Structuring social relationships

Music-making and group identity

Rebecca Naomi Whiteman
Corpus Christi College

This thesis is submitted for the degree of Doctor of Philosophy
April 2020

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

Structuring social relationships: Music-making and group identity

Rebecca Naomi Whiteman

This thesis is about groups and their boundaries: how we bond with some people, but are separated from others who do not belong. It is also about social interaction - the building-blocks of this group identity. In particular, I investigate music's role in our social landscape. Making music together is a powerful way of establishing and structuring these relationships; I argue that it can bring people together, but can also reinforce the divisions between them.

First, I present a new synthesis, drawing on relevant literature about our capacity for sociality, analyses of social interaction, and a history of the research on social groups. I outline a helpful framework by which to understand different forms of social engagement, depending on the nature of the interaction *goal*; I also clarify the concepts of interdependence and categorisation as distinct processes in group formation. Following this, I suggest that when our interaction is primarily affiliative, or relational, in goal (with little or no external goal focus), then it brings people together via relationships of interdependence; when we aim to communicate something more precisely (i.e. we have an external goal), then the need to maintain our common ground might instead form the basis for social division via categorisation.

Second, I report an initial empirical project which tests some of these predictions. My experiments show that music-making enhances affiliation, especially when there is no external goal focus. Adding a goal contributes to social division - affiliation on the basis of common team membership - but only when the interaction task was a success. When it was not successful, or caused embarrassment for those involved, participants instead seem to distance themselves from any associated group identity. This experimental work is supported by video analyses, in which I show different patterns of behaviour in interaction with and without an external goal.

This thesis is an important starting point in understanding the nature both of social groups and of music. It highlights the potential for music-making to structure our social world through either affiliation or division.

For my parents

Acknowledgements

This thesis would not exist without the support and friendship of many in Cambridge and beyond. I would like to thank my supervisor, Professor Ian Cross, for all of his wisdom, guidance, encouragement and reassurance for the past three and a half years. Thanks are due to all who have formed the community at the Centre for Music and Science, especially Juan Pablo Robledo del Canto, Erica Cao, Satinder Gill, Sarah Hawkins and Xi Zhang. I thank them all, together with my advisors, Jeremy Begbie and Antonio Rodrigues, for helping me to grow as a researcher, musician and scientist.

I am indebted to the many participants without whom I could not have developed and run my experiments. I am grateful also to those who helped me with all the logistics on these occasions: Nicholas Lee, Carla Cao, Louisa Denby, Gabriele Ceccheti, and Kaamya Varagur. I would also like to thank Miles Eastwood and Mustafa Beg, whose technical expertise helped greatly in setting up my video analysis work.

Throughout my PhD, I have enjoyed many hours of making music with St. John's Voices, and with many other friends through the University's Instrumental Award Scheme for chamber music. I am thankful for these opportunities. I have been blessed with a few very good friends, and am thankful especially for Jonathan.

My family have provided endless encouragement and support; thank you to my Mother and Father.

Note

I have used '' for simple quotations from a text. I have used "" for quoting something second-hand (i.e. something quoted within a text), or a colloquial turn of phrase.

Table of contents

Introduction

Humans have always been sociable animals. Every day, we engage with others; we think, speak or act according to the norms of our community. We are interdependent - we learn from others, acquire food with them, and rely on their help in bringing up our children. We each build a sense of our own identity - both as individuals, with our quirks and peculiarities, and as members of society, wanting to feel part of something bigger. This thesis is about how we relate to others; specifically, about the everyday, face-to-face, small-scale interactions we have with others to enact these social relationships.

We interact with others in all sorts of ways - passing people in the street, standing together in a lift, holding a conversation or committing to a lifelong partnership. Musical interaction is an important part of this social world. Playing in an orchestra, listening to a playlist, attending a concert or "jamming" with friends are all places where we build connections with others. This work concerns the role of music-making in forming social groups; specifically, how does music - as a form of social interaction - help us to *bond* with others, and also to form our sense of social *identity*? And also - how does music relate to other forms of interaction, such as speech?

Answering these questions has been a multi-faceted process, drawing on insights from many different fields of study - observational, theoretical, ethnographic, and empirical. Understanding humans - and especially *groups* of humans - requires knowledge of our ontogeny (and phylogeny); understanding of our social cognitive capacities, from the underlying neuroscience to their behavioural manifestations; and insights into how our particular cultural context influences how we expect to engage with others, in speech and music. As such, this work does not fall easily into the discipline of musicology, or even social psychology - rather, it involves the convergence of these with anthropology, evolutionary psychology, ethnomusicology, and sometimes sociology.

The project has two main parts: first, a literature review, synthesising research from a variety of fields (Chapters 1 to 4); and second, an empirical project, using novel experimental methods to offer initial insights into a perhaps ongoing investigation (Chapters 5 and 6). I shall use this introduction to present some definitions, before providing a more detailed overview of the thesis as a whole.

Understanding musicality

I start with the particularly thorny issue of how to define music, beginning with the commonplace (Western) idea of music as sound, as an artwork to listen to or even to "consume". I argue that this is far too narrow a definition for our purposes - where the broader, social function of music is

concerned. Instead, I present a basis for thinking of music as an interactive activity - a (potentially universal) *practice*, rather than a specific product or "piece". Once this is in place, we can consider the social (and even survival) consequences of music-making - including, of course, the potential efficacy of music in forming group identity.

The Western music scene

If we were to approach an unsuspecting individual and ask how they define "music", they are likely to suggest that it is primarily an auditory phenomenon: perhaps "organised sound", a backing track, or something they enjoy listening to in their leisure time. The first entry for the Oxford English Dictionary defines 'music' as:

> The art or science of combining vocal or instrumental sounds to produce beauty of form, harmony, melody, rhythm, expressive content, etc. ("music", OED, 2003)

Music for us is a pleasing combination of sounds - or at least something to which we might ascribe aesthetic or "artistic" value. On special occasions, we pay to attend concerts, to see and appreciate a particular performer or piece; we sit quietly and listen to how the pattern unfolds (and complain when the presence of other audience members spoils this experience; see, e.g. Molleson, 2013).

Now that recorded music is easily available, especially over the internet, music can be literally all around us. People carry with them their own musical world, playing through noise-cancelling headphones as they walk, jog or commute. Music has become a commodity (see Katz, 2010; also Sterne, 2006) - we pay for the right to "own" or curate that world. We can build and share playlists which give a sense of who we are and what we enjoy (see e.g. North & Hargreaves, 1999). At the click of a button, we can regulate our emotional state, using music to lift our spirits or evoke a sense of nostalgia (e.g. Saarikallio & Erkkilä, 2007). We can use this to influence others - perhaps even against their better judgement (e.g. Ziv et al., 2007). Thus, music has become a sonic resource - something we draw on as individuals, but also of commercial value for businesses to exploit.

Our musical tastes signify our own social identities, while musical genres more widely often align with socio-economic or national stereotypes (see e.g. Bryson, 1996, or Shevy, 2008). Music can have political import, in its association with either minority groups (like the subversive anti-government sentiment in popular music of 1980s Brazil) or oppressive forces in society (think of the Nazi adoption of Wagner's rousing melodies in public displays of power - an association which remains in Israel today).

Locating the survival function of sound

In these ways, music is socially and politically significant. However, attempts to understand what underlies this, and the potential origins of this behaviour, from a psychological, biological, even evolutionary perspective, all highlight a puzzle. Namely, music does not have obvious survival value, and so does not fit easily into traditional models of natural selection. As noted by Charles Darwin himself:

> As neither the enjoyment nor the capacity of producing musical notes are faculties of the least direct use to man in reference to his ordinary habits of life, they must be ranked amongst the most mysterious with which he is endowed. (Darwin, 1871)

Perhaps, music is a by-product of other, more useful capacities. Pinker (1997) famously characterised it as 'auditory cheesecake' - our enjoyment of music is a pleasurable accident, drawing on useful capacities for language, emotional signalling and auditory perception, but without any value in its own right (also compare Roederer, 1984, and Hauser & McDermott, 2003). Others have afforded music slightly more significance: Miller (2000) proposed a costly signalling account, arguing that music could be adaptive in sexual selection - performers demonstrate their fitness through impressive sonic displays. However, given the lack of sexual dimorphism we see in musical performance, this account seems somewhat limited. More recently, Patel (2010) suggested that music is a 'transformative invention' or 'technology': it was a chance occurrence, a one-off, which then happened to have biological, social significance, and impacts on our cultural landscape - in the same way as e.g. fire.[1]

These accounts, characterising music as a by-product, or something that emerged by accident, without biological significance or roots (apart from, perhaps, Miller's account), do not sit well with our everyday experience of music. We report strong emotional effects (Gabrielsson, 2001), with neurobiological correlates in central reward-processing regions (Blood & Zatorre, 2001); we exhibit early proclivity for rhythm perception (Winkler et al., 2009); and close ties between music and ritual are apparent across most (if not all) human societies (see summary of possible musical universals reviewed by Trehub et al., 2015). Music is likely older, with deeper roots than they acknowledge: the 'Isuritz Pipes' exemplify sophisticated musical instruments (with precise placing of finger-holes, and possibly an off-centre playing position), dating back to the Aurignacian (d'Errico et al., 2003). The more recent discoveries of bone and ivory flutes in southwestern Germany corroborated this: by 35

---

[1] Hodgson and Verpooten (2014) also proposed a costly signalling account of music, but with its adaptive value limited to a cultural sphere (see Chapter 1 for an evaluation of the validity of the distinction between biology and culture): within models of cultural evolution, music and other ritual practices were selected due to their value in displaying the fitness of the society or group as a whole (see also Hodgson 2017).

000 years ago, our Europeans ancestors already had a well-developed musical tradition (Conard et al., 2009).

The discrepancy we find between music's social (and personal) significance, its ancient heritage, and the suggestion that it is no more than a by-product with no biological significance in its own right, might be the result of a problematic starting point: our definition of the phenomenon we seek to explain. If we understand music purely as sound, then it is harder to understand why this might be so important. However, if we have reason to expand this concept, and to think of music in broader terms, then we might discover more productive avenues to address questions of the social, psychological role of music in group bonding and identity.

Music beyond the West: expanding our definition

The idea of music as a commodity, or an auditory art-form, is widespread now but had very specific, Northern European, origins. The industry of copying manuscripts, a rising Medieval trade proliferating particular versions of church liturgy and providing objects of prestige for the wealthy, might be an early example of this thinking. Notated music here became part of the economy, even though what was written down was only an extract from a rich, flexible orally-transmitted tradition (see overview in Cross, forthcoming). It was much later that the idea of "the music itself" began to surface. Musical engagement (influenced by figures such as Immanuel Kant) involved 'disinterested' listening, as we contemplate the beautiful (even sublime) in the musical form (see Rohlf, 2018, for an overview of Kant's philosophy). A musical work was unique, and of great cultural and artistic value; the master composer who created it sometimes became an almost mythical being - a "musical genius" whose oeuvre transcended time to become part of a virtual museum of musical artefacts (see Goehr, 1992). Our approach to music - in the West - is heavily influenced by this philosophy. When we attend a concert or download an MP3, we are buying music as if it were an artefact, an object with an identity of its own.

However, this perspective does not capture the focus of most musical practice, in the West and elsewhere. Children sing together with no training; people dance or clap along, engaging actively with music more often than they try 'disinterested listening'. Thomas Turino (2008) proposed a helpful distinction between our product-focussed, or presentational, musical culture and more *participatory* fields of music-making. The former encompasses the idea of music as an artefact, or even commodity, produced by one group of people, 'the artists', for a different group, 'the audience'. This is what we are used to when we attend a concert, for example, or study the Beethoven string quartets. The latter definition of music instead prioritises '*the doing*' (original emphasis), where *all* participants are contributing and have a role to play in the musical, social, event

(Turino, 2008: 26, 28).[2] This is clear for the typical church congregation (where participants seldom *perform*); however, even considering a classical concert, we miss half the story if we consider only the musical 'work' which is on display. The actions of the audience, tapping their feet, recognising or humming along to a melody, even the decision to go to and participate in that specific musical event, is part of the richer activity which Christopher Small (1998) calls 'musicking'.

We should also consider how music is done and thought about beyond what is familiar in the Western world - not in stylistic terms but in the very definition of what the behaviour involves (Cross, 2012a). A crucial dimension of this is *practice*:

> The vast majority of the philosophical literature has treated music as aesthetic object from the perspective of a listener - as sound, as score, as, at best, traces of musical behaviours. Very little attention has been devoted to accounting for music in action - to accounting for how a capacity to make music is acquired and how it is exercised, as well as what constitutes such a capacity in the first place. ... In any case, the majority of non-Western societies have tended to rely on informal, mimetic, enculturative processes of music learning...just as is the case across contemporary Western societies outwith the academy. (Cross & Tolbert, forthcoming)

The freedom and uncertainty embodied in many improvised musical traditions provide a good demonstration of this: free collective jazz performance relies primarily on the spontaneous, mutual adjustment of co-performers, drawing only on the 'implicit mental model' which comes with familiarity with one's partners at the outset, rather than any musical "work" (Canonne & Aucouturier, 2015, p. 545; see also Wilson & MacDonald, 2016). However, the same applies even to scripted performances: music is realised in the performance event, where people align their actions and understandings with one another in real time (Cook, 2017). Recently, approaches in music psychology have situated this 'online' processing in the context of embodied and enactive approaches to music cognition, more generally: the creation of musical meaning (whether scripted or spontaneous) is socially-situated, *emergent* from the collective actions of those involved (Bishop, 2018; see also van der Schyff et al., 2018, and Moran, 2014, on the value of a process- over product-oriented approach).

Ethnomusicologists have acknowledged this for a long time. In the middle of the last century, Alan Merriam (1964) advocated awareness and discussion of the social context of musical practices, not a sole focus on formal features of the "music itself". He proposed a three-stage model of music - the concept, behaviour, and, only then, how these are manifested in sound. This *anthropology* of music was taken up later in the work of John Blacking, who noted:

---

[2] In an essay, Turino (2009) links these fields of practice to different social habits; he argues that the participatory field contributes more to a sense of community, where everyone can and *should* take part in the performance.

> Music is a synthesis of cognitive processes which are present in culture and in the human body: the form it takes, and the effects it has on people, are generated by the social experiences of human bodies in different cultural environments. Because music is humanly organized sound, it expresses aspects of the experience of individuals in society. (Blacking, 1973: 89)

In his ethnography of Venda children's songs, Blacking (1967) describes how their "musical" characteristics or form can be understood in the light of the children's everyday *social* experiences; he identifies, for example, the influence from Venda language and national dance on what and how those children sing. A decade later, Anthony Seeger (1987) further developed this approach. He described how music is, in fact, more than just a *manifestation* of cultural context; rather it is 'part of the very construction and interpretation of social and cultural relationships and processes' (Seeger, 1987: xiv). The musical traditions and practices of the Suyá (a hunter-gatherer tribe in the Brazilian Amazon) provide his case study.[3] Musical performance here is not a single event, isolated or set apart from daily routines; rather, music-making is closely associated with social roles, identities, and ceremonies central to the life of the community: 'Suyá society was an orchestra, its village was a concert hall, and its year a song' (Seeger, 1987: 140). In this way, the *practice* of musical interaction is a key part of social relationships, even the functioning of society as a whole.

In this thesis, I am considering the roots of music-making, its possible role in social group identity, and the psychological, biological, evolutionary mechanisms underpinning this. For these purposes, we need to consider human universals over cultural peculiarities, and so require a definition of music which is cross-culturally valid (the traditional, rather Western-centric, understanding of music as "organised sound" is too limited for these questions).[4] Music as *practice* (i.e. music-making), in one form or another, might be a human universal. In their cross-cultural review, Trehub et al. (2015) note convergences in various practices and definitions of musical behaviour: it is often central to society, to social and ritual norms, enacting social bonds from intimate caregiving up to large-scale gatherings. This approach aligns with the ethnographic data, and offers a perspective relevant to exploring face-to-face, social relationships and identity.

Music-making: a working definition

Therefore, I define musical performance, or music-making, primarily as a form or mode of interaction, a social, performative behaviour creating a context in which we can relate to one or

---

[3] Note that 'Suyá' is an external label; their self-denomination is 'Kĩsêdjê' (Everett, 2012: 20).
[4] Much of the psychological literature has come up against this very problem, as many experimental results recruit exclusively undergraduate students at American Universities (Henrich et al., 2010). Heine and Norenzayan (2006) call for cross-cultural differences to be taken into account, and propose methods for investigating psychological universals (also Norenzayan & Heine, 2005).

many others. Cross (2012b) talks about how music lies within a single 'communicative toolkit', which supports spoken, sung, and all other ways of relating to and with others.

Specifically, musical interaction can be seen as 'embodying, entraining, and transposably intentionalising time in sound and action' (Cross, 2003). It is likely to involve some degree of interpersonal synchrony, as people tend to *entrain* with one another, moving (and thus sounding) together in time (Clayton et al., 2005). When we isolate the sound dimension, we see that it generally involves discrete pitches, unequally spaced, perhaps with a preference for consonant over dissonant intervals (although see McDermott et al., 2016). Importantly, however, this communicates some sort of emotional expression - an *embodied* property rooted in animal signalling behaviour (see literature review in Juslin & Laukka, 2003). Although we can (and do) abstract this musical sound from its social, communicative context, even the audio signal alone maintains those 'human-like properties' which demonstrate its embodied and interactive origins (Morley, 2013: 315; see also Molnar-Szakacs & Overy, 2006). In these ways, the "musical" types of communicative, intentional behaviour can be a particularly flexible, and easy, way of engaging with others. This is the music I focus on; a *behaviour* which is intricately tied to our sociality and to the performance of our communities.

