditions. *Perception & Psychophysics*, 52(6), 691–704. <https://doi.org/10.3758/BF03211706>
- McEllin, L., Knoblich, G., & Sebanz, N. (2018a). Distinct kinematic markers of demonstration and joint action coordination? Evidence from virtual xylophone playing. *Journal of Experimental Psychology: Human Perception and Performance*, 44(6), 885–897. <https://doi.org/10.1037/xhp0000505>
- McEllin, L., Knoblich, G., & Sebanz, N. (2018b). Imitation from a joint action perspective. *Mind & Language*, 33(4), 342–354. <https://doi.org/10.1111/mila.12188>
- McNamara, M. (2017, November 15). The Speed Trap: How To Stop Rushing the Music. Retrieved June 15, 2018, from Dance Spirit website: <https://www.dancespirit.com/how-to-stop-rushing-the-music-and-keep-the-tempo-2509515224.html>
- Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature*, 414(6859), 69–73. <https://doi.org/10.1038/35102060>
- Merrill-Mirsky, C. (1986). Girls' Handclapping Games in Three Los Angeles Schools. *Yearbook for Traditional Music*, 18, 47. <https://doi.org/10.2307/768519>
- Michael, J., & Wolf, T. (2014). Why Apply a Hierarchical Predictive Processing Framework to Music Perception and Performance ? *Empirical Musicology Review*, 9(3), 177–182.
- Miton, H., Claidière, N., & Mercier, H. (2015). Universal cognitive mechanisms explain the cultural success of bloodletting. *Evolution and Human Behavior*, 36(4), 303–312. <https://doi.org/10.1016/j.evolhumbehav.2015.01.003>
- Miton, H., & Mercier, H. (2015). Cognitive Obstacles to Pro-Vaccination Beliefs. *Trends in Cognitive Sciences*, 19(11), 633–636. <https://doi.org/10.1016/j.tics.2015.08.007>
- Mountjoy, M. (2009). Injuries and Medical Issues in Synchronized Olympic Sports: *Current Sports Medicine Reports*, 8(5), 255–261. <https://doi.org/10.1249/JSR.0b013e3181b84a09>
- Naveda, L., Gouyon, F., Guedes, C., & Leman, M. (2011). Microtiming Patterns and Interactions with Musical Properties in Samba Music. *Journal of New Music Research*, 40(3), 225–238. <https://doi.org/10.1080/09298215.2011.603833>
- Novembre, G., Varlet, M., Muawiyath, S., Stevens, C. J., & Keller, P. E. (2015). The E-music box: an empirical method for exploring the universal capacity for musical production and for social interaction through music. *Royal Society Open Science*, 2(11), 150286.

- Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, 108(52), 20947–20952. <https://doi.org/10.1073/pnas.1108155108>
- Okano, M., Shinya, M., & Kudo, K. (2017). Paired Synchronous Rhythmic Finger Tapping without an External Timing Cue Shows Greater Speed Increases Relative to Those for Solo Tapping. *Scientific Reports*, 7(March), 43987. <https://doi.org/10.1038/srep43987>
- Ölveczky, B. P., Andalman, A. S., & Fee, M. S. (2005). Vocal Experimentation in the Juvenile Songbird Requires a Basal Ganglia Circuit. *PLoS Biology*, 3(5), e153. <https://doi.org/10.1371/journal.pbio.0030153>
- Palmer, C. (1989). Mapping Musical Thought to Musical Performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15(12), 331–346.
- Palmer, C. (1996a). Anatomy of a Performance: Sources of Musical Expression. *Music Perception: An Interdisciplinary Journal*, 13(3), 433–453. <https://doi.org/10.2307/40286178>
- Palmer, C. (1996b). On the Assignment of Structure in Music Performance. *Music Perception: An Interdisciplinary Journal*, 14(1), 23–56. <https://doi.org/10.2307/40285708>
- Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language*. <https://doi.org/10.1006/jmla.2000.2780>
- Pecenka, N., & Keller, P. E. (2009a). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*. <https://doi.org/10.1111/j.1749-6632.2009.04785.x>
- Pecenka, N., & Keller, P. E. (2009b). The relationship between auditory imagery and musical synchronization abilities in musicians. *ESCOM 2009: 7th Triennial Conference of European Society for the Cognitive Sciences of Music*.
- Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*. <https://doi.org/10.1007/s00221-011-2616-0>
- Pieslak, J. (2007). Society for Music Theory Re-casting Metal: Rhythm and Meter in the Music of Meshuggah. *Source: Music Theory Spectrum*, 29(2), 219–245. <https://doi.org/10.1525/mts.2007.29.2.219>
- Pollard, H. F. (1968). Time delay effects in the operation of a pipe organ. *Acta Acustica United with Acustica*, 20(4), 189–199.