Making music for survival

With this new lens, considering music-making as an activity (rather than an artefact), we can now re-address the question of its evolutionary origins. There are a number of possible biological or evolutionary accounts. For example, music could be an important signalling behaviour. While Miller's sexual selection account focussed on the possible mating advantage offered by the ability to produce music (defined in this case as an impressive sonic display), others have developed and nuanced this approach. For example, Merker (2000) explored the incidence of synchronous chorusing amongst primates and other animals (e.g. crabs); this not only attracts females, but also indicates male co-operation. Hagen and Bryant (2003) suggested that music and dance might have functioned as a multi-modal 'coalition signalling system', displaying the group's co-operation, cohesion and strength, and thus contributing to defence against neighbouring, competing tribes. Hagen and Hammerstein (2009) corroborated this with the observation that social carnivores such as primates and wolves also use co-ordinated vocalisations to aid territorial defence. Thus, the ability to produce co-ordinated musical displays offers clear selective advantages. However, these accounts by and large still focus on the "music itself" - a signal which conveys useful information to others. Instead, I am interested in the function of the musical *behaviour*.

Other work has considered the capacities behind this music-making - our musical "instinct", perhaps, or our "musicality". Fitch (2015) identified four key features of 'bio-musicality': song and vocal

learning, tool use and drumming, synchronisation, and dance. Each of these might demonstrate continuity (homologous or analogous) with other animal behaviour, such as bird and whale song, ape tool use, and various chorusing behaviours; Fitch argued that we should consider the origins, phylogeny and ontogeny, and various ecological pressures, which lie behind each of these aspects of musical behaviour - rather than considering "music" as a whole. Trainor (2015) suggested that our ability to understand auditory signals (necessary in many musical contexts) is likely to be grounded in auditory scene analysis; this capacity has clear survival value - we need to be able to recognise and respond to potentially significant sound in our surroundings. Of course, beyond our auditory capacities, music-making involves moving in time, entraining with others around us. Synchrony in sound and movement is likely to be advantageous in forming strong social bonds *within* the group - not just signalling coalition strength to outsiders (McNeill, 1995).

All of this work acknowledges that 'musicality' is a synthesis of many components, each of which is likely to have been of evolutionary, selective value. Indeed, it is important to remember that music-making is not necessarily a separate category of behaviour at all. We tend to talk about music, or song, as distinct from spoken genres (we even describe songs as comprising music and lyrics, as separate components). However, this dichotomy is not always appropriate. Consider, for example, church liturgy, or poetry, which fall somewhere in the middle. The vocal forms of the Suyá in Brazil include genres which span both (what we would call) speech and music, encompassing 'instruction', 'everyday speech', 'everybody listens speech', 'invocation' and 'song' (Seeger, 1987: 26). All of these could be said to be simultaneously spoken and sung, to different extents. Thus, we should consider music-making a very general class of behaviour, including all sorts of communicative practices (from spoken to instrumental forms, and everything in between).

Iain Morley discusses the emergence of musical behaviour in this broad sense, considering the origins of contemporary hunter-gatherer musical ritual alongside those of mother-infant interaction. He provides a comprehensive review of the biological, cognitive capacities underpinning all sorts and components of interactive musicality - including emotion contagion (or the collective modulation of emotion and arousal), entrainment (moving and attending together in time), and an understanding of this shared experience (through joint intentionality or Theory of Mind). He also draws on archaeological evidence to show how these various abilities emerged incrementally over a long period of time. Vocal production, whether more musical or speech-like, is tied to our emotional experience and expression, and also our ability to control and co-ordinate our movements with others. This embodied, emotional vocal control (including a larger tonal range, capacity for sustained duration and articulatory control) is likely to have older evolutionary roots than any capacity for linguistic syntax. This more ancient and holistic, multi-facetted, musical instinct has great social significance, supporting our ability to relate to and bond with others - especially important in times

of uncertainty, such as child-rearing, inter-group meetings, perhaps rallying against threats to food or other resources (Morley, 2013).[5]

Roederer had already provided a similar suggestion in his exploration of the emotional content of music (albeit defined as sound). Music:

> ... can contribute to the equalization of the emotional state of a group of listeners just as a speech may contribute to the equalization of the intellectual state (knowledge) of the audience. ... [it is] a means of establishing behavioural coherency in masses of people. In the distant past this could indeed have had an important survival value, as an increasingly complex human environment demanded *coherent*, collective actions on the part of groups of human society. (Roederer, 1984: 356, original emphasis)

Ian Cross (2012b) also considers this relational quality of music-making as central to its historic and present societal significance. He speculates that modern musical practices might have come about alongside the progressive altricialisation of our species: human infants are born relatively under-developed, thus the social context of engagement with those around them is crucial in their cognitive formation. These musical modes of engagement remain important into adulthood, as they continue to provide a flexible context to relate to others - especially in 'situations of social uncertainty'. Music remains a sort of signalling behaviour - but one where the ambiguity of its meaning, or its 'floating intentionality', is a crucial part of its social function (Cross, 1999). Cross locates the survival value of music-making not just in biological terms, but also in the formation of our capacity for cultural understanding:

> ... human musicality appears to provide co-ordinative and open frameworks for interaction, maintaining and perhaps forming, the intellectual and social flexibility that is manifested in our capacity for shared intentionality. Musicality does not give rise to social justice, but at an evolutionary timescale it is likely that it has provided space for the emergence of concepts that bear on how humans can, and ultimately perhaps should, interact. (Cross, 2005a:124)

In this thesis, I continue this investigation into the social efficacy of music-making. Music, when defined in terms of social performance or communicative, interactive practice, is of great evolutionary and sociocultural significance in the development of modern human behaviour. It is appropriate, therefore - and indeed necessary - to explore it through the lens of experimental social

---

[5] Contrast the account provided by Gary Tomlinson (2015), who also discusses the interdependence of music and speech, situated in the context of the emergence of human sociality. However, he defines music as 'the systematization of discrete pitch' (Tomlinson, 2015: 265), exploring it at a more syntactic, symbolic level (rather than as communicative practice), aligned with linguistic forms in speech. He suggests that music is in fact the 'final coalescence' of our abstract, symbolic thought (which is combined with more basic mechanisms such as entrainment; Tomlinson, 2015: 268). Here, instead, we consider musicality as a capacity and a performative context in which we form social relationships; this has clearer evolutionary roots, with more direct selective significance for individuals and groups (see discussion of the group selection debate in Appendix 4).

psychology. I use the remainder of this introduction to present some key considerations and debates within this discipline, to justify and delimit the scope offered by this approach.

Understanding sociality

The question of how music-making supports or generates a sense of group identity could be explored in a myriad of ways. We require a multi-disciplinary approach, bringing together many different insights to inform our predictions about how musical interaction might relate to group formation. I shall also contribute to this scholarship with empirical - experimental and observational - data, collected using methods devised in line with most social psychology. In this introduction, I shall review some of the issues which have shaped this discipline in particular. With this in mind, I then provide an overview of my thesis - the main theory I explore, and how I devised a suite of experiments to test my hypotheses.

Approaches to social psychology

In the past century, two broad themes have characterised debate around the approach and scope of experimental social psychology. I outline first the apparent role of psychology in establishing human universals, contrasted with an awareness of substantial cultural variability, and note how researchers have attempted to balance these considerations. Second, I explore the potential of experimental methods - whether they are limited to clearly measurable, observable phenomena, or can tap into underlying factors, documenting some aspects of our "subjective" inner experience.

Human universals or cultural specifics?

In general, psychologists aim to elucidate human universals in our way of perceiving and understanding our environment. They assume the existence of at least some human-wide, sometimes species-specific mechanisms which shape how we act - and that these can be uncovered through experimental, behavioural, or neuroscientific methods, including animal models. These mechanisms would have evolved: they should provide (directly or otherwise) some sort of survival or selective advantage in the face of ecological pressures or threats, whether from other species or through intraspecific competition with other individuals or groups.[6]

Solomon Asch (1987) wrote on the aims and methods of social psychology more specifically, highlighting the importance of considering universals. Psychologists are tasked, he said, with formulating a 'theory of man' - namely, 'a comprehensive conception of human nature' (Asch, 1987: 5). However, this is held in tension with the fact of significant cultural divergences.

---

[6] The controversy over group selection is explored in Appendix 4.

> If one wants to understand human beings it is not sufficient to analyse their individual capacities; one must study men in their setting, which is society, and observe the forces that social conditions exert upon them. Society...is the prime reality that provides goals and gives form to the character of men. Man is socially determined; to understand men we must start with society. (Asch, 1987: 16)

Social psychology should tread between these universalising and specific accounts. It has to examine the interrelationship between the 'psychological facts' and the 'social environment' (Asch, 1987: 38).

It has recently become evident that most participant samples, used for experiments into so-called universals, come from exclusively WEIRD populations (Western, Educated, Industrialized, Rich, and Democratic) - often undergraduates at American institutions (Henrich et al., 2010). Researchers' lack of awareness of their own cultural specificity - and how this has the potential to shape their performance in experiments - has limited the conclusions we can draw from this work. To address this, and to ascertain what might be universal (and rooted in our species-wide cognitive and biological capacities), we need to replicate any experimental results in other cultural contexts (see strategies set out in Norenzayan & Heine, 2005; Heine & Norenzayan, 2006). This is especially important in investigations of *social* psychology: differences in cultural context likely have particular significance when it comes to our way of relating to or interacting with others.

This task does not just involve noticing parallels across various social contexts, but actually understanding human cultural diversity - how the underpinning psychological tendencies are manifested in specific locales. This has historically been the domain of social anthropology - or, in our case, ethnomusicology: case studies shed light on particular social contexts, but also provide insights into the commonalities of human (musical) experience. In the light of this work, we can suggest possible universal traits - as in the aforementioned work by Morley (2013) and Trehub et al. (2015).[7]

Some interesting insights on this issue have also come in the guise of evolutionary psychology - specifically, theories of 'social niche construction' (see Chapter 1). Social behaviour is not just an outcome of particular (biological or cultural) evolutionary mechanisms, but rather forms the environment, exerting its own selection pressures which shape the development of norms, practices, perhaps even psychological needs and capacities (see Laland et al., 2001). For all of these reasons, understanding an aspect of our social psychology - specifically, the role our social (musical) interactions play in our group identities - requires insight not only into our human biology, but also into the concrete contexts in which we perform these processes, day-to-day.

---

[7] Earlier, and within anthropology more generally, Brown (1991) enumerated many cross-cultural universals (or near-universals). There is some controversy over whether universals should be considered among concrete practices, as Brown does, or whether they are only on the level of underlying capacities (see e.g. Ekman, 1994, vs. Russell, 1994 - cited in Norenzayan & Heine, 2005).

Internal processes or outward signs?

Another important debate has encompassed the field of psychology, as a whole. Looking back a little over a hundred years, we have seen dramatic and far-reaching shifts in the overall approach and scope of the discipline. At the start of the last century, scientists such as Watson, and later, Skinner, promoted a psychology purely based upon external behaviour. Writing in 1913, Watson argued that a focus on stimulus and response (and how one might predict another) is a more reliable, and useful, form of data - or 'observable fact' (Watson, 1913: 167) - than could be gained from any studies relying on introspection. This idea limited the scope of psychological investigations:

> The time seems to have come when psychology must discard all reference to consciousness; when it need no longer delude itself into thinking that it is making mental states the object of observation. (Watson, 1913: 163)

Instead, he proposed:

> What we need to do is to start work upon psychology, making *behavior* [sic], not *consciousness*, the objective point of our attack. (Watson, 1913: 175-176, original emphasis)

Insights could then be drawn from how behaviour changes, how habits are formed in ways predictable from the external environment (the presence of rewards or punishments, perhaps). No interpretation of this is needed - we should not propose any internal workings, mental processes or mechanisms. Indeed, no mention need be made of our (or another animal's) 'stream of consciousness' (Watson, 1913: 160) at all - a subject on which scientists up until that point had only been unhelpfully vague (see also an overview in Graham, 2019).

These ideas brought us an empirical method: a psychology which could be considered 'a purely objective, experimental branch of natural science', on a par with physics or chemistry (Watson, 1913: 176). However, when this was applied to more complex behaviour, identifying just the observables could not provide adequate insight. Notably, Skinner's radical behaviourist approach to language - a capacity, he suggested, which was acquired through simple reinforcement - attracted serious criticism:

> The magnitude of the failure of this attempt to account for verbal behavior [sic] serves as a kind of measure of the importance of the factors omitted from consideration, and an indication of how little is really known about this remarkably complex phenomenon. (Chomsky, 1959: 28)

These 'omitted' factors were the internal workings - what we might now call cognition, memory, or perhaps learning:

> One would naturally expect that prediction of the behavior of a complex organism (or machine) would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes input information and organizes its own behavior. (Chomsky, 1959: 27)

By this point, of course, scientists began to have at their disposal new models, or analogies, for how this internal structure might operate. New computing machinery was coming to the fore - initially, those systems devised by Alan Turing which were instrumental in the British war effort during the 1940s. These provided a conceptual tool to theorise what lay between the stimulus input and output behaviour - i.e. the workings of the "mind" - triggering a so-called "cognitive revolution". Thus, we could start to understand how we hold, process and respond to information as a series of connections, formed and stabilised by frequent co-occurrence - a process summarised in the well-known Hebbian maxim: "Cells that fire together wire together." The idea that we gradually acquire this network of associations (supported by the results of various priming studies) formed the basis for new models of language (e.g. Dell, 1986; see overview in Thagard, 2018). These approaches have been refined over the years, but remain plausible in the light of corresponding advances in neuroscience (e.g. Song et al., 2000).

These two broad standpoints - a focus either on behaviour or on inner processes - remain important in contemporary experimental paradigms. In a critique of social psychology, Baumeister et al. (2007) questioned an almost exclusive focus on self-report or apparently irrelevant 'finger-tapping' tasks - measures which are assumed to relate in some way to our inner attitudes or mental processes. They call for researchers to remember that the 'actual behaviour' is what we are ultimately interested in.

> Once upon a time, perhaps, psychologists observed behavior and reported what they saw, along with their theories about why it happened. The emergence of competing theories, and therefore competing explanations, led psychologists to push each other to show what happened inside the person to produce the behavior. Gradually the focus shifted on these debates about inner processes, and journals started publishing studies that made significant contributions about demonstrating inner processes. Somewhere along the way, it became acceptable to publish data on inner processes without any real behavior included at all, which eventually became the norm. (Baumeister et al., 2007: 400)

Of course, such 'finger-tapping' or more introspective measures might well provide insight into some internal mechanism; Baumeister and colleagues recommend a balanced approach - investigating both these processes and other, more relevant, external behaviours.

> Everyone would probably agree that the ideal paper would report both direct observation of behavior and measurement of inner processes that mediate and produce those behaviors. (Baumeister et al., 2007:401)

In the present project, I aim for a similar balance. This is an investigation into the relationship between certain behaviours - social, including musical, interaction - and an inner attitude or feeling, namely our sense of group identity. The latter can, however, be measured both as a behaviour - displaying preference towards the in-group, in some way - or as an underlying attitude - for example, the feeling of "connectedness" to specific others. My own empirical data include self-report and other measures of internal attitudes (including implicit measures), as well as video analysis of what my participants are actually doing as they engage with one another.

A helpful reductionism?

Psychological experiments offer the rigour and precision of statistical methods to test specific hypotheses about the effect of particular conditions or activities on an outcome. They can provide quantitative evidence that changing one variable causes a significant change in some other dimension, as indexed using certain scales or measures. However, the conditions and measurements must be carefully devised to ensure validity and generalisability, while maintaining a certain level of control. In general, therefore, all aspects of the experiment design - the research hypotheses, as well as the test conditions or activities, and the chosen outcome measures - must be informed and guided by theory. I devote a chapter to how I translate theory to practice (see Chapter 5); however, I provide a brief overview of this issue here.

In another critique of the prevalent methods in social psychology, Rozin (2001) describes how a focus on experimental precision, at the expense of broader observations and description, can become far removed from the actual phenomena in question.

> Any social laboratory experiment involves making a large number of rather arbitrary choices, including instructions to participants, the particular human and environmental situation, the selection of the manipulation and control, parameters of time and sequence, and the particular instantiation of the issue at stake. These necessary decisions entail two risks: (a) they allow for the possibility that the results will not bear on real social situations and (b) they may generalize to only a very narrow range of apparently similar experimental situations. (Rozin, 2001: 9)

Modelling musical interaction (or, indeed, conducting any sort of social interaction) in a laboratory is a task which raises these same questions of ecological validity: how can we use data on the ontogeny, phylogeny of music-making and current musical practice, to develop activities which test the effect of these activities on our social relationships? Piloting extensively the many possible instruments and instructions I could use to this end helped in identifying an experimental paradigm which could test - easily, and realistically - the effect of music-making, broadly defined. The video analysis project, reported in Chapter 6, also provided insights into what this experimental task involved.

Measuring social relationships is also a difficult task. For this, there was already an extensive empirical literature I could draw on; however, my research required a distinction to be drawn between group bonding effects, i.e. a direct feeling of connectedness towards specific others, and a sense of group or social identity, i.e. a feeling of affiliation towards others on the basis of common membership. Devising adequate measures for this purpose required an awareness of how these types of social relationships differ - and so what dimensions for explicit self-report, or other more implicit tests, would be valid ways to ascertain the degree of bonding for each.

These considerations shaped how I devised my experiments and video analysis. I present a multi-method empirical project which aimed to provide "triangulation" of evidence, which I can situate within the broader interdisciplinary research context. This is only the beginning of the empirical work necessary to test my predictions; however, it provides some precise findings which inform scholarship going forward. In addition, the methodological paradigm I have thus developed might be of more general value in furthering research on social interaction and bonding.

The structure of this thesis

Now that I have outlined the scope of this enquiry - rooted in cross-disciplinary ideas, but using the research methods of experimental social psychology - and some definitions - notably, what I mean by "music" - I can describe in more detail the aims and structure of this thesis.

Research on the social and evolutionary significance of music-making provides a basis for suggesting that musical interaction might well be important in establishing a sense of group identity. The purpose of this thesis is to investigate (a) the characteristics of musical interaction that contribute to its social efficacy, and (b) the nature of the group identity it might create. I shall argue that music is effective both at bonding people together into interdependent groups, but also instrumental in forming the category divides which limit the scope of affiliative interaction.

I begin, in Chapter 1, by outlining our social capacities - why and how we interact in the first place; these attributes have a significant role in our development and wellbeing in the ecological and cultural landscape. First, I outline the evolutionary significance of human sociality. After a brief historical survey of the application of Darwin's theory to human culture, and an outline of how this can be understood productively through the lens of social niche construction, I present evidence that prosociality and the formation of co-operative groups was crucial in our phylogeny; this provided the context for the emergence of complex cognition which, in turn, supported this environment. Second, I describe the mechanisms which underpin our *need to belong* in a social environment. Animal models provide insights into the neurochemistry of dyadic and group-level attachments, and highlight the importance of interaction in forming and maintaining these fundamental bonds. Third, I

present a summary of the complex cognitive skills which are necessary for social interaction - skills which we acquire through those attachment relationships discussed above. These include perceptual skills, such as our automatic understanding of self-other equivalence; alongside cognitive flexibility, enabling us to infer when others differ from us in belief or emotional state. Importantly, these skills are interactive - our capacity for *shared intentionality* allows us to engage with others, from establishing shared attention to social learning. An overview of these capacities provides a basis, before we consider how social interaction unfolds in practice, and how it affects our relationships with others.

In Chapter 2, I give an overview of how these capacities are enacted in the context of one-to-one spoken conversation. I discuss the role of common ground, how grounding takes place moment-to-moment through verbal and non-verbal behaviour, and the distinction within speech between transactional and relational goals. Sharing these goals can be thought of in terms of common ground: perhaps shared perspective on the physical world, shared knowledge of local conventions, or in the joint process of *face-work*. I then compare speech with musical interaction, discussing how the same goal orientations emerge in each. We can incorporate external conventions in musical interaction, just as in speech; however, music often foregrounds the phatic. These observations form the basis for a framework of interaction type - characterising musical (and other) forms of interaction as along a continuum of goal focus - from the *extrinsically-* to the *intrinsically-*oriented.[8]

Then, in Chapter 3, I provide a review of the research on social groups. While, in Chapter 2, I focussed on dyadic interaction, I now address the study of social collectives, and how they have been thought about during the course of the last century. The first half of this chapter is structured broadly as a historical overview, considering the range of approaches there have been to groups, within the discipline of social psychology. I consider, in turn, concepts of cohesion, interdependence, and entitativity; the post-war inter-group field studies by Sherif and colleagues; through to the development of the Minimal Group Paradigm and its conflicting interpretations. In the second half, I review in greater depth the controversy between theories of group formation based on interdependence or on categorisation. I discuss the distinction between these processes, but argue that we should consider them as interlinking processes which work in tandem.

In Chapter 4, I present the core of my thesis, bringing together material from both Chapters 2 and 3. I suggest that social, musical interaction contributes to group identity through both interdependence and categorisation. Specifically, I draw a parallel between those group formation processes -

---

[8] I use the terms *intrinsically-* and *extrinsically-*oriented to describe the nature of the interaction, as a whole. Intrinsically-oriented interaction privileges relational goals, whereas interaction that is extrinsically-focussed foregrounds the transactional. Most interaction comprises both relational and transactional goals, more or less simultaneously. (See Cross, forthcoming.)

interdependence and categorisation - and the continuum of social interaction - from more intrinsically- to more extrinsically-oriented engagement. Interaction which foregrounds the relational, phatic dimension facilitates face-to-face group bonding through interdependence; however, a more transactional, focus, necessitates *external* forms of common ground for the interaction to be successful (even within forms of music-making), thus limiting the scope of this interdependence and forming the basis for category divides. In this way, music-making can bring people together in forming social groups, but can also demarcate who belongs and who does not.

In Chapter 5, I introduce my empirical project, which builds upon and tests this theory. I devise three experiments; these investigate (a) the role of musical versus non-musical group activities on social group bonding; (b) the efficacy of musical group goals on both group bonding and identification with a team; and (c) the efficacy of musical interaction goals in a dyad on partner bonding and an implicit measure of team identification. I also outline a video case study, contrasting musical interaction with and without a goal on patterns of gaze and interpersonal synchrony. I describe first how I developed this methodology, before reporting my experiments and the results of my video analysis in Chapter 6. In brief, my experimental work provides initial evidence that more intrinsically-oriented interaction is indeed more efficacious for group bonding, whereas the presence of an external, transactional focus or goal can reinforce processes of categorisation. This work also highlights the role of embarrassment (perhaps a form of perceived interaction success) in influencing our group identity. My behavioural analysis from video data also demonstrates substantive differences between interaction conditions: modes of interaction which do not have an external goal generally have more interpersonal synchrony, involving more focus on mutual points of convergence than on any turn-taking structure. I conclude by summarising my main findings and suggesting avenues for future research.

Chapter 1 Social capacities

As soon as we enter the world, we are immersed in a complex and changeable landscape. We try to find order, perceiving patterns when things are predictable and consistent, responding to any unexpected events, and learning how to manipulate or change our surroundings. However, we are not just occupied with passive objects; in fact, we spend more time and attention on our dealings with other people. We are social creatures, seeking out and enjoying the company of others. These are more complicated, and potentially risky, encounters. Other people have minds of their own; we need to co-operate with them and trust that they will do the same in return.

In this chapter, I explore our social capacities - what they are, and where they came from. I start with a broad vantage point, considering the place of human sociality in the evolution of our species. The application of theories of natural selection to human social, or cultural, experience, and to social groups, have been fairly controversial; I review these difficulties, and suggest that theories of social niche construction are helpful in coupling social (or cultural) behaviours - like music-making - with their biological, evolutionary context.[9] After this, I focus on the biological basis for our social instinct, and the cognitive capacities which allow us to interact and engage with others. I bring together research on the neurochemistry of attachment, theories of neural resonance and action understanding, as well as work on our Theory of Mind, our experience of empathy, and how we can establish shared intentionality.

This provides an overview of our social capacities, answering the *why* and the *how* of social interaction, before we discuss (in Chapter 2), *what* social interaction comprises; and then (in Chapter 3) how this interaction occurs in practice, in a complex social landscape of in- and out-groups.

The evolutionary landscape

Music-making is an activity that has been significant for thousands, if not tens of thousands of years, across cultures, and for individuals' whole life-span (see e.g. Morley, 2013). To consider the role of musical interaction in our experience of living in and belonging to social groups, we need to start by exploring the evolution of our social capacities, in general. Our tendency to interact and affiliate with others, to want to belong, is something peculiarly human. I shall explore why this might have been beneficial, or adaptive, and how it might have, itself, shaped the emergence of our modern behaviour.

---

[9] Culture is a difficult term to define. Here, I treat the social and the cultural as overlapping; culture is comprised of wider social norms and shared symbolic systems, itself formed by (and shaping) face-to-face social interaction. I discuss this in more depth below. Readers should note that the relationship between the social and the cultural in anthropology and psychology has not been straightforward (see Johnson, 1994).

Evolutionary paradigms

Understanding the processes of evolution - not to mention the evolution of human capacities - has been a challenging topic throughout the past century and a half. The theory of natural selection espoused by Darwin set the agenda for subsequent work (Darwin, 1859). It has been variously applied, and used to explain all sorts of phenomena (see review in e.g. Jablonka & Lamb, 2007).

In this section, I shall summarise the various forms in which Darwinism, or neo-Darwinism, have appeared. I shall note some controversies that still plague the field; although I do not attempt to solve them, an awareness of current debates helps to develop an informed approach to the evolutionary significance of human sociality and group identity.

Understanding humans by their genes

In general, we first encounter Darwin's theory in the form generated by the 'Modern Synthesis' of the 1930s, where the idea of natural selection was combined with Mendelian genetics. Since this development, genes (specific regions of DNA, each coding for a particular trait or attribute) have been considered the primary units of selection, transmitted via reproduction. Phenotypes, or the expression of these traits in the appearance and behaviour of organisms, are determined in large part by genotypes (i.e. one's genetic make-up). Indeed, changes in small sections of DNA can cause significant changes in external features, which are then inherited by offspring (from whether we are biologically male or female, to the colour of our eyes). When diversity appears (through random genetic mutations), certain traits (i.e. genes) which fit better, or offer more adaptive benefit in the current environment, are more likely to be selected, and passed down to offspring. There is little acknowledgement here of the role of the external environment in this process - genetic variation and selection is paramount (see historical survey by Jablonka & Lamb, 2007).

Scientists have applied this approach to various social traits as well, seeking, for example, a genetic basis for antisocial behaviour (e.g. Brunner et al., 1993), or even musicality (e.g. Mariath et al., 2017). Additionally, more extensive efforts to apply this thinking to social or cultural phenomena have involved devising a different but analogous unit of selection - one that is not biological and transmitted via reproduction, but rather a cultural entity which is passed from generation to generation by social learning. Dawkins (1976) proposed that 'memes' might fulfil this role: these are replicating units which code for specific cultural traits.[10]

Alongside models of genetic evolution, therefore, we have theories of cultural evolution which account for the development of all sorts of human behaviour - from language to basket-weaving (e.g.

---

[10] See also the notion of 'conceptual genes' in Caporael & Brewer (1991).

Jordan & Shennan, 2003); of course, musical instruments and styles of performance have also been approached in this way (Tëmkin & Eldredge, 2007; Le Bomin et al., 2016).

Beyond the gene - importance of environmental factors

However, both this "nature" and our "nurture" are important in our formation; indeed, factors in our external environment influence and interact with the expression of genetic attributes. Organisms have 'developmental plasticity' (West-Eberhard, 2003): they are responsive both to their genes and to the developmental context in which these genes are expressed. Our models of evolution needed, therefore, to be expanded to encompass these insights:

> ... if recurrent phenotypes are as much a product of recurrent circumstances as they are of replicated genes, how can we accept a theory of organic evolution that deals primarily with genes? (West-Eberhard, 2003: 4)[11]

Our understanding of these epigenetic processes has grown since then. For example, maternal grooming shapes rat stress responses (Weaver et al., 2004); in humans, environmental stress can trigger the onset of schizophrenia (a genetic condition; the effects of this stress on the individual are also heritable - see Roth et al., 2009).

If genes alone cannot account for heritable, biological conditions, then it does not make sense to use genetic selection-based models to explain even more complex human behaviour in a deterministic manner.

> We cannot base an evolutionary psychology on an impoverished biology. If genetic reductionism is inadequate to account for integrated structure at the level of the organism, direct genetic explanations for complex social arrangements should be inconceivable. (Carporael & Brewer, 1991:189)

Similarly, it becomes difficult to characterise social, cultural behaviour as the expression of discrete, gene-like entities, which are subject to analogous processes of selection and transmission. Our investigation of one such cultural "trait" - namely music-making (see Introduction) - has shown that this is not a single attribute, which is passed on in discrete units, but rather a synthesis of many interlinking capacities, then performed in a rich social environment. If neo-Darwinian genetic theory does not completely explain processes of biological evolution, Dawkins' version of this modern synthesis, applied to the cultural sphere, is scarcely appropriate.

---

[11] Oyama et al. (2001) had a similar idea a couple of years earlier. In their 'developmental systems theory', they proposed that evolution occurs in 'cycles of contingency': 'development and evolution [are] processes of *construction* and *reconstruction* in which heterogeneous are contingently but more or less reassembled for each life cycle' (p. 1). Thus, multiple factors shape the expression and transmission of traits - again beyond the genetic code.

Other evolutionary approaches have been developed to account for this more complex picture. Selection might not occur just on the level of the gene (or gene equivalent); instead, different individuals, kinship groups, or even populations might be units (or 'vehicles') of selection. In other words, some genes, behaviours, or even groups, might offer higher 'fitness' and thus outcompete others in that environment. This idea - namely, multi-level selection - was formulated by Wilson and Sober (1994) as a way to defend (or temper) the more controversial theory of group selection (see Appendix 4). Although evolution still acts at the level of individuals, and genes remain the unit of inheritance (unlike in group selection proper; only individuals can reproduce and pass on their genetic material to the next generation), group-level behaviours or environments are key contributors to this process. We can explain this further using niche construction theory: this acknowledges that the selection environment (the conditions or pressures which favour one individual over another) is not unchanging: rather, it is itself formed through organisms' characteristics and behaviour. Organisms do not just respond to their environment; rather, they play an active role, changing the ecology and thus the acting selection pressures (see Odling-Smee et al., 2003). For example, when a beaver constructs a dam, this does not just provide protection from external threats (an adaptive response to ecological pressures), but this new habitat itself becomes the selective environment, changing water drainage patterns, plant ecosystems, and even potential predation risks (see Naiman et al., 1988, for an account of how beavers change their environment; and Turner, 2000, for many other examples of this sort of behaviour). This is a holistic account, involving much more than just genes.

If these theories can accommodate broader environmental factors, implicating even learned behaviours which manipulate in some way the external world, then there is more scope to consider the place of social or cultural phenomena - human behaviour which is difficult to explain at a genetic level - in the evolution of our species.

Characterising culture

Humans are a product of their biology just as much as are other animals. There are some aspects of our social lives, though, where we might hesitate to apply evolutionary theory - such as our artistic culture, our musical tastes, our morality or sense of spirituality. However, this traditional distinction between the arts and the sciences, or between our culture and biology, is not always helpful. Tim Ingold, working in social anthropology, discusses how our cultural practices - those activities we learn through socialisation or enculturation, such as a manner of walking or playing a Bach cello suite - do not operate in a separate sphere from our psychology, but rather emerge from the intersection of our sociality and biology, forming learned behaviour which is fundamentally embodied. He argues

that we need to 'dissolve the boundaries' between the social and biological disciplines, adopting a so-called 'obviation approach':

> ... the human being is not a composite entity made up of separable but mutually complementary parts, such as body, mind and culture, but rather a singular locus of creative growth within a continually unfolding field of relationships. (Ingold, 1998: 23)

To understand the place of our human sociality, or the impact of certain social activities - like musical interaction - then we need to situate them in context. This means understanding their deep-seated biological, or cognitive, roots, and their wider evolutionary functions. In what follows, I shall review a couple of theories which accommodate cultural phenomena within evolutionary psychology.

First of all, there is the theory of cultural evolution developed by Richerson and Boyd (documented in Richerson & Boyd, 2005; and earlier, Boyd & Richerson, 1985). In this, they retain some of the more simplistic, Darwinian, almost memetic approach to culture, with specific cultural traits being selectively transmitted (see a critical review by Fuentes, 2006). However, they defend their use of natural selection as a *model* - acknowledging that the actual mental representations involved, and the means of transmission by social learning, are fairly loose analogies of the genetic process (Henrich et al., 2008). Importantly, their theory does not involve cultural evolution in isolation, but rather gene-culture *coevolution*: both cultural practices and genetics influence the environment, shaping the selection pressures acting both on cultural variation (more quickly), and on genetic change (over a longer period).

> ... genes and culture [are] obligate mutualists, like two species that synergistically combine their specialized capacities to do things that neither one can do alone. ... Genes, by themselves, can't readily adapt to rapidly changing environments. Cultural variants, by themselves, can't do anything without brains and bodies. Genes and culture are tightly coupled but subject to evolutionary forces that tug behaviour in different directions. (Richerson & Boyd, 2005: 194)

One well-known example of this is the prevalence of lactose tolerance within societies with a history of dairy farming. The gene which codes for the enzyme necessary to break down lactose (beyond childhood) is relatively uncommon in humans. However, with the Neolithic advent of dairy farming in Europe, more individuals with this gene thrived in the new environment, being more likely than others to pass on their genetic make-up to the next generation. Gradually, more and more could process (and benefit from) the milk and cheese that were then available; there is now a preponderance of this gene in certain populations (but not in others; see Beja-Pereira et al., 2004).

Second, given that humans are particularly adept at social learning - a skill which enables us to work together to construct or change our environment - niche construction theories are particularly applicable. In building shelter, managing our food supply, even developing traditions of cultural,

perhaps moral, norms or expectations, we alter the selection pressures which act to give some individuals survival (and social) advantage over others (Laland et al., 2000). Certain religious traditions, for example, favour those who are more inclined to co-operate over others; these social conditions then can also enable collective efforts to change the physical landscape, diverting the water supply, for instance - as necessary in the construction of complex rice terraces in Bali (see Lansing & Fox, 2011). Even routine foraging and farming methods affect the local ecology, concentrating natural resources within a relatively small area with large population density (Rowley-Consy & Layton, 2011). These sorts of activities can have significant (and relatively rapid) impact on our surroundings, on the acting selection pressures - potentially prompting us to adapt the way we generally function and behave (see Laland et al., 2001).

To emphasise, this not only implicates our physical surroundings, but also the social or cultural environment. Our suitability to our habitat involves all of our being - not just our biology, or how able we are to fight disease or predators, but also our sociability, our ability to understand, relate to others and co-operate with them. We can find examples of this even in the contemporary world. Broadly speaking, our Western culture focusses on individuals, encouraging self-reliance or autonomy. Having an independent self-construal might be more adaptive in this individualist setting. However, elsewhere, in collectivist societies, having a sense of oneself as interdependent with others in community is more congruent, or a better fit, with the environment. This selection process, especially the interaction between individuals and their own selective environment, exemplifies *social* niche construction (Yamagishi & Hashimoto, 2016).[12]

In these ways, all of our behaviour - from our biology to our sense of what is "right" - is subject to, and forms an active part of, the selection pressures and other evolutionary forces which have shaped the development of our species as a whole. We are not just concerned with changes or pressures from the external environment, but also the corresponding neural adaptations, and developments in our cognition (which in turn influence the social or physical ecology, comprising a positive feedback loop; see Iriki & Taoka, 2012, for an example of this in tool use and the evolution of symbolic cognition). This is a context in which to understand the significance of our human sociality - including our abilities to interact with others, and to develop a sense of group belonging or identity. These capacities were adaptive within a group context, and also maintained this new social environment.