- Poudrier, È., & Repp, B. H. (2013). Can Musicians Track Two Different Beats Simultaneously? *Music Perception: An Interdisciplinary Journal*. <https://doi.org/10.1525/mp.2013.30.4.369>
- Pressing, J. (1999). The referential dynamics of cognition and action. *Psychological Review*, 106(4), 714–747. <https://doi.org/10.1037/0033-295X.106.4.714>
- Pressing, J. (2002). Black Atlantic Rhythm: Its Computational and Transcultural Foundations. *Music Perception*, 19(3), 285–310.
- Ragert, M., Schroeder, T., & Keller, P. E. (2013). Knowing too little or too much: The effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Frontiers in Psychology*, 4(June), 368. <https://doi.org/10.3389/fpsyg.2013.00368>
- Ramenzoni, V., Riley, M. a, Davis, T., Shockley, K., & Armstrong, R. (2008). Tuning in to another person's action capabilities: perceiving maximal jumping-reach height from walking kinematics. *Journal of Experimental Psychology. Human Perception and Performance*, 34(4), 919–928. <https://doi.org/10.1037/0096-1523.34.4.919>
- Ravnani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-016-0007>
- Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A Meta-Analysis. *Zeitschrift Fur Psychologie / Journal of Psychology*. <https://doi.org/10.1027/2151-2604/a000252>
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88(2), 622–641. <https://doi.org/10.1121/1.399766>
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969–992. <https://doi.org/10.3758/BF03206433>
- Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 57(January 2016), 499–521. <https://doi.org/10.1080/02724980343000369>
- Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science*, 27(3), 423–456. <https://doi.org/10.1016/j.humov.2008.02.016>

- Repp, B. H., Knoblich, G., & Knoblich, G. (2004). Perceiving Action Identity: How Pianists Recognize Their Own Performances. *Source: Psychological Science*, 15(9), 604–609.
- Repp, B. H., London, J., & Keller, P. E. (2012). Distortions in Reproduction of Two-Interval Rhythms: When the “Attractor Ratio” Is Not Exactly 1:2. *Music Perception: An Interdisciplinary Journal*, 30(2), 205–223. <https://doi.org/10.1525/mp.2012.30.2.205>
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403–452. <https://doi.org/10.3758/s13423-012-0371-2>
- 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>
- Rink, J. (2002). *Musical performance: a guide to understanding*. Retrieved from <https://dx.doi.org/10.1017/CBO9780511811739>
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27, 169–192.
- Robert B. Zajonc. (1965). Social Facilitation. *Science, New Series*, 149(3681), 269–274.
- Rzeszutek, T., Savage, P. E., & Brown, S. (2012). The structure of cross-cultural musical diversity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1733), 1606–1612. <https://doi.org/10.1098/rspb.2011.1750>
- Sabu, S., Curioni, A., Vesper, C., Sebanz, N., & Knoblich, G. (in prep.). *How does a Partner's Motor Variability Affect Joint Action?*
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112(29), 8987–8992. <https://doi.org/10.1073/pnas.1414495112>
- Savage, P. E., Tierney, A. T., & Patel, A. D. (2017). Global Music Recordings Support the Motor Constraint Hypothesis for Human and Avian Song Contour. *Music Perception: An Interdisciplinary Journal*, 34(3), 327–334. <https://doi.org/10.1525/mp.2017.34.3.327>
- Schmidt, R C, Bienvenu, M., Fitzpatrick, P. a, & Amazeen, P. G. (1998). A comparison of intra- and interpersonal interlimb coordination: coordination breakdowns and coupling strength. *Journal of Experimental Psychology. Human Perception and Performance*, 24(3), 884–900. <https://doi.org/10.1037/0096-1523.24.3.884>

- Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 227–247. <https://doi.org/10.1037/0096-1523.16.2.227>
- Schmidt, Richard C., & Richardson, M. J. (2008). Dynamics of Interpersonal Coordination. In *Coordination: Neural, Behavioral and Social Dynamics* (pp. 281–308). https://doi.org/10.1007/978-3-540-74479-5_14
- Schmidt, Richard C., & Turvey, M. T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics*, 70(4), 369–376. <https://doi.org/10.1007/BF00200334>
- Schoonderwaldt, E., & Altenmüller, E. (2014). Coordination in Fast Repetitive Violin-Bowing Patterns. *PLOS ONE*, 9(9), e106615. <https://doi.org/10.1371/journal.pone.0106615>
- Schulze, H.-H., Cordes, A., & Vorberg, D. (2005). Keeping Synchrony While Tempo Changes: Accelerando and Ritardando. *Music Perception*, 22(3), 461–477. <https://doi.org/10.1525/mp.2005.22.3.461>
- Schwarz, K. R. (1980). Steve Reich: Music as a Gradual Process: Part I. *Perspectives of New Music*, 19(1/2), 373. <https://doi.org/10.2307/832600>
- Scott-Phillips, T., Blanke, S., & Heintz, C. (2018). Four misunderstandings about cultural attraction. *Evolutionary Anthropology: Issues, News, and Reviews*, 27(4), 162–173. <https://doi.org/10.1002/evan.21716>
- 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>
- Semjen, A., Schulze, H. H., & Vorberg, D. (2000). Timing precision in continuation and synchronization tapping. *Psychological Research*, 63(2), 137–147. <https://doi.org/10.1007/PL00008172>
- Sternad, D., Dean, W. J., & Newell, K. M. (2000). Force and timing variability in rhythmic unimanual tapping. *Journal of Motor Behavior*, 32(March), 249–267. <https://doi.org/10.1080/00222890009601376>

- Summers, J. J., Davis, A. S., & Byblow, W. D. (2002). The acquisition of bimanual coordination is mediated by anisotropic coupling between the hands. *Human Movement Science*. [https://doi.org/10.1016/S0167-9457\(02\)00151-3](https://doi.org/10.1016/S0167-9457(02)00151-3)
- Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, 37(5), 343–349. <https://doi.org/10.1016/j.evolhumbehav.2016.02.004>
- Taylor, S. A. (2003). Ligeti, Africa and Polyrhythm. *The World of Music*, 45(2), 83–94.
- Taylor, S. A. (2012). Hemiola, Maximal Evenness, and Metric Ambiguity in Late Ligeti. *Contemporary Music Review*, 31(2–3), 203–220. <https://doi.org/10.1080/07494467.2012.717362>
- Thomson, M., Murphy, K., & Lukeman, R. (2018). Groups clapping in unison undergo size-dependent error-induced frequency increase. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-017-18539-9>
- Tognoli, E., Lagarde, J., DeGuzman, G. C., & Kelso, J. A. S. (2007). The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(19), 8190–8195. <https://doi.org/10.1073/pnas.0611453104>
- Trehub, S. E., & Trainor, L. (1998). Singing to Infants: Lullabies and Plat Songs. In C. K. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), *Advances in Infancy Research* (Vol. 12).
- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, 16(3), 285–295. [https://doi.org/10.1016/0163-6383\(93\)80036-8](https://doi.org/10.1016/0163-6383(93)80036-8)
- Van Der Steen, M. C. M., & Keller, P. E. (2013). The Adaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, 7(June), 253. <https://doi.org/10.3389/fnhum.2013.00253>
- Van Der Steen, M., Jacoby, N., Fairhurst, M. T., Keller, P. E., & Van Der Steen, M. C. (2015). Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Research*, 1–22. <https://doi.org/10.1016/j.brainres.2015.01.053>
- van Rijn, H. (2016). Accounting for memory mechanisms in interval timing: A review. *Current Opinion in Behavioral Sciences*. <https://doi.org/10.1016/j.cobeha.2016.02.016>

- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*. <https://doi.org/10.1016/j.neunet.2010.06.002>
- Vesper, C., Schmitz, L., Safra, L., Sebanz, N., & Knoblich, G. (2016). The role of shared visual information for joint action coordination. *Cognition*, 153, 118–123. <https://doi.org/10.1016/j.cognition.2016.05.002>
- Vesper, C., Schmitz, L., Sebanz, N., & Knoblich, G. (2013). Joint Action Coordination through Strategic Reduction of Variability. *Annual Meeting of the Cognitive Science Society (CogSci)*, 1522–1527.
- Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211(3–4), 517–530. <https://doi.org/10.1007/s00221-011-2706-z>
- Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2012). 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>
- Vicary, S., Sperling, M., von Zimmermann, J., Richardson, D. C., & Orgs, G. (2017). Joint action aesthetics. *PLOS ONE*, 12(7), e0180101. <https://doi.org/10.1371/journal.pone.0180101>
- Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. *Handbook of Perception and Action*, 2, 181–262.
- Welsh, T. N., Wong, L., & Chandrasekharan, S. (2013). Factors that affect action possibility judgments: The assumed abilities of other people. *Acta Psychologica*, 143(2), 235–244. <https://doi.org/10.1016/j.actpsy.2013.04.003>
- Wilson, M., & Knoblich, G. (2005). The Case for Motor Involvement in Perceiving Conspicifities. *Psychological Bulletin*, 131(3), 460–473. <https://doi.org/10.1037/0033-2909.131.3.460>
- Wing, A. M., & Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, 14(1), 5–12. <https://doi.org/10.3758/BF03198607>
- Wing, A. M., & Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Perception & Psychophysics*, 13(3), 455–460. <https://doi.org/10.3758/BF03205802>

- Wöllner, C., & Keller, P. E. (2017). Ch26 - Music With Others: Ensembles, Conductors, and Interpersonal Coordination. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition* (pp. 313–324). Routledge.
- Wolpert, D., & Kawato, M. (1998). Multiple Paired Forward and Inverse Models for Motor Control. *Neural Networks*, *11*, 1317–1329.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 593–602.
<https://doi.org/10.1098/rstb.2002.1238>
- Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, *17*(2), 312–321.
<https://doi.org/10.1038/nn.3616>
- Zago, M., Pacifici, I., Lovecchio, N., Galli, M., Federolf, P. A., & Sforza, C. (2017). Multi-segmental movement patterns reflect juggling complexity and skill level. *Human Movement Science*, *54*, 144–153. <https://doi.org/10.1016/j.humov.2017.04.013>
- Zamm, A., Pfordresher, P. Q., & Palmer, C. (2015). Temporal coordination in joint music performance: effects of endogenous rhythms and auditory feedback. *Exp Brain Res*, *233*, 607–615. <https://doi.org/10.1007/s00221-014-4140-5>
- Zamm, A., Wang, Y., & Palmer, C. (2018). Musicians' Natural Frequencies of Performance Display Optimal Temporal Stability. *Journal of Biological Rhythms*, *33*(4), 432–440.
<https://doi.org/10.1177/0748730418783651>

ACKNOWLEDGEMENT

TO EXTERNAL FUNDING AGENCIES

CONTRIBUTING TO PHD DISSERTATIONS

Name of doctoral candidate: Thomas Wolf

Title of dissertation: *Joint Music-Making and Temporal Coordination in Joint Action*

Name of supervisor(s): Gunther Klaus Knoblich, Natalie Sebanz

External funding agency: **European Research Council**

Acknowledgement:

This research has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) under **grant agreement No 616072 – JAXPERTISE**.

This research has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) under **grant agreement No 609819 – SOMICS**.