Therefore, we can (and should) attempt to understand human social and cultural behaviour in the light of its evolutionary roots. Theories of natural selection are not restricted to genetic processes,

---

[12] This is also a good example of gene-culture coevolution: these cultural systems or self-construals seem to be associated with a polymorphism for a serotonin transporter protein. The short allele, also linked to affective disorders such as anxiety, is more prevalent in collectivist cultures; however, this social niche acts as a buffer, actually reducing the level of anxiety in the population (see Chiao & Blizinsky, 2009; although, compare comment by Eisenberg & Hayes, 2011).

but in fact offer insights into the role of sociality both in response to, and in creating, our "natural" habitat. Specifically, our musical traditions and other community practices are not divorced from our biology; rather, they form a potentially integral part of the evolution of our species. In the next section, I consider the basis for this claim, discussing the emergence and the effect of our social, group behaviour, before turning to its constituent parts in the second half of this chapter.

The emergence of group behaviour

Now that we have made the case for exploring social behaviour through an evolutionary lens, we consider the biological, phylogenetic roots of our social capacities - capacities which lie behind our sense of group identity, and our enjoyment of making music together. In this short section, I shall present our current understanding of the role of our social instinct in the evolution of modern human behaviour and cognition - both its adaptive benefits in response to environmental pressures or threats, and how it created a protective social niche, providing the context which allowed an ever-expanding cultural complexity.

The evolution of prosociality

Engaging with others always relies upon the fundamental motivation to co-operate with them. Interacting, sharing information with others, helping them, or just giving up some time and energy to be with them, is primarily prosocial - we are motivated to do it, and not just because it offers us some direct advantage.

Researchers studying the evolution of human co-operation in general have approached it from a number of perspectives. Working from a traditional understanding of natural selection, investigations have identified certain "selfish" motivations or explanations for co-operation, such as kin selection, direct or indirect reciprocity. Individuals who behave prosocially are more likely to be reproductively successful, as either their kin will survive to continue the family line (as in models of inclusive fitness or kin selection), or others will see their good works and help them in return (see summary by Nowak, 2006, and early models by Trivers, 1971; and Axelrod & Hamilton, 1981). Given that we regularly help those to whom we are not related, kin selection accounts have been expanded to describe how we might use similarity, perhaps our common possession of a hard-to-fake 'tag', to direct (and limit) our altruism towards those within our wider social circle (Cohen & Haun, 2013).

In fact, co-operating with others might be beneficial for individuals *because* it temporarily serves the collective. Co-operating or collaborating with another is actually a rewarding experience in its own right - not just because of the future benefits we might get as a result (Rilling et al., 2002). Groups with more altruists are likely to be more successful, whether in defending territory or hunting, than others. We might say that social groups with norms of co-operation are likely to function more

successfully than other groups, leading to the proliferation of more co-operative tendencies in subsequent generations: this is a *group-level selection* account of prosociality (see original theory of group selection proper presented by Wynne-Edwards, 1962, and its adaptation as part of multi-level selection theory by Wilson & Sober, 1994).There has been ongoing controversy surrounding group selection; however, the idea that group-level processes support individual fitness has been supported by subsequent work - notably, the suggestion that emergent, group-level processes are important traits which affect individuals' chances of survival (Smaldino, 2014), and that groups form the selective environment - as in the niche construction theories we considered above (see Appendix 4).

An environment in which mutual collaboration is expected would favour those individuals with more prosocial motives who act in line with these norms. Systems of rewards and sanctions, and a general bias for conformity, might reinforce these processes - selecting those who are motivated to co-operate, and who possess the capacities for social understanding (see Boyd & Richerson, 2009, 2010, on cultural group selection; and Molleman et al., 2013, for a model of how the group selection of co-operation might interact with processes of social learning; a similar process is explained by Gintis, 2011). Thus, prosociality, the motivation and the ability to interact with others, is prevalent, and indeed necessary, for social living - whether it is the genetic predisposition, an individual personality trait, a social behaviour or activity, or a whole group, which is selected.

Benefits of group living

To consider where this social world might originate, we turn to the ecological pressures which might have triggered the construction of this co-operative, protective niche. The nature of the food supply, for example, has been linked to the development of egalitarian (or other) social systems (Woodburn, 1982), and to costly behaviour to defend the group's territory (see Brown, 1964, on his theory of economic defendability). Marean (2016) explores the possibility that abundant and predictable food resources located in marine coastline regions off northern and southern Africa (probably around 110 000 years ago) prompted intergroup territoriality and conflict, as well as 'hyperprosociality' within the group to co-ordinate this sort of behaviour.

Moreover, Collard et al. (2016) argue that environmental factors such as these drove increasing cultural complexity. They argue against the 'demographic turn' hypothesis (or the idea that an abundance of cultural forms came merely as a result of population size or density), instead suggesting that environmental risk, such as competition or other potential threats to available resources, led societies to develop more and more complicated technologies and social systems to manage, and potentially tackle, these problems. In these ways, the co-ordination of group efforts

was necessary in generating a material, and a social, niche which moderated any threats, and increased the chance of those individuals surviving in their (potentially hostile) environment.

The start of a species?

It is probably the case that there was no sudden shift, no big cultural explosion that coincided with the onset of *Homo sapiens* in the world. Rather the origin of our species should be understood more in terms of a series of small transitions, in an environment where many hominids were living and developing relatively successfully (Foley et al., 2016). However, the development of prosociality and norms of co-operation, in response to these selection pressures, then prompted further shifts in the social cognitive capacities of *Homo sapiens*, in particular.

According to the Vygotskian Intelligence Hypothesis, our abilities to engage with others are based, at root, on these co-operative intentions. Whereas non-human primates can point, this is generally only for instrumental purposes, i.e. "I want that", in response to potential competition from others. On the other hand, human infants, already within their first couple of years, show willingness to co-operate - pointing in order to *share* information with others (see e.g. Franco, 2005, and discussion below). This is the motivation they need to share joint goals with others, to achieve and maintain together a state of shared intentionality. For this, however, we need much more complex cognitive representations: we need to understand that another has a different perspective to us, try to imagine what that might be, and then act with this in mind. The impulse to want to help or to affiliate with others is therefore the basic context which supports the development of our more extensive social cognition: the understanding and awareness of other people which then enables us to learn from and to collaborate with them (see Moll & Tomasello, 2007).

Modern humans are distinguished not only by this shared intentionality, but also the "scaling up" of this to a whole-group level, forming *collective* intentionality. While the ability to co-ordinate with specific others, in shared intentionality, provided interdependence in defence, foraging and hunting, the emergence of group norms which guided these activities required a sense of collective identity and intentionality - a sense that the whole group can act as a unit (Tomasello et al., 2012; see a similar two-stage model in Tomasello & Vaish, 2013).

These capacities - whether being able to represent another's perspective or the idea of group norms - rely on a symbolic type of understanding (rather than something that is iconic or indexical). In other words, we need to be able to have mental images which do not link to anything concrete in the external world, but rather comprise an abstract idea or concept. This symbolic cognition probably emerged from our actions in the world: making and using tools to manipulate the environment extends our natural body schema and relies on (and develops) skills such as recursion - paving the

way for the reasoning central to use of language and syntax (Stout & Chaminade, 2012). At the same time, the pair bond - the co-operative niche in which infants learn how to engage with the world and with others in this way - favoured the selection of individuals with a more developed prefrontal cortex (Deacon, 1997). Thus, our social skills developed in a self-reinforcing cycle - with new actions, neural developments, then underpinning corresponding cognitive capacities, in a process of 'triadic niche construction' (Iriki & Taoka, 2012).

Alongside a more complex symbolic cognition, we developed other skills within this niche which enabled us to *learn*. High-fidelity imitation, especially of the whole action process (rather than just achieving a particular outcome), enables quick and effective learning, from generation to generation, as well as amongst the current social community. Instead of discovering everything each time anew, we could develop the group's resources in a so-called 'ratchet effect'. This, alongside our bias for conformity (and the presence of potential sanctions if we act otherwise), resulted in a 'cumulative culture' (Tennie et al., 2009; also Sterelny, 2011; and see model in Lewis & Laland, 2012). The willingness to co-operate with one or many others, therefore, provided the seed from which we developed a social landscape, with many material, technological, and symbolic innovations. This protective niche became more and more complex and diverse, shielding us from any external environmental pressures, and providing considerable selective advantage over our competitors.

The core capacities (that arose from and were reinforced in this environment) are summarised clearly in Marean's 'three pillars of human uniqueness'.

> No other living animal has this combination of hyperprosociality, heightened social learning, and advanced cognition. They are the unique descriptors in our species abstract and, in my opinion, are what archaeologists were chasing with their fuzzy concept of modern human behavior. (Marean, 2015: 541)

Our tendency towards prosociality, co-operation, and the corresponding abilities in social learning and symbolic cognition, all seem to play a large part in the emergence and development of our modern human behaviour. Next, we explore the basis for social interaction, and group identity, more specifically - mindful of the broader significance of these capacities within the evolutionary landscape. If this section addressed the question "why", the next shall explore the "how" of our social capacities: I shall discuss the nature of the biological, and cognitive, mechanisms which support our ability to interact with those around us, forming social groups through activities like music. I explore first what might drive our predisposition for sociality, including the role of low-level reward circuits; and second the perceptual skills we need to understand others' minds. Most research we draw on here focusses on the skills required as we *observe* the actions of others. In Chapter 2, however, we expand our scope beyond those 'spectator' theories to address how these individual

capacities are played out in an *interactive* context (see Schilbach et al., 2013, for a discussion of this distinction, and Hari et al., 2015, for suggestions about how experimental work should account for this).

Our social instinct

In the light of this evolutionary picture, it is clear that our enjoyment of social interaction, even a need to be in the company of others, is deep-seated. We also know this through everyday experience. From our earliest years, we seek out and are deeply affected by the positive interactions we have with people around us.[13] We maintain relationships with kin, as well as developing wider friendship groups outside of our family circle. We might talk to strangers on the street, but we are also likely to dance with people we don't know well in a night-club, or sing with them in church. This is such an integral part of our lives that we feel negative emotions, a sense of loneliness, when we lack social interaction or miss the company of others. This is acknowledged in the psychological literature in the concept of our 'need to belong':

> … human beings have a pervasive drive to form and maintain at least a minimum quantity of
> lasting, positive, and significant interpersonal relationships. (Baumeister & Leary, 1995: 497)

Others have referred to this as a 'concern with connectedness' (Pickett et al., 2004: 1095).[14] It is not enough to interact now and again; we must know that our relationships endure, with regular, positive contact with certain others, reinforcing the knowledge that we care for one another's wellbeing. Baumeister and Leary argue that this motivation could account for a vast range of behaviour, perhaps even underlying our need for power, validation, as well as intimacy:

> … the belongingness hypothesis would suggest that human culture is at least partly adapted to
> enable people to satisfy the psychological need to live together (Baumeister & Leary, 1995: 498)

We can now measure this need in a single-item self-report scale (Nichols & Webster, 2013); prior to this, researchers showed that a higher 'need to belong' correlated with accuracy in understanding facial emotion (Pickett et al., 2004); an experience of rejection led to better memory for social events (Gardner et al., 2000).

Above, we explored the emergence of this 'hyper-sociality', and its role in the development of modern humans (beyond the capacities of other hominids, or what we see in present-day non-human primates). Here, I shall outline the mechanisms and capacities which underpin our motivation as a peculiarly sociable species; in fact, social interaction might be our 'default mode' of functioning -

---

[13] An inability, or even just a disinclination, to do this, we view even as pathological.
[14] This construct of connectedness is used in the present empirical project - see Chapter 5, and results in Chapter 6.

it does not emerge from these mechanisms, but forms the context for their development (Hari et al., 2015). In turn, I discuss research on relationships (a) where individuals become attached to another in a pair bonding process, and (b) where they are bonded in a close-knit social group. In reviewing a wide range of literature in this chapter, I show how interpersonal interaction has great emotional salience, and deep evolutionary roots.

Music-making is one mode in which we relate to many different others, and is an activity which can be central to maintaining stable relationships within our community. Understanding the broader mechanisms behind why and how we approach, bond, and interact with others more generally will help us to understand, in turn, why we make music, the capacities which are involved, and - ultimately - the effects this might have on us.

Pair bonding

Our most salient bonding experiences are generally with one other person. We become attached first to our caregiver, who looks after us, providing our "way in" to the community at large (Bowlby 1984). They make sure our physical needs are met, providing us with food and protection. Just as important, interacting with our caregiver helps us to practise and develop the social skills we need to function effectively in the social world. Even as we go on to engage with lots of other people, in adulthood we again seek dyadic attachments, whether in the form of romantic relationships or with close friends. Monogamous pair bonds also occur in non-human animals, such as some species of voles. Combining evidence from our human experience with animal models of the underlying neuroanatomy helps to build a convergent account of how this might work.

Neurochemistry of attachment: oxytocin and the reward network

We have subcortical brain structures which respond to beneficial environmental stimuli - like food, drink, or sexual contact - with a reward response, reinforcing certain behaviours or encouraging us to seek out that stimulus in the future.[15] Negative responses to certain environmental events are also adaptive: we feel fear when there is a potential threat, something we would rather avoid. Even in our fairly comfortable Western society, where we do not need to worry about seeking food or escaping predators, this primitive approach-avoidance system is still at work. Emotional reinforcements of these kinds are implicated in all sorts of learning (the process of modifying or creating new synaptic pathways - see Schultz, 2000), even guiding our supposedly "rational" decision-making (Damasio, 2006). Forming attachments with others relies on this network; it motivates us to approach others, and then to continue to seek out their company after this, if it was a positive experience for us (Insel & Young, 2001; Lim & Young, 2006).

---

[15] Note that the 'liking' of a certain stimulus is distinct from the 'wanting' - these are supported by dissociable neural reward pathways (see e.g. Berridge, 1996).

The neuropeptide hormone oxytocin is implicated in this process, alongside the dopamine pathways stimulating the experience of reward - a process that Ruth Feldman terms 'crosstalk' (Feldman, 2017). Oxytocin modulates these reward responses, depending on the social context; it shapes our response depending on what (or who) we see, underpinning the social recognition memory necessary for attachment to *specific* others (Groppe et al., 2013). While it enhances our reward response when we see our loved ones (Scheele et al., 2013), it also inhibits this response, or even stimulates a fear response, if we see others with whom we do not want to engage (see Maroun & Wagner, 2016, on how oxytocin acts in different parts of the amygdala to balance these approach or avoidance tendencies).

A neurochemistry triggered by interaction

We find interaction, even with strangers, a rewarding experience. We have a reward response following all sorts of social stimuli, whether just making eye contact (Kampe et al., 2001), initiating shared attention (Schilbach et al., 2010), or mutual co-operation in an economic game (e.g. Rilling et al., 2002; McCabe et al., 2001). The rewarding nature of our interactions is manifested in our tendency for play - something also exhibited by other animals. Trezza et al., (2010) reviews the neurobiological correlates of play behaviour in rats, implicating dopamine and opioid networks (see below on the role of endorphins).

Interaction with our loved ones, often infants (and romantic partners), has some distinctive features - exaggerated, musical vocalisations, expressions of positive affect, direct gaze and touch. The individuals are closely co-ordinated, interacting with and responding to each other, displaying a high degree of 'biobehavioural synchrony' (Feldman, 2017).[16] This is rewarding behaviour, and again is shaped by the release of oxytocin (Feldman, 2012).[17] Even very small-scale, perhaps hardly noticeable, behaviour is relevant here: synchronised pupil dilation leads to mutual trust of interaction partners - another effect which is modulated by oxytocin (Kret & De Dreu, 2017). Oxytocin improves the fluency of our interactions, whether tapping synchronisation (in a leader-

---

[16] There are various definitions of synchrony which are used in the literature; Feldman, in her general approach, defines synchrony in terms of temporal concordance in behavioural patterns and physiology; this has potentially broad applications: 'Synchrony...describes a critical component of close relationships that builds on familiarity with the partner's style, manner, non-verbal patterns, personal rhythms, behavioural preferences, and pace of intimacy.' (Feldman, 2012: 382)

[17] In Feldman's model of attachment, these bottom-up effects, stemming from the interaction itself, are also modulated by more complex, higher-order representations, such as ideas of commitment. If our mode of social engagement creates the immediate conditions in which we might bond with others, then these conditions can be shaped from the top-down, with other more symbolic or normative, perhaps even moral, concepts or imperatives. These might encourage us to stay within an attachment relationship that is actually maladaptive, or help to re-wire our avoidance responses to heal from previous negative experiences (Feldman, 2017). The types of group formation which I explore later in this thesis - the way our interaction mode facilitates immediate bonding, and then how our ideas of group identity moderate this bonding process - parallel this distinction.

follower paradigm - Gebauer et al., 2016) or even inter-brain synchronisation (measured using EEG - Mu et al., 2016).

The same neurochemical response characterises maternal 'licking-and-grooming' caregiving in rats. This behaviour has long-lasting, even cross-generational significance: the amount of caregiving received influences later stress responses, and even the amount of caregiving they give to their own offspring (see Champagne, 2008, for a review). Thus, the experience of interaction within pair bonds is mediated by these neurochemical mechanisms, enacting and establishing attachment within the dyad.

Attachment and our view of the world around us

Our attachment relationships, although performed within an exclusive, one-to-one context, are relevant to our experience in the wider world. Human attachment behaviour corresponds to social interactions and stress response later in life. Sroufe (2005) noted longitudinal effects of infant attachment style (trait-like characterisations, developed from work by Ainsworth) on later self-reliance, emotional regulation and social competence; more specifically, Feldman (2007) reported that the degree of synchrony in mother-infant face-to-face interaction predicted capacity for empathy even in adolescence.[18]

Attachment in adulthood also shapes how we respond to events or stressors in the environment. James Coan's 'social baseline theory' (2008) posits that proximity to our significant others helps to regulate stress through load sharing or risk distribution. Relationships of trust and interdependence reduce the risk from outside threats - we can afford to worry less, as we know others around us can respond and protect us, as we could do for them. We even exhibit a smaller neural response to threat when we are with someone we feel close to (Coan et al., 2006).

Belonging to a social group is also a context in which we develop relationships of trust and expected interdependence with others around us. There is a strong, biological basis for pair bonding, enacted through the experience of interacting and co-ordinating with another. The feeling of group belonging might come about in a very similar way.

---

[18] See also Carter (2005) on the role of oxytocin in mediating the effects of neglect.

Group belongingness

Although attachment relationships are most often considered only in the context of pair bonds, we can have a similar experience towards a social group (see Smith et al., 1999 for an application of attachment theory to group membership).[19]

The purpose of this thesis is to address the question of how our group bonds, our sense of belonging to, or identifying with, a group, comes about through the way we interact with others. Now that we have outlined the basic mechanisms behind attachments in general, we can consider how the same biological processes might support either our bonding to other group members, or an identification with the group as a whole.

Oxytocin in group attachment

In her work on attachment, Feldman (2017) includes the role of 'biobehavioural synchrony' in dyadic up to group-level interactions. Importantly, the same biological correlates are involved. For example, the regulatory effects of oxytocin, shaping our attention towards, and salience of, certain social cues are just as important in the context of a group. We are motivated to act prosocially, experience empathy, and have a decreased stress response when we are amongst members of our own group - again, this offers protection, moderating potential risks from environmental threats. This is a selective response, just as with specific attachment figures (Shamay-Tsoory & Abu-Akel, 2016).

In this case, the social cues need not be people with whom we have interacted. Rather, we can feel attachment towards the group identity, moderating our response to others on the basis of cues for in-group or out-group membership. Oxytocin is responsive to these types of social cues, motivating selective prosociality, conformance to specific group norms, as well as aggression towards the out-group (De Dreu & Kret, 2016).

Furthermore, there is a large literature comparing the effect of intranasal exposure to oxytocin, versus a placebo condition, on co-operative behaviour in economic games. MacDonald and MacDonald (2010) review a number of these studies, highlighting the overall effect of oxytocin on biasing individuals towards prosocial behaviour (see Kosfeld et al., 2005, and Zak et al., 2005, and Barraza et al., 2011, for specific investigations of this; although, compare meta-analysis by Nave et al., 2015, who suggest that these findings might be somewhat exaggerated). Baumgartner et al. (2008) showed that oxytocin reduced the response to a breach of trust - participants in this condition continued to act prosocially, even when their partner had not (corresponding to decreased activation of the amygdala).

---

[19] The 'Inclusion of the Other in the Self' scale is applied to both individual and group-level contexts (see Aron & McLaughlin-Volpe, 2001).

Grooming at a distance

Robin Dunbar also considers how our social network is maintained through interaction - with important psychopharmacological and cognitive correlates (Dunbar, 2018). He argues that primate attachment, so-called 'social grooming', behaviour can again be extended to incorporate much larger-scale human communities. This implies a similar continuity between one's close attachments and the wider membership of our social network.

Primate group sizes are correlated with neocortex volume (Dunbar, 1992) - perhaps a proxy for our social cognitive capacities. According to the Social Brain hypothesis (see overview in Dunbar, 2007), the number of relationships we can maintain at once depends upon our capacity to keep track of those links, an information-processing capacity whereby we can keep in mind not only third-party relationships, but also how others behave towards us. Social relationships are 'implicit social contracts', an understanding of reciprocal support as and when it is needed (Dunbar, 2018). In other words, our social relationships are built on trust. This depends upon our ability to understand others (capacities we shall explore further in the next section of this chapter), and, crucially, has important behavioural manifestations.

In particular, primate group size is correlated with time spent grooming (Dunbar, 1993).[20] This is a social behaviour which requires single-minded attention devoted to a partner. It is a significant time investment (up to 20% of their day), requiring mutual trust; 'a statement of commitment' (Dunbar, 1996: 44). Vervet monkeys are more likely to go to another's aid if they are a recent grooming partner (see work by Seyfarth & Cheney, reported in Dunbar, 1996); grooming can even act as a form of reconciliation following a conflict or breach of trust (see Dunbar, 1996: 26). The action of physical grooming creates 'the psychological environment' for this social reciprocity (Dunbar, 2010): it lowers the heart rate, reduces stress, and triggers the release of endorphins (e.g. Keverne et al., 1989; also Machin & Dunbar, 2011). In primates, this sort of close relationship supports small coalitions, 'buffering' individuals from external threats (see overview in Dunbar, 1996; also, compare this idea to Coan's Social Baseline Model of attachment).

The time demands of this activity impose a limit on group or coalition size (Dunbar et al., 2009). Thus, the larger networks we see amongst humans necessitate forms of interaction which have the same effects as this dyadic social grooming, but are more efficient - extending to more individuals at once. Dunbar hypothesised that language evolved to meet this need:

---

[20] Although, compare Grueter et al., 2013; with a reply from Dunbar & Lehmann, 2013.

> ... language evolved as a "cheap" form of social grooming, thereby enabling the ancestral humans
> to maintain the cohesion of the unusually large groups demanded by the particular conditions
> they faced at the time. (Dunbar, 1993: 689)

Through speech, we can engage with multiple individuals simultaneously (to be exact, three times as many - in proportion with the size of human as compared to other primate groups - see Dunbar et al., 1995), while doing other things; speech also enables us to learn about others and their relationships without direct observation, i.e. in gossip (see overviews in Dunbar, 1993, and 1996).

However, language alone was unlikely to 'bridge the bonding gap', as Dunbar puts it. Just exchanging words with another, or building up our knowledge of whether or not to trust certain others in our society, do not give the same endorphin 'kick' as does the action of physical grooming (Dunbar, 2004). Other types of interaction might enact more effective 'grooming-at-a-distance'. For example, laughter releases endorphins (as measured by proxy in pain threshold, by Dunbar et al., 2011), is particularly salient (even contagious) in a social context (Devereux & Ginsburg, 2001), and remains an important part of our social encounters - promoting affiliation and cohesion, 'smoothing' our interactions with others (see Gervais & Wilson, 2005). Moreover, behavioural synchrony increases liking and trust for others, mediated by ratings of self-other overlap and opioid release (with pain threshold as a proxy - see Lang et al., 2017); synchronised rowing also causes an endorphin surge (Cohen et al., 2010). Dunbar and his colleagues have found that singing, drumming and dancing together also result in endorphin release (again, measured using pain threshold); the crucial part of this is the activity, rather than more passive engagement with the so-called "music itself" (Dunbar et al., 2012). In particular, moving in synchrony with others seems to contribute in large part to this bonding effect, making participants more likely to help or reward their partners (whether after rowing or dancing together - see Cohen et al., 2010; Tarr et al., 2015, 2016; although compare Davis et al., 2015).[21]

Thus, the endorphin system might play a more important role than oxytocin, being implicated in dyadic grooming behaviour as well as in relationships beyond the exclusive infant-caregiver or romantic partner contexts. In summary, though, whether it is oxytocin or the endorphin system which plays a primary role, group belongingness seems to be underpinned by the same emotional mechanisms as our experience of attachment to a significant other - and interaction is important across the board. Our experience of bonding or attachment towards other individuals seems to operate with similar mechanisms, whether we are in a dyad or a group, or whether we know the

---

[21] The use of fire is also likely to have created extra time for these activities to be built into the community routine. While it was an adaptive technology for cooking, protection and making tools, it also lengthened the waking day, creating time which could not easily be spent hunting or foraging for food, but rather could be dedicated to social bonding activities (such as dancing, chatting, story-telling) with everyone brought together around the campfire (Dunbar & Gowlett, 2014).

other person intimately or not at all. These are significant experiences, implicating deep-seated (and evolutionarily-ancient) emotional resources. Now I have explained our basic motivation to interact with others, I shall next outline in more detail the skills and awareness we need to engage with them successfully.

Social cognition

We are motivated to interact with other people. We seek out others, especially those select few to whom we feel attached, or those who with whom we share the same group association or membership, and desire to feel close to others, or that we belong. However, the skills and understanding we need to engage with others are more involved than this. These capacities are generally encapsulated in our 'social cognition'; this is an umbrella term which includes a whole range of perceptual abilities. In their review, Happé and colleagues define it as:

> ... the processing of stimuli relevant to understanding agents and their interactions (Happé et al., 2017: 244)

They include more general skills as part of this processing ability - including memory and attention - as well as our general social motivation (see above). More specific abilities include imitation, emotion recognition, empathy, social learning, and Theory of Mind (Happé et al., 2017).[22] Frith and Frith (2007) have a similarly broad definition:

> Social cognition is the sum of those processes that allow individuals of the same species...to interact with one another. ... Essentially it depends upon the exchange of signals. (Frith & Frith, 2007: R724)

In this section, I shall outline some of what is involved in recognising and exchanging these signals; this ranges from quick, automatic responses, to more deliberate, top-down processing. These same capacities are involved regardless of how we engage with others - whether it is avoiding someone on the street, holding an argument, or playing a piano duet. I offer here a very general review, addressing, in turn, how we perceive and simulate others' experience, how we understand another's *different* perspective, and then how we interact or work with them in the light of this.

Understanding others: self-other equivalence

Work in social neuroscience provides insights into a low-level, neural basis for social cognition. This is, at root, the understanding that others are 'like me' (in the words of Meltzoff, 2007) - a starting point by which we can relate to others, on the basis of our own experience.

---

[22] Happé et al. (2017) also describe how cognition denotes a level of explanation which lies between neural processes and behaviour, but crucially depends on and relates to each of these.

The discovery of mirror neurons

In the 1990s, a research group investigating the neural correlates of actions performed by monkeys noticed - by chance - that the very same neurons were stimulated when the experimenter moved the relevant objects into place in setting up the experiment. A subsequent investigation of this phenomenon revealed that there are neurons in the F5 region of the premotor cortex which fire both when the monkey performs specific goal-directed hand movements and when it just observes those same actions (Di Pellegrino et al., 1992).

This was a highly significant finding, and one which has had relevance in research on the mechanisms underpinning action understanding, Theory of Mind, empathy, as well as our ability for joint action (thus interaction) with others. It has been described as:

> … a neurophysiological account of the experiential dimension of both action and emotion understanding. (Gallese et al., 2004: 396)

Importantly, in a review of recent findings, Keysers (2009) notes that mirroring systems seem to be at work in various parts of the human brain, but so far there is only animal evidence for the presence of actual mirror neurons. Therefore, Keysers recommends that we speak of 'a *putative* mirror neuron system.' With this reservation in mind, it seems likely that the capacity for neural simulation, drawing on the neural resources we have in common with our interaction partner, is likely important in our ability to understand and co-ordinate with them. I shall now review some of the evidence for this claim.

Simulating actions and emotions

If we have a mechanism by which we simulate others' actions when we observe them, it means that we can draw on our own experience of performing those actions to infer their goal, predicting what the other might be intending to do or to achieve. In a way, we experience the same actions - even though we do not actually carry them out.

> With this mechanism we do not just "see" or "hear" an action or an emotion. Side by side with the sensory description of the observed social stimuli, internal representations of the state associated with these actions or emotions are evoked in the observer, "as if" they were performing a similar action or experiencing a similar emotion. (Gallese et al., 2004: 400)

Even just hearing the outcome of an action activates the equivalent motor regions of the brain (e.g. Keysers et al., 2003). This mechanism is most effective for actions we are already familiar with (see a comparison of novel and familiar actions in an fMRI study by Brass et al., 2007).

Moreover, we do not just simulate another's movement, but also actually mimic them. Often, we do this without realising it; it occurs frequently when we are conversing with others.

> One may notice using the idiosyncratic verbal expressions or speech inflections of a friend. Or one may notice crossing one's arms while talking with someone else who has his or her arms crossed. Common to all such cases is that one typically does not notice doing these things - if at all - until after the fact. (Chartrand & Bargh, 1999: 893)

We do this to aid the fluency of our interactions, to understand or predict what the other is doing, and to make the other person like us more (but only when they do not notice us doing it). People who score higher in tests of "perspective-taking" mimic others more (Chartrand & Bargh, 1999). We also use mimicry to enhance affiliation with our partner, especially when we feel we need to compensate for an otherwise unsuccessful encounter (see e.g. Lakin & Chartrand, 2003).

Motor simulation mechanisms are also important in our recognition, and experience, of others' emotional states. The same mirroring, 'experiential insight' (Gallese et al., 2004: 401), lies behind emotional mimicry (for example, mimicking another's facial expressions), and even emotional contagion (where we actually *feel* the same emotions as others around us). For example, we recognise that another feels either disgust or fear, just based on whether they have widened or narrowed their eyes (Lee & Anderson, 2017); we even feel others' pain, with the same neural correlates - activation of our insula - as we would expect from first-hand experience (Lamm et al., 2011; although modulated by experience and context, e.g. Cheng et al., 2007). Again, this helps us to affiliate - we bond with interaction partners, even in a group context, through a shared emotional experience (see Hess & Fischer, 2013; Barsade, 2002).

A neural basis for language?

This mirroring, or perception-action coupling, forms an automatic link between us and others. I have given examples of how this is important in signalling - whether showing affiliation, in recognising another's intentions or how they are feeling.

Some have argued that this same mechanism can support more complex signalling, such as the skills of perception (and prediction) involved in the back and forth of a conversation. This implicates higher orders of intentionality; in other words, we do not just signal and recognise others' signals, but we also signal in the knowledge that others are recognising us, and we recognise others' signals knowing that they know we are doing this etc. - in a recursive system (see Dunbar, 2007, for an illustration). Rizzolatti and Arbib (1998) were among the first to suggest that our ability to simulate and recognise others' actions, alongside some mimicry of those actions - enabling the other to recognise that we

are simulating them in this way - is the start of a bridge between the two communicators which can then support dialogue.

Since then, Arbib (2005) has traced a possible evolutionary development, from action recognition, to imitation, and then forms of proto-language. Gallese (2008) formalised this in his neural exploitation hypothesis: our mirror neuron system comprises the root of our language abilities, which are thus fundamentally 'embodied'. Understanding others' movements requires us to simulate and represent the structure and hierarchy of their action, and infer its goal; this same mechanism, only in more abstracted form, lies behind our perception and performance of the hierarchies and structures of language:

> ... the "words" of the premotor vocabulary...are not only assembled and chained to form intentional "action sequences"...; they can also be assembled and chained to structure language sentences and thoughts. (Gallese, 2008: 328)

A similar line of thought is used to link our early tool use to the symbolic cognition recruited in language (see e.g. Arbib, 2011, and Vaesen, 2012). Others, however, are much more sceptical of the link between a mirror neuron system and more complex forms of social cognition (see e.g. Cook et al., 2014, and Jacob & Jeannerod, 2005).

In general, though, our basis for relating to others is that they are "like us". On top of this, we shape our perception of others according to more complex social factors or considerations - whether we are aware that they have a different perspective, for example. I turn next to address these moderating influences.

Understanding others: beliefs and emotions

When we interact with other people, it is important to understand that our partner is not always just like us - but is distinct, with a different perspective, and thoughts and beliefs which might decidedly differ from our own. If our basis for engaging with and understanding others is an awareness of self-other equivalence, then we need to complement this with self-other distinction or control, taking into account other, external factors which might influence them (even if they do not influence us). This balance of self-other equivalence and control is important in (a) our understanding of Theory of Mind, and (b) our ability to empathise.

From mirroring to mentalizing

Our ability to mentalize is a sort of mind-reading ability - being able to understand, or 'infer' another's mental state or emotions (Brüne & Brüne-Cohrs, 2006). A couple of fMRI studies have implicated 'mirror neuron areas' of the brain when participants observe *social* actions. For example, Becchio et al. (2012, in a study of the same name) found that observation of competitive or co-

operative actions (i.e. two people reaching for the same object, or working to build a tower together) correlated with greater activation of mirror neuron areas (e.g. the inferior frontal gyrus), and mentalizing brain regions (including the prefrontal cortex), than observation of individual actions, moving objects irrespective of the other person. Moreover, Iacoboni and colleagues (2005) found that some of the same mirror areas (again, the inferior frontal gyrus) were activated in observing an action that was socially 'embedded' i.e. for cleaning, or drinking, in comparison to the same action observed with no context. In all of these social contexts, when we observe someone else, we are trying to discern what their intentions might be, so we can predict what they might do and respond appropriately. Our mirror neuron system - i.e. the understanding of self-other equivalence - is implicated in this process: 'the mirror and the mentalizing systems work in concert as to code for others' social intentions' (Becchio et al., 2012: 247).

However, self-other mirroring is not enough. Others might hold beliefs that are different to ours, and perhaps even different to what we know to be true about the world: it is important to recognise another's perspective and how it might be shaped by external conditions or access. Theory of Mind tasks, including the famous 'false belief paradigm', are used to assess our understanding of others' different minds. In the 1980s, Wimmer and Perner developed a task in which children observed an object being placed somewhere, and had to predict where a second observer thought the object might be. The twist was that the second observer left for a time, and, in that period, the object was moved - without their knowing. Children at around age 3 did not acknowledge this fact in their answers, not seeming to recognise that the other person had a false belief about where the object was; however, by age 4, half of the participants began to recognise that their beliefs were different; most of the 6-9 year-old participants could answer correctly. This provides evidence that this ability to mentalize comes with some social experience,[23] emerging consistently around the age of 4 (Wimmer & Perner, 1983; see meta-analysis of studies by Wellman et al., 2001).[24]

Moreover, in interacting with another, it is not always helpful to mimic them, but rather to take complementary roles. Hasson and Frith (2016) argue that we should understand interaction in terms of 'coupled dynamics' more than simple mirroring: consider the joint action of passing and catching a ball, or playing piano duets where one of us is to provide the off-beat. Whether predicting another's false belief, or acting in a different but complementary way, we in fact need to inhibit our tendency to mimic or mirror others. Brass et al. (2005) found just this - in a task where we need to *not* imitate

---

[23] This also implicates the extended altricial, or juvenile period, in human lives, within the context of the infant-caregiver bond - a context in which the infant learns these important social skills (see Brüne & Brüne-Cohrs, 2006, for details).

[24] Some argue that task demands mean that this ability is only expressed explicitly at this age, but the understanding might be present earlier, by just 13 months. Clements and Perner (1994), and Onishi and Baillargeon (2005) introduce implicit measures which might demonstrate an earlier ability to mentalize, but this is difficult to interpret (see Heyes, 2014).

others, in other words, where we inhibit our automatic mirroring system, we activate the same neural areas recruited in Theory of Mind tasks. In this way, we seem to have two layers, perhaps systems, in our social cognition - (1) an automatic, self-other equivalence, stemming from neural mirroring mechanisms, and (2) a more cognitively-demanding and flexible ability to perceive social states according to context (see Apperly & Butterfill, 2009, and Evans, 2003).

Empathy

Empathy is a multi-faceted construct which recruits both of these layers in its different components.[25] Preston and de Waal define empathy as:

> ... any process where the attended perception of the object's state generates a state in the subject that is more applicable to the object's state or situation than to the subject's own prior state or situation. (Preston & de Waal, 2002: 4)

Or, more concisely:

> ... the ability to share the feelings of others (Bernhardt & Singer, 2012: 1)

This of course includes a degree of self-other equivalence: just observing another's emotional state or affliction primes our own representations as if we were in their position - recruiting the same perception-action links discussed above (see, for example, the parallel brain networks activated in observation of disgust, in von dem Hagen et al., 2009). As we saw above, these lie behind emotional contagion (e.g. Barsade, 2002 - see above), real feelings for the other person which motivate prosocial behaviour to alleviate that state (see contrasting accounts of this by Batson, 1987 vs. Cialdini et al., 1987; also, de Waal, 2008).

Alongside this 'emotional empathy' component, there is also 'cognitive empathy'. It is not always helpful just to feel and act according to our own sympathetic distress; rather, we should sometimes inhibit these feelings to then be able to act appropriately (Preston & de Waal, 2002). This is not always a positive thing, though - inhibiting our emotional empathy might be due to motivations for revenge, envy, or even because the other person belongs to an out-group (Gutsell & Inzlicht, 2010; see also reviews by Bernhardt & Singer, 2012, and Singer & Lamm, 2009).[26]

In summary, our complex forms of social cognition - our response to social signals or stimuli - comes from low-level self-other merging mechanisms, relating to others as if we are like them, alongside

---

[25] This skill is also associated with early parental interactions - see Gonzalez-Liencres et al. (2013) for one possible account of this.

[26] A high score on traits in the so-called 'dark triad' of personality, e.g. psychopathy, is linked to selective deficits in emotional (or affective) empathy, but without any impairment in cognitive empathy (Wai & Tiliopoulos, 2012).

cognitive flexibility, taking account of context, even being inclined to dissociate from others in some circumstances. Before the close of this chapter, I shall apply these skills to an interactive setting: beyond just understanding others, or predicting what they are thinking, we need to consider how we work to align our intentions and beliefs with them *in situ*.

Working with others: sharing their intentions

Michael Tomasello and his colleagues have suggested that what makes human social cognition different from that of our nearest primate relatives is the capacity for shared intentionality:

> ... a suite of social-cognitive and social-motivational skills...[which] refers to collaborative interactions in which participants share psychological states with one another (Tomasello & Carpenter, 2007:121)

It is important not just to investigate our ability to perceive and interpret others' social signals, but also to address how we work together in interaction, actively responding to and influencing the other in real time. Gallotti and Frith call this the 'we-mode':

> ... interacting agents share their minds by representing their contributions to the joint action as contributions to something that they are going to pursue together, as a "we". (Gallotti & Frith, 2013: 163)

This builds upon our skills of perception, the bridge we form with others when we realise we are like them, and the cognitive flexibility we have in recognising that they might differ from us in meaningful ways. It also acknowledges how we each *act* to manipulate the other and their mental states, in a mutual exchange.

Shared attention

The simplest form of this 'we-mode' is our tendency to share focus, to look at something together. We intuitively follow another's gaze, using this signal to understand what they are attending (apes can also do this, but rely more on head movement: Tomasello et al., 2007). Beyond this, though, we work together, directing each other's gaze in order to achieve this state of joint (or shared) attention:

> ... joint attention is not just two people experiencing the same thing at the same time, but rather it is two people experiencing the same thing at the same time *and knowing together that they are doing this* (Tomasello & Carpenter, 2007: 121, original emphasis)

The aim is not just to look at something together, but to *know* that we have this mutual focus:

> ... both individuals are aware of each other's object of attention and each of them will use the other's gaze direction to check that both attend to the same target (Itier & Batty, 2009: 848)

Siposova and Carpenter (2019) have recently offered more clarity about the nature of this shared state, and the type of common knowledge (or common ground) it involves. They prefer the term 'social attention', and unpack it into a number of levels, all of which 'depict some form of triadic relation between self, other, and an object of their attention' but vary in the degree of 'jointness'. Monitoring attention is an individual experience, tracking and manipulating the attention of another; common attention involves the mutual knowledge of a shared focus, but again still on the level of the individual. In contrast, *mutual* attention involves more explicit shared focus: 'both of them directly experience each other attending both to that thing and to their own attention to the thing and each other'; and, finally, *shared* attention requires the intentional communication about that state (Siposova & Carpenter, 2019: 261-262).

These begin to emerge from infancy: we use pointing gestures to try and orient another's attention in a particular direction, and then check back to their partner to make sure this has been successful (see Franco, 2005). The associations we learn through doing this are important in language acquisition (see e.g. Brooks & Meltzoff, 2005). In this way, we co-operate with one another (see above on Vygotsky Intelligence Hypothesis) to build our common ground about the world around us. We take into account what we know about the other (based on our own experience, but also from an awareness of their own, perhaps differing, perspective), working to ensure that we are on the same page. We shall see later (in Chapter 2) that achieving and maintaining this common ground is still crucial in ever more complex forms of social interaction.

Social learning

We use all of these skills - from an understanding of action intentions, Theory of Mind, and the inclination to want to share attentional focus - to learn from one another, building up our own knowledge and understanding of the world (and how we can act upon it - with tools, for example). Our self-other equivalence is crucial here. Sommerville and Hammond (2007) set up a task whereby an adult showed a child how to build toys. After working collaboratively (taking turns), the child displayed a memory bias, reporting more often "I did it" for the experimenter's actions, presumably having represented them cognitively as their own.

However, this ability to mimic is again recruited flexibly. Wang and Hamilton (2012) present a framework to understand the top-down modulation of imitation (again implicating the prefrontal cortex - see mentalising section above), which depends on the social context. We are much more likely to imitate (and thus learn from) higher-ranking individuals (who are presumably more expert or reliable; see theory of prestige presented by Henrich & Gil-White, 2001, or evidence from Rakoczy et al., 2010). Again, we use imitation deliberately to build affiliation - we even over-imitate, demonstrating our adherence to certain norms, even if they do not add any instrumental value to

the action (Over & Carpenter, 2013). Most children do not just imitate, but can also switch roles, demonstrating that they have understood what is required (e.g. Carpenter et al., 2005); they readily adapt to teaching their peers the skills they have just acquired (as in e.g. Whiten & Flynn, 2010) and they even recognise when others are adhering to their learnt norms (and object if they are not - Rakoczy et al., 2008).

Whiten (2017) offers a comprehensive comparison of children's abilities with those of chimpanzees. There are some similarities, and continuities, but the high fidelity with which children copy the form of other's actions (rather than just emulating the result), the ability to build upon and develop those learnt techniques, and the over-imitation of certain norms all lie behind a striking disparity in the diversity and complexity of culture displayed by humans, compared to our primate relatives - and perhaps how a cumulative culture might have developed (see discussion above, and also e.g. Kendal, 2011, and Lewis & Laland, 2012).

Summary: our interaction engine

I shall summarise by drawing on the image of an 'engine' - a holistic and multi-dimensional concept, coined by Levinson (2006). The multiplicity of social, perceptual capacities, which all together comprise our social cognition, lies behind our ability to understand and engage with others. These capacities integrate more primitive reward-based representations with our more complex symbolic abilities, implicating neural pathways between older, subcortical regions and the more recently-developed prefrontal cortex (see Damasio's somatic-marker hypothesis - Damasio, 2006; and work outlining a 'social brain' by Adolphs, 2009, and Insel & Fernald, 2004).[27]

At root, our social cognition relies upon a willingness to co-operate, to approach others, a desire to engage with them in the first place - underpinned by the neurochemistry of attachment, our social instinct that we reviewed above. However, interacting with others is then a more complex task: we rely on subtle cues to infer a great deal about what another is thinking or feeling, try to influence their attentional focus, and in turn generate signals for their benefit. Levinson and Holler (2014) discuss the many layers of this 'stratified' system. Just holding a conversation with one other implicates deep gestural, embodied capacities, which are then co-opted into our symbolic systems of thought and language. We do not just talk - rather, we observe and understand others through their actions and align ourselves with them, generating a sense of 'shared intentionality' in order then to establish and build upon our common ground.

---

[27] The importance of these neural interactions is revealed in certain social deficits (and other traits) characteristic of autism (Gaigg, 2012).

With the repertoire of capacities which make this intricate performance possible, our species has a greatly expanded potential for learning, dissemination of techniques or technologies within and between generations. As a result, we have much more complex and diverse social systems, cultural differences in the way we are expected to interact - a multitude of sociocultural niches (see above). Importantly, social interaction - and the particular way in which it is manifested in our community - is an art that we have to practise. We learn these skills, developing our empathy and Theory of Mind, acquiring our local language of interaction, through experience with others around us. Humans have a long period of 'altriciality': in other words, we are born relatively underdeveloped, so our neural and cognitive development occurs as we engage with others, situated within a particular locale, from our earliest years (see Joffe, 1997).

It is the performance of these interactive skills which we are interested in here. This chapter was devoted to understanding our individual, social capacities, and their place in the evolution of our species. However, our conversations, and our music-making, are all enacted in a rich social context - they recruit these individual capacities, but in ways which are tailored to the specific cultural world in which they are performed. In order to understand how social - specifically, musical - interaction pertains to our relationships with others, we need to address how our social cognition is played out, in real time. In my next chapter, I turn to address this - shifting focus from research on our biological, cognitive capacities, to draw on observations of our everyday social encounters.

Chapter 2 Social interaction

If Dora had been alone she would have called out at once to Toby, so little was she embarrassed and so much amused and pleased by what she saw. But the proximity of Michael, which she had for a moment forgotten, made her pause, and turning to him she had a sense of embarrassment, not so much because of his presence as on his behalf, since he would perhaps imagine some embarrassment in her. (The Bell, Iris Murdoch, p. 75)

Humans are social animals, with the motivation and capacity to engage with others. We grow up in this peculiar niche, developing these skills from our infancy. Holding a conversation is not a simple task. It involves multiple systems operating in parallel, with precisely-timed movements and actions, and a high level of interpersonal and symbolic understanding. We start to be able to do this, with apparent ease, before we can walk.

In the previous chapter, I gave an overview of some of the mechanisms which underpin this extraordinary ability - the neurobiological architecture, and the basic cognitive skills involved in our 'shared intentionality'. Importantly, these are *enactive* mechanisms - we develop and demonstrate these skills in a social context, through interacting with others (see Schilbach et al., 2013); this is our 'default mode' of engaging with the environment (Hari et al., 2015). Now, I shift focus to explore the behaviour involved in this activity - a rich (and somewhat messy) picture. Most research in this area investigates dyadic, spontaneous conversation between two adults, who often have no previous acquaintance. I draw out key observations and themes from this work, and present one lens through which we might understand the possible *goals* of this behaviour. This (albeit simplified) picture is helpful in building my theoretical (and, ultimately, empirical) paradigm; but, the categories we establish here are, in reality, much more flexible and constantly overlapping.

First, I present an overview of the intricate behaviour involved. Since the first micro-analyses, conducted during the 1960s, there has been a growing body of work deepening our understanding of what goes on, moment-by-moment, as a conversation unfolds. We act in synchrony with our partner, co-ordinating when and where we look, each taking our turn to speak while the other assures us that they are listening. Second, I identify what might be the main goals of our interaction. The aim to maintain and build common ground with our partner is an overriding concern; however, the encounter might be focussed more or less towards either transactional or more relational goals. My third section explores one relational focus. As we interact with another, we consider the impression we are giving, being sure to maintain the other's positive view of us. Where our relationship is affiliative, we also show a degree of tact to protect the other's reputation, or *face*.[28] I provide an

---

[28] Where we are motivated by certain power dynamics, however, things can be a bit different.

overview of this goal, as well as the function of embarrassment - a social emotion we experience when these things go wrong. All of these components of conversation serve the process of *grounding* - maintaining our shared focus, with a common goal to communicate and to affiliate with our partner. We hope that we both hold positive impressions of each other; when we both realise something has gone awry, one (or probably both) of us feel ashamed or embarrassed.

Although all of this work so far focusses on spoken conversation, it is relevant for our understanding of music-making. In the second half of this chapter, I present the case for thinking about music and language as part of the same communicative system, but - again - with possible differences in their primary focus or goal orientation. The same framework of goal type - contrasting extrinsically- with intrinsically-oriented interaction - is useful in distinguishing music and speech as tending towards opposite poles on a single continuum; while speech highlights the transactional, music foregrounds the relational goals of engagement. I summarise these ideas, and suggest how they might also characterise different types of musical interaction. In turn, these musical goals might structure our social relationships - perhaps contributing to our sense of group identity (see Chapter 3 for an overview of approaches to social groups, and then Chapter 4 for a synthesis of these ideas and the resulting experimental hypotheses).

Everyday interaction: conversation

When we think about speech, we often immediately think of our use of language - the common code we use to communicate certain information to others, a vocabulary we acquire early on, and then continue to expand throughout our adulthood. However, speech is more than just syntax and semantics. Not only do we act through words, but we also gesture with our hands, nod our heads and shift our posture - communicating a lot through our body language. With a combination of these channels, we might aim to communicate something about the world, or we might just be modulating our relationship with the other, starting a conversation, changing topic, or trying (as delicately as we can) to extract ourselves from a conversation that has gone on far too long. By way of introduction, I shall introduce three key aspects of conversation which frame how I shall think about it for the rest of the chapter: as a shared activity, a multi-modal activity, and as an activity with both transactional and relational goals.

Whenever we enter a place with at least one other person in it, we are aware of them, and shape our behaviour accordingly (think of the sometimes stressful experience of entering an elevator with someone we do not know well). We might aim to engage with them, or to avoid such an encounter. But, as soon as we begin an interaction (or they approach us), we have to attend to and co-ordinate with them, going through the motions to fulfil our obligations as a conversation partner. Here, I talk

primarily about those 'focussed encounters'; in the words of sociologist Irving Goffman, these involve:

> ... the reciprocal influence of individuals upon one another's actions when in one another's immediate physical presence (Goffman, 1959:27)

Conversation is a *shared* activity, something we do together in the same way as passing a ball in football, or working together to lift a table (see Sebanz et al., 2006). For this reason, Garrod and Pickering describe speech as 'a particularly well-integrated form of joint action', involving alignment of actions, and symbolic or cognitive understanding (Garrod & Pickering, 2009: 293).

This joint activity involves holistic patterns and perceptions of behaviour. When we speak on the telephone, we only have the auditory information to go on (and the world of email and instant messaging is even more restrictive) - but it is always easier to communicate clearly in person, face-to-face.[29] In this context, we have access to facial expressions, information about what someone is looking at, as well as the way in which they are moving their hands and shifting their posture. Often, our observations are key in understanding what they are talking about or how they are really feeling.[30] Thus, we should approach conversation as a *multi-modal* activity. Levinson and Holler describe the role here of a whole 'ensemble of linked systems': 'language is the tip of an iceberg riding on a deep infrastructure of communicational abilities' (Levinson & Holler, 2014: 1-2). This claim is supported by investigations of the phylogenetic development of our 'interaction engine', especially the 'gesture-first' hypotheses of the evolution of language (see Levinson & Holler, 2014, for an overview, and also a presentation of this thesis by Tomasello, 2008).

When we view speech as an interactive, collaborative, and multi-modal form of engagement, it is then clearer that the meaning we construct together can be both transactional - sharing information about something in the world - and relational - modulating how we relate to each other in the process. In just the same way as we can use all our both our linguistic and our visual gestures to portray meaning which might be iconic, indexical or symbolic,[31] the *way* in which we act through all of these channels comments upon or changes the nature of our social engagement. These are, in Goffman's words, 'a special set of acts and gestures comprising communication about communicating', which frame the whole encounter according to a 'working consensus' (Goffman,

---

[29] Researchers have only recently started investigating the differences we experience between co-present face-to-face interaction and conversation mediated by video over the internet (as on platforms such as Skype - see e.g. van der Kleij et al., 2009).

[30] See, for example, the 'integrated message model' of Bavelas and Chovil (2000): they discuss how a variety of auditory and visual signals can together comprise synchronised 'acts of meaning'; Clark (1996) alludes to a similar phenomenon in his idea of a 'composite signal' (see Clark, 1996: 156).

[31] According to Peircian Semiotics, icons derive their meaning through similarity, indexical signals through a physical link or co-occurrence (such as pointing), and symbols on an arbitrary convention (see Atkin, 2013).

1966: 96, 99).[32] Thus, we enact both transactional (or instrumental) and affiliative goals, to differing degrees (from e.g. greeting, then passing on a message, to ending the conversation, we highlight the relational, transactional, and then the relational dimension again). These stages, we might say, have different *goal orientations* - two types of communicative intention which might become more or less important at different stages in our conversation.

In what follows, I shall explore each of these facets in more depth. In turn, I address how this joint action is performed (e.g. in the timing of our words, or where we look); I discuss how this work supports a process of grounding, helping us to align our understanding and intentions; and then I unpack the parts of conversation primarily oriented towards either the more transactional or relational goals. Our motivation to 'save face' is one example of how we manage an important relational dimension of the encounter.

Performing conversation: evidence from micro-analysis

Conversation is an 'organised, skilled performance' (Argyle & Kendon, 1967). It is co-ordinated and precisely-timed - a bit like (or perhaps very like) a musical ensemble. I shall give an overview of the types of behaviour which are involved in conversation, especially the roles of synchrony and mutual gaze.[33]

Turn-taking - a universal conversation format

It is a fairly basic feature of conversation that participants take alternating turns and, in general, avoid overlapping talk (see Schlegloff, 2000). It is normal (and indeed expected) that we give the other chance to voice their thoughts, and that we actively (and obviously) listen to them while they do this. Interrupting (even coming in with our turn slightly too soon) and talking non-stop are generally inadvisable and sometimes frowned upon.

 A general model of turn-taking was formulated from audio recordings by Sacks et al. (1974). This characterised the system as (1) 'locally managed' (at each point in the dialogue, the participants individually track potential transition points), and (2) 'interactionally controlled' (each participant shows an awareness that the other is tracking possible transition points, and takes account of this in the way they talk), being (3) 'sensitive to recipient design' (participants construct their turn according

---

[32] Abiding or not to these norms might make the encounter a positive or a distinctly negative experience; this implicates an understanding of etiquette, and even a moral dimension, to our conduct in interaction (see Goffman, 1966).

[33] Modern observational techniques now allow us to track the joint action involved - from the eye movements of multiple conversation partners, up to the synchronisation of whole-body movements. Before this precision was available, however, some of the seminal studies which set the stage for conversation analysis were conducted with a single video camera, careful placing of mirrors, and markings on the film itself to denote synchronised or other movement of interest. This methical work was crucial in the development of this field, and is of relevance here.

to the orientation, knowledge - or common ground - of their partner). These features comprise a format which is 'a basic form of organization for conversation':

> ... "basic", in that it would be invariant to parties, such that whatever variations the parties brought to bear in the conversation would be accommodated without change in the system, and such that it could be selectively and locally affected by social aspects of context. (Sacks et al., 1974: 700)

In other words - it has 'the important twin features of being context-free and capable of extraordinary context-sensitivity', or both 'general abstractness and local particularization potential' (Sacks et al., 1974: 700). Moreover, an analysis of conversations in ten different languages (by Stivers et al., 2009) suggests that this format is a universal - all examples displayed avoidance of overlapping talk, and minimal silence between turns (relative to the length that seemed "normal") - with cultural variants only in overall 'tempo differences'. They conclude:

> Our empirical evidence suggests robust human universals in this domain, where local variations are quantitative only, pointing to a single shared infrastructure for language use with likely ethological foundations. (Stivers et al., 2009: 10587)

Levinson (2016) expands on this infrastructure. He points out that turn-taking involves complex multitasking: in order to take our turn "on time", we have to start formulating our response, and launching language production while still listening to and comprehending the speaker, predicting what they will say next and when they will finish. Thus, it does not make speech easier; instead, we adapt our way of speaking to aid this process - for example, using particular intonations to signal the type of utterance as soon as we start. Despite this complexity, turn-taking is present from proto-conversation in infancy, and in sign languages - where it is not strictly necessary in the same way as it might be in speech; it also resembles the 'duetting' of primates such as monkeys and marmosets. Turn-taking is therefore not just an easy way to organise our speech; instead, it is likely to be an 'older' or more basic structure which provided a frame for the evolution of language itself.

Co-ordinating turns: interactional synchrony

Within this structure of turn-taking, it is important to take our turn at precisely the right moment. If we are too quick, we might cut the other off, interrupting them and seeming rude. If, on the contrary, we wait for too long, it might seem that we are not interested, or even give the impression that we disagree with what they have just said. If the timing of our talk can communicate all sorts of meaning, we must be able to - by default - take our turn at the right moment, to avoid possible misunderstandings. From experience, we know that this feels automatic, or "natural".

Duncan (1972) identified some signals we use to do this. We show that we are passing on the floor to the other, or 'turn-yielding', with a change in pitch intonation, alongside some sort of stereotyped expression (e.g. "you know"), slowing our speech and finishing our gestures; if we want to keep our turn, stopping the other person from coming in, we generally keep our gestures going (even if our speech momentarily stops). The listener supports our turn through 'back-channel' responses such as "mm", maybe completing our sentences for us.

However, as we have mentioned, there is more to taking a turn than just responding to the speaker (by which time it might well be too late). As well as the predictive mechanisms suggested by Levinson (see above), our timing is guided by a regular "beat", aligning the speaker and listener through a mechanism of entrainment (see Clayton et al., 2005, on how this occurs in all sorts of communicative, musical, and other behaviour). Condon and Ogston were the first to notice such 'interactional synchrony' (Condon & Ogston, 1967: 230).[34] In an analysis of conversation around the dinner table, between three family members, they note that one person would be moving their fork in time with the syllables of another's speech. They remark:

> ... the three interactants looked like puppets being moved by the same set of strings. (Condon & Ogston, 1967: 229)

Later work provided insights into potential mechanisms lying behind this. Wilson and Wilson (2005) suggest that entrainment happens on a neural level. 'Endogenous oscillators' (i.e. neurons with regular, periodic activity) in the brains of both participants entrain to the speaker's syllable rate. This effectively forms a beat (here called an 'entrained cyclic pattern'), which guides the timing of turn-taking: listeners will come in "on time" with this beat; their 'readiness' is supposedly 'counterphased' with that of the speaker to avoid simultaneous starts.[35] If they miss their entrance (or give a back-channel response in its place), the baton is handed back to the speaker. On the other hand, the listener can indicate that they want to speak, starting to gesture and form words at regular intervals. (If they are cut off by the speaker again and again, this can become a frustrating experience.) In this way, turn-taking is a mutual, carefully timed performance, involving the co-ordination of both participants to ensure it runs smoothly.

Kendon makes the parallel with music-making, suggesting that listeners know when to take their turn:

> ... much as a musician may begin to move conspicuously with the music, as he readies himself to enter with his part at the right moment. (Kendon, 1970: 114)

---

[34] Synchrony and entrainment are often used interchangeably in the literature. I use the term "entrainment" to denote a coupling mechanism, from which "synchrony" is often the outcome.
[35] Sometimes, two people can be unhelpfully "in phase": both start speaking at the same time, stop and then start again at precisely the same moment.

He built upon the earlier work of Condon and Ogston, providing more examples in his observations of movement among conversation groups in a hotel lounge. Importantly, Kendon notes that the mirroring of gestures, and synchronisation of the listener's body movements with the flow of speech, mark out a dyad from the rest of the group:

> By mirroring the movement of the speaker, the person directly addressed thus at once
> differentiates himself from the others present, and at the same time he heightens the bond that
> is being established between him and the speaker. (Kendon, 1970: 113)

Schober and Clark (1989) note a similar distinction between the addressee and any overhearers. The direct engagement between the conversation partners enacts a process of grounding - alignment of action corresponds to cognitive or symbolic alignment, and thus mutual understanding. In this way, the "musical" co-ordination and patterning of movement in time is a crucial part of speech, supporting our mutual understanding or common ground within the dyad (see footnote 37, p. 65).[36]

Joint action - back-channelling

The speaker and listener, swapping roles, are always both involved in maintaining this "beat" structure. Even if only one individual is talking at once, they depend upon the feedback of the listener, through the latter's 'back-channel' responses. A conversation depends upon this reciprocal, collaborative process: the listener reassures the speaker that they understand and want them to continue.

This is the case even in a very one-sided conversation. Bavelas et al. (2000) instructed one participant in a dyad to tell a 'close-call story', while the other was just to listen (and, in some cases, to do a counting task at the same time). The listeners who gave their undivided attention (i.e. who did not have a simultaneous counting task) contributed frequently with responses that were closely related to the story's content, smiling and showing excitement at the relevant points, even gesturing to depict the scene being described to them. On the other hand, when listeners were distracted by the counting task, not only did the frequency of these responses drop, but the speakers told their stories significantly less well (according to external raters). Listeners, therefore, are actually involved in the telling of the story:

> Even in highly asymmetrical dialogues, speaker and listener roles are not fixed and separate.
> Rather, their relationship is reciprocal and collaborative, in that the narrator elicits responses

---

[36] Also important is the interpersonal 'bond' that Kendon mentions. Synchrony in conversation enhances partner affiliation or degree of 'rapport' (Bernieri, 1988). This is crucial in infant-caregiver relations, with correlates in attachment style (secure attachments were associated with well-timed interactions, whereas more insecure relationships were characterised by less smooth exchanges - see e.g. Isabella & Belsky, 1991) and the development of empathy (Feldman, 2007). This will be discussed in greater depth in relation to musical interaction more specifically, and how this effects prosociality, especially in the context of a group.

from the listener and the listener's responses affect the narrator. In spontaneous storytelling, the interlocutors interact together to produce the narrative. (Bavelas et al., 2000: 951)

The researchers also investigated the role of gaze behaviour in a similar, asymmetrical dyadic context (Bavelas et al., 2002). They find that the speaker 'seeks a response' by looking at the listener. The listener (who is likely to be looking already), meets the eyes of the speaker, and produces their supportive response at this point, then the speaker looks away and continues with their story. They argue that these 'gaze windows' are important check-points in the grounding process: if the listener does not respond confidently at these points, the speaker then repeats themselves or elaborates until the listener shows something more convincing (an example of repair - see below). This 'microcollaboration' guides how the speaker's narrative will proceed.

Gaze behaviour

This study shows just one instance of the role of gaze; the eye movements of both the speaker and the listener play a disproportionate role in the structure and experience of dialogue. Meeting the eyes of someone else (whether just in passing, or within a conversation) is something we do not like to do for too long, but is a particularly useful way of understanding their intentions or mental state. Humans have unusually-shaped eyes in comparison to all other primates. They are more oval-shaped, with an extremely prominent white sclera. Whereas most primates camouflage where they are looking (as an adaptation against predators), humans display this clearly (Kobayashi & Kohshima, 1997). The opportunities this affords for signalling, and thus effective social communication, was important in our adaptation to group living (where the risk of predation is reduced by virtue of creating a protective social niche; see Emery, 2000, and also Chapter 1).

We focus on the eye region when we try to recognise another person (or just below the eyes - Peterson & Eckstein, 2012), or discern what emotions they might be experiencing (e.g. Lee & Anderson, 2017). This is important when coupled with information about someone's gaze direction, helping us to learn about the environment - the affective significance of objects around us. Infants can detect the difference between direct and averted gaze, later being able to orient their attention to what another is looking at (and vice versa - see section on shared attention in Chapter 1). This aids language acquisition, and is important in the development of Theory of Mind (with noticeable differences among individuals with e.g. autism spectrum disorders; see review by Itier & Batty, 2009).

Direct gaze (also called mutual gaze, or eye contact) is more salient in our interpersonal interactions. From birth, infants are sensitive to, and prefer, faces which are looking directly at them (e.g. Farroni et al., 2002); we also find it easier to remember those faces and read their emotions, and even rate them as more attractive (see Itier & Batty, 2009). Meeting someone's gaze activates a different

cognitive and neural system compared to when we see them look elsewhere. While averted gaze activates the superior temporal sulcus (important in our perception of context and motion, i.e. the orienting of our attention in the environment), eye contact implicates the amygdala and the fusiform gyrus (important in emotion and facial memory processes; George & Conty, 2008).

Direct gaze can signal affiliation or attraction, or otherwise communicates dominance or threat, depending on the context (see e.g. Ellsworth & Carlsmith, 1968; Exline, 1963). We tend to regulate the amount of eye contact we have with another in order to maintain an appropriate level of intimacy in our conversation (adjusted in relation to other factors such as physical proximity - see Argyle & Dean, 1965 - and amount of smiling - as in Kendon, 1967). We might use it for social control, perhaps persuasion or seeking friendship, or fostering a spirit of co-operation (see review of all of these functions in Kleinke, 1986).

Kendon (1967) showed that mutual gaze was particularly important in regulating the flow of a conversation. He analysed video film of conversations between two unacquainted individuals, who had been instructed to "get to know one another", and found that it functioned:

> ... both as an act of perception by which one interactant can monitor the behaviour of the other, and as an expressive sign and regulatory signal by which he may influence the behaviour of the other. (Kendon, 1967: 24)

One can tell that the other is about to speak when they look away; after that, the speaker then monitors the listener, looks for a response to signal understanding (see discussion of gaze windows, above), and indicates when they are coming to the end of their turn by looking again to the listener. Mutual gaze thus indicates and effects a turn transition; it is also a means for grounding - ensuring continued mutual understanding throughout the exchange.

This grounding process is shown in more detail by Goodwin (1980), who reveals an interesting pattern between a speaker's words and the listener's gaze behaviour. When one starts speaking, their utterance is rather jumbled and fragmentary - until they see that they have the other's attention (i.e. they achieve mutual gaze), at which point they restart and give the complete sentence. The speaker thus only takes their turn when they know the other is exhibiting 'proper hearship', signalling to the speaker that their words will be recognised and understood.

This tracking signal is important not only to validate the speaker in an exchange, but also for the listener to demonstrate that they want to be involved. In a more complex eye-tracking study of three-person conversations, Holler and Kendrick (2015) find that the 'unaddressed' participant looks to the current speaker, shifting to the other just before their turn ends, perhaps to ensure that they

are kept up-to-date with the conversation, and also 'to display recipiency' to those who are presently more involved.

In this way, the reciprocal patterning of actions - even just eye movements - in time creates and maintains an affiliative context in which we share both our experience of being in that relationship, and also the information or ideas that we want each other to know. Now that I have detailed the format of this performance, we can examine in greater depth the goals of the encounter - primarily, the goal to maintain common ground.

Achieving alignment: maintaining our common ground

We talk about "being on the same wavelength" with someone, or whether someone else "gets" what we are talking about. When we talk, we share knowledge with another, perhaps aiming to impart some of our wisdom (as in instruction), or to persuade someone who thinks differently (as in an argument or debate of some kind). Whether or not we eventually agree, our goal is, at least, to understand how another is thinking, so that we could think like them if we wanted to. In other words:

> ... dialog is a form of joint action in which interlocutors have the goal of aligning their understanding of the situation (whether they agree about every detail or not) (Garrod & Pickering, 2009: 295)

Aligning our understanding is not just about ideas and concepts. It is likely that, in fact, our fast, automatic mirroring mechanisms are a core part of this process (see Chapter 1). Our comprehension system and that involved in the production of our own spoken responses are interlinked via what Garrod and Pickering (2004) call the 'perception-behaviour expressway'. Basic motor resonance mechanisms might underpin even higher-level, shared cognitive representations (lexical, syntactic or semantic alignment) - perhaps with the different levels of representation reinforcing each other.[37] All of this creates our 'common ground' (Clark, 1996: 12).

Starting a conversation: discerning our common ground

Holding any conversation depends on sharing some form of communicative code or language.[38] Other cultural, or group-specific, knowledge is also often assumed, and this guides how we talk and what words we use (e.g. a mathematician will talk about their work very differently if they think their

---

[37] See the 'mechanistic account of dialogue' in Pickering & Garrod (2004). Garrod & Pickering (2009) suggest that the lower-level, automatic processes of imitation and entrainment are key to this process, underpinning more abstract cognitive representations (others go further and argue that the higher-order processes are not necessary at all - see e.g. dynamical systems theories of Shockley et al., 2009)

[38] We can communicate well in the absence of this, however. For example, while on holiday in a country where we do not speak the native language, it is possible to use a variety of gestures to ask someone to take our photograph in front of a specific landmark.

partner does or does not have a degree in the subject). On a more localised level, we take into account what another can or cannot see from where they are standing (although not always reliably - see Keysar et al., 2003). Overall, for two people, their common ground is 'the sum of their mutual, common, or joint knowledge, beliefs and suppositions' (Clark, 1996: 93). Importantly, this concerns our beliefs about our common ground, or what we *think* we share, based on some sort of evidence (e.g. what we remember from our previous encounters with them).

This is a recursive process: not only do we know what they know, but we also know that they know what we know, etc. Clark argues that it is impossible to hold in mind the infinite iterations of this, proposing that our level of knowledge is limited to the following levels:

> ... *p* is common ground for members of community C if and only if: (1) every member of C has
> information that basis *b* holds; (2) *b* indicates to every member of C that every member of C has
> information that *b* holds; (3) *b* indicates to members of C that *p*. (Clark, 1996: 94)

Relevance

The awareness of what we do (and what we don't) share with our partner at the outset, is then crucial in how we then interact. We form utterances on the basis of what we think they know, or what perspective they have - according to the principle of 'recipient design' (see above on turn-taking). More generally, Sperber and Wilson (2012) suggest that we hold expectations of relevance; these are not rules, but form a part of 'the very act of communicating'; they pertain to the knowledge, awareness (or common ground) held by each partner as a basis by which new information might be processed:

> An input is *relevant* to an individual when it connects with available contextual assumptions to
> yield *positive cognitive effects*: for example, true contextual implications, or warranted
> strengthenings or revisions of existing assumptions. (Sperber & Wilson, 2012: 6)

Their theory comprises two main principles: the cognitive principle of relevance - that 'human cognition tends to be geared to the maximisation of relevance' - and the communicative principle - 'every act of overt communication conveys a presumption of its own optimal relevance.' (Sperber & Wilson, 2012: 6). This forms a heuristic that makes communication much more efficient: we can assume that the speaker formed an utterance on the basis of what is most relevant, or salient, to them - as well as their knowledge of what is relevant to us, the recipient - and pick the interpretation which is thus likely to be the most obvious to us both. In this way,

> ... relevance-oriented inferential processes are efficient enough to allow for a much greater slack
> between sentence meaning and speaker's meaning, with sentence meaning typically being quite

fragmentary and incomplete, and speaker's explicit meaning going well beyond the minimal proposition arrived at by disambiguation and reference assignment. (Sperber & Wilson, 2012: 5)

Moreover, we can arrive at this richer meaning without too much effort:

> According to the relevance-theoretic comprehension heuristic, [a hearer] should follow a path of least effort, and stop at the first overall interpretation that satisfies his expectations of relevance. (Sperber & Wilson, 2012: 7)

Grounding and repair

On the basis of what we think we share, including a common understanding of the types of actions which are relevant for our discourse, we proceed to exchange and share information of one sort or another. In the process of speaking, we build up our common ground in a mutual, interdependent process of 'grounding' (Clark, 1996: 221). This happens incrementally. We do not just send information to another, but we communicate it in parts, checking at every stage for signals from them that they have understood (see above on back-channelling). These processes do not always work, however. If they look at us blankly, request clarification, or say something nonsensical, we know that we then need to go through a process of 'repair' - and one which minimises the disruption to our exchange. In general, this is initiated in some way by the recipient - namely, 'other-initiated repair':

> ... a recipient of a linguistic message signals that there is a problem understanding or hearing what was said, and the sender then "fixes" it. (Dingemanse et al., 2015: 2)

Schober and Clark (1989) demonstrated the role of these processes empirically by comparing differences in the understanding held by 'addressees' (those who are directly involved and can respond in conversation with another participant) and 'overhearers' (others who are not involved in the conversation). Participants were given a simple task to arrange shapes into a particular configuration. Because the addressees could signal their understanding (or request clarification), and thus could collaborate with the speakers in building up a shared vocabulary to describe the obscure shapes at their disposal, it was much easier for them to construct the right picture, in comparison to the overhearers (who still had access to all of the same information exchanged between the conversation partners). In this way, holding a conversation involves a reciprocal process of grounding, building up shared understanding between the interactants, which in turn guides their ensuing talk.

This grounding does not just happen in language use. Clark presents a 'ladder of joint actions', from the movement or articulation itself, up to its symbolic import (Clark, 1996: 221), and argues that

alignment must occur on all of these levels for the grounding to be ultimately successful (and the message to be jointly understood).

> To see how grounding works, we must look beyond language use. There are general principles about how people discharge intentions in performing any action, both autonomous and joint actions. If so, they should also apply to signaling [sic] and recognizing, presenting and identifying, executing behaviours and attending to them. (Clark, 1996: 222)

Our 'interactive alignment' (Garrod & Pickering, 2004) occurs through a mutual recognition of all sorts of gestures and actions, with language remaining just the tip of the iceberg (Levinson & Holler, 2014). The multi-modal performance of precisely-timed movements and gestures are oriented towards this goal - to share understanding or experience. However, the nature of this goal varies widely with context - as we shall see next.

Communicative goals

Conversation is best described as a form of joint action, a participatory activity involving many levels of action - from the mirroring of movements and gestures, up to alignment of our cognitive representations, as shown in our use of symbolic and abstract language. Beyond the mode (or multi-modality) of this activity, however, we need to consider the goal of this joint action. Some experimental work uses 'joint action paradigms' to investigate the dynamics of collaborative interdependence, or interaction. The tasks might involve individual (but simultaneous) responses to particular stimuli which appear on a screen (e.g. Sebanz et al., 2003), playing a piano duet in time with one another (or with a recording - see e.g. Novembre et al., 2014), or jointly tracking a moving object on a screen with a cursor (Knoblich & Jordan 2003). In these, participants work together towards some pre-determined, and external, goal (or are mutually aware of each other's separate tasks). Speech is a form of joint action - albeit, a more complicated, and intricately detailed, one; indeed, the goal of our everyday interaction tasks (whether spoken or musical) is much harder to pin down.

Transactional and relational goals

Most of our discussion of common ground so far has concerned our mutual understanding *about* something in the world. We take into account our partner's knowledge, what they can see, and our previous history of interaction in how we talk; we check regularly to see if they affirm what we are saying, that they know what we are talking about. In this way, we work together, using gestures or words, to co-create and share meaning, referencing things outside the interaction according to iconic, indexical or symbolic relations (see footnote 31, p. 58). Grice (1969) argued that this meaning can be understood in terms just of the speaker's *intentions*, which are uncovered by the listener through inference:

"U meant something by uttering *x*" is true iff, for some audience A, U uttered *x* intending (1) A to produce a particular response r; (2) A to think (recognize) that U intends (1); (3) A to fulfill (1) on the basis of his fulfillment of (2). (Grice, 1969: 151)

Sperber and Wilson expanded this to include the listener's actions, in their discussion of how meanings and intentions are understood *collaboratively* in terms of overall relevance (see above). Both perspectives, though, highlight the instrumental goal, a transactional dimension of our interaction - a little like the goals devised in those joint action paradigms. However, our aim in conversation is not solely to communicate something. We greet our partner, ask them how they are, and, while we talk, we nod, smile or move our hands in ways that do not actually add anything substantial to the information we exchange.[39] Thus, our conversation comprises different, but often simultaneous levels - multiple interaction goals at any one moment.

A helpful distinction is often made between the syntactics, semantics, and pragmatics of language. While syntax refers to the way in which information is transmitted in particular codes or sounds, and semantics the conventions by which certain signs have agreed significance, pragmatics refers to the relationship between the participants, i.e. a focus which is:

... less on the sender-sign or receiver-sign relations and more on *the sender-receiver relation, as mediated by communication*. (Watzlawick et al., 1967: 4, original emphasis)

Watzlawick et al. elaborate on this particular comparison, also presenting it as between the *content* and the *relationship*. Along the lines of Bateson's earlier work, they use a physiological analogy:

... let *A*, *B*, and *C* be a linear chain of neurons. Then the firing of neuron *B* is both a "report" that neuron *A* has fired and a "command" for neuron *C* to fire. The report aspect of a message conveys information and is, therefore, synonymous in human communication with the *content* of the message. It may be about anything that is communicable regardless of whether the particular information is true or false, valid, invalid, or undecidable. The command aspect, on the other hand, refers to what sort of a message it is to be taken as, and, therefore, ultimately to the *relationship* between the communicants. All such relationship statements are about one or several of the following assertions: "This is how I see myself...this is how I see you...this is how I see you seeing me..." and so forth in theoretically infinite regress. (Watzlawick et al., 1967: 32-33, original emphasis)

While we have the 'content' on one level, on another we have the 'metacommunication', the '*information about this information*' (Watzlawick et al., 1967: 34-35, original emphasis), or, as Goffman put it, the 'special set of acts and gestures comprising communication about communicating' (Goffman, 1966: 99). This is always important: a relational frame or lens via which

---

[39] Sometimes, we even nod and agree when we have no idea what the other person is talking about.

the information is shared. Sometimes, though, relational goals, or the phatic dimension, have greater prominence in our exchange.

The phatic

The term 'phatic communion' was first used by Malinowski (1923) to describe:

> ... a type of speech in which ties of union are created by a mere exchange of words (Malinowski, 1923: 478)

He elsewhere describes his idea of 'pure social intercourse', which is:

> ... when the object of talk is not to achieve some aim but the exchange of words almost as an end in itself (Malinowski, 1923: 475)

Although his main object of study is the 'primitive' language of 'savage tribes' (Malinowski, 1923: 451) he acknowledged that this apparently "basic" form of communication applies extensively elsewhere too:

> A mere phrase of politeness, in use as much among savage tribes as in a European drawing room, fulfils a function to which the meaning of its words is almost completely irrelevant. Enquiries about health, comments on weather, affirmations of some supremely obvious state of things - all such are exchanged, not in order to inform, not in this case to connect people in action, certainly not in order to express any thought. (Malinowski, 1923: 476)

The goal of these types of communicative exchanges is for bonding and affiliation - serving what Malinowski calls 'the fundamental tendency which makes the mere presence of others a necessity for man' (Malinowski, 1923: 477; compare Baumeister and Leary's 'need to belong' - Chapter 1).[40] Jakobson (1960) also employed the concept of the phatic, writing on the distinction between linguistics and poetry. He presented a triadic model of the function of language, carrying referential, emotive and conative functions - all of which concern 'the information it carries' from the addresser to the addressee (Jakobson, 1960: 354). Beyond this, he describes the contact or the phatic function as:

> ... messages primarily serving to establish, to prolong, or to discontinue communication, to check whether the channel works ("Hello, do you hear me?"), to attract the attention of the interlocutor or to confirm his continued attention ("Are you listening?" or in Shakespearean

---

[40] Importantly, the inability to communicate in this way can be very unhelpful: 'Now speech is the intimate correlate of this tendency, for, to a natural man, another man's silence is not a reassuring factor, but, on the contrary, something alarming and dangerous**.** The stranger who cannot speak the language is to all savage tribesmen a natural enemy.' (Malinowski, 1923: 477) I shall go into more detail on how the potential (or inability) to interact forms an important basis for category divides in Chapter 4.

diction, "Lend me your ears" - and on the other end of the wire "Um-hum!").' (Jakobson, 1960: 355)

This has a similar, relational purpose.[41] Later, in a review of how the concept of 'phatic communion' had developed since this earlier work, Coupland et al. (1992) expanded it beyond just affiliation to encompass other 'relational goals'. In general, these are:

> Goals of talk that relate to building, modifying, or dissolving personal relationships, and, on the other hand, those that have to do with the definition and redefinition of own and others' identities as interacting beings (Coupland et al., 1992: 211)

Particular ways of communicating, including nonverbal behavioural cues such as gaze or posture, do not communicate "information" but rather act to define relational *frames*, which then continue to shape the whole encounter. These frames can be categorised along mutually exclusive dimensions of affiliation and dominance, alongside a degree of involvement (see Dillard et al., 1996; 1999), or according to Burgoon and Hale's (1984) more extensive list of twelve relational messages (including task orientation, involvement and affection. The goal of these communicative acts is therefore to define the relationship, which in turn shapes how we interpret the words spoken, the content of the speech, within this frame.[42]

In this way, we have some communicative acts which convey information about the world, while other signals indicate and shape the relational context - they have, respectively, transactional or relational *goals*. Importantly, however, this is not a sharp divide; the same types of behaviour might be used in both categories, according to context. For example, nodding one's head is often just a signal that you are listening, that you are following (or to give the impression that you are engaging), i.e. it serves a relational goal. However, this immediately changes when you are asked a question - nodding then has a transactional purpose, showing that you do agree, or that you will do something. These different goals surface at different points in the exchange; we might say that the interaction moves from being more *intrinsically-* to more *extrinsically*-oriented (see Cross, forthcoming: 11). At the start of an encounter, or at the end, our actions enact a relational goal, whereas, at other points, this behaviour remains in those back-channel responses, acting as a "channel" for the information, but the transactional is emphasised - the interaction becomes extrinsically-oriented.

Coupland et al. (1992) used the example of "how are you?" questions. The distinction between treating this question either as relational (perhaps rhetorical), or as an occasion where the speaker really wants to know the truth of the matter (i.e. where a transactional goal is more salient), is

---

[41] Albeit, as part of Jakobson's more traditional approach - describing our use of specific 'messages' conveyed from speaker to listener.
[42] See Burgoon et al. (1984) for examples of particular clusters of nonverbal cues; and the discussion of this 'meta-communicative' dimension above.

particularly prominent in the context of health interviews - the data that Coupland et al. analysed. They found that the different goals behind this particular question are negotiated over the course of the conversation, as both participants work together to decide what the current, salient goal of the exchange should be: they 'feel their way toward or away from the phaticity of their conversational exchanges' (Coupland et al., 1992: 219).

These two interaction goal orientations both still rely on (and themselves comprise) forms of common ground: the participants decide jointly whether certain actions are relevant in a relational, or in a transactional dimension. In just the same way as transactional goals necessitate grounding about the external world, achieving alignment on some idea, establishing a relational frame also becomes common knowledge for both participants, who understand more about how they relate to each other. Even the awareness of whether some behaviour or utterance itself is intended in a relational or a transactional sense should be something which is shared. (Think of the possibility that someone might ask "how are you?" to be met with an unwanted account of various financial or health-related difficulties.) Thus, I suggest that common ground is necessary, both in relational and transactional dimensions of the interaction.

Next, I shall explore the experience of embarrassment. The motivation to avoid this, and to 'save face' (Goffman, 1967: 5), is, I argue, a particularly important relational goal which shapes the way we interact and associate (or not) with other people; embarrassment could be construed as a shared understanding (a common ground) of the failure to achieve this end.[43] This interesting case will become more important in the light of research into social identity and group formation (to be discussed in the next chapter); I shall explore how this moderates the effect of interaction on group formation in my experimental work (Chapter 6).

Creating good impressions

When we interact with other people, we are not just concerned with communicating something well, but also with keeping a good image of ourselves in their eyes, and in the eyes of others who might be listening (or who they might talk to later). This motivation is important for our own self-esteem; even this is a highly social phenomenon - it is tied to how we think others perceive us, based on our experience of interacting with them. This is an important motivation in unfocused interactions - we feel embarrassed if we trip up clumsily in front of others, for example, and hope that no-one we know was there to see us. Here, however, I focus on the need to save face, to protect our self-

---

[43] Of course, the cause of our embarrassment might be something external. For example, I might inadvertently steal your shopping trolley, or spill wine over your carpet. This shared external focus, however, has consequences for what you think of me - and thus for our relational goals. In this way, however, embarrassment spans both of these categories, testing the limits of our present framework.

esteem, within more focussed interactions - the conversational exchanges that have been our primary focus so far.

Face-work

Irving Goffman describes how, in general, we aim to project acceptable 'selves', or positive images of ourselves (Goffman, 1956: 268). This is an important concern in our everyday interactions - not just in what we tell others (the transactional domain), but in how we interact with them, respecting ritual etiquette in order to preserve an affiliative relationship. Penelope Brown and colleagues (1987) discuss this in terms of 'politeness': we take great care over the way we say things. Instead of getting straight to the point, we might introduce any request we have, for example, with:

> "Look, I'm terribly sorry to bother you, would it be awfully inconvenient if..." (Brown et al., 1987: 57)

This is a 'strategic' use of language to ensure we gain (and maintain) our partner's approval (Brown et al., 1987: 56). We are not concerned only with saving our own 'face' (to use Goffman's term once again), but also fulfil our obligations to our partner to protect their image, perhaps tactfully avoiding certain topics or trying to continue a normal conversation despite another's errors of social judgement.

> The elements of a social encounter, then, consist of effectively projected claims to an acceptable self and the confirmation of like claims on the part of the others. (Goffman, 1956: 268)

Sometimes, though, this careful performance starts to unravel. An error of judgement - a joke that falls flat, or, worse, one that is taken as an insult - threatens this state of affairs. We first try to gloss over the event, maybe hiding the fact that we noticed their *faux pas*. Laughter, teasing or other forms of 'banter' are techniques we use to diffuse the potential tension, trying to 'reduce the seriousness of conflict by denying reality to the situation' (Goffman, 1956: 271). Edelmann (1987) also describes how the embarrassed individual and their 'audience' together use laughter as a coping mechanism 'in order to change the meaning and focus of the situation', transforming it from something embarrassing to something merely funny (Edelmann, 1987: 86). In other words, we try to maintain focus on our previous common ground - the ideas we had previously about each other and our relationship.

However, sometimes this does not work. Goffman describes the experience of 'breaking frame' within a conversational group, perhaps with 'paroxysms of laughter':

> The little social system they created in interaction collapses; they draw apart or hurriedly try to assume a new set of roles. (Goffman, 1956: 267)

Embarrassment

The experience of embarrassment is a key part of this process. It is an emotion which is generally shared, in response to the failure of one interactant or the other to save face, to protect their perceived self-esteem (or that of the other). Individuals are motivated to hide their embarrassment, in an attempt to maintain the previous state of play. They might avoid another's gaze, while regaining composure. However, once the event has come into the open, and individuals are appropriately embarrassed, the participants have together acquired new common ground, the 'new set of roles' and relational understandings which they then have to deal with in some way.

Specifically, embarrassment:

> ... occurs whenever an individual is felt to have projected incompatible definitions of himself before those present (Goffman, 1956: 264)

Importantly, this discrepancy, or the new understanding of a person, is new information which is then *shared* (or thought to be; this is information which we had previously tried to hide). For Modigliani (1971), embarrassment is understood as:

> ... a feeling of inadequacy precipitated by the belief that one's presented self appears deficient to others (Modigliani, 1971: 15)

It comprises a loss of self-esteem in a particular situation, crucially related to the perceived observation and negative evaluation from others. Modigliani provides evidence for this definition from an experimental study in which subjects experience either success or failure, which is either mitigated or unmitigated, and either in public or private. Embarrassment following failure is generally stronger in the public condition (although a mild form might occur in private, perhaps due to the anticipated presence of others); when it is stronger, it was followed by more attempts to explain away, excuse, or even deny the failure in some way.

Felt embarrassment, a wish to retreat from the situation, or an experience of shame,[44] is often complemented by an external, visual signal, predominantly in the form of blushing. Given our motivation to hide our embarrassment, this would seem to be a maladaptive response. However, this might in fact function to *restore* one's image in the eyes of others (perhaps in the same way as our attempts to regain some esteem through excusing our actions). The blush might act as a signal of our acknowledgement of failure, or of our misconduct, thereby showing others that we do in fact share those values or norms that we have transgressed. It is an honest signal of important moral common ground.

---

[44] Although, embarrassment and shame are arguably dissociable to some extent - see discussion in Castelfranchi & Poggi (1990).

> Blushing lets others know that some value against which we are, or are believed to be, inadequate is nonetheless a value to us, a value we sincerely share. Thereby they can distinguish us, sincere maintainers of the group's values, from other potentially dangerous individuals who only pretend to share such values, but in fact do not, and therefore do not have the right to belong the group. (Castelfranchi & Poggi, 1990: 241)

In summary, the motivation to avoid or hide our embarrassment (as part of the general performance of 'face-work' - Goffman, 1967), and the display of embarrassment when it is mutually recognised, serve to maintain or manage certain aspects of our relational common ground. This is a mutual process, as it concerns not only our perceived self-esteem, but also our ideas about what the other can perceive about us, as a result of our interaction. We control these perceptions together, deploying coping strategies or distracting our attention - always jointly.

This is just another case of how our spoken, dyadic conversations are intricate, reciprocal performances, responding to the actions of others in real time, in order to work towards a range of possible goals, simultaneously, which might be more or less transactional or relational. Maintaining a positive relationship is one relational goal, achieved through 'face-work' and part of a multi-modal process of grounding. In the event of social mishaps, this positive common ground might be threatened, perhaps resulting in the uncomfortable feeling of shame or embarrassment (often shared by all present). While this has possible restorative functions (in the display of blushing), this subjective experience is likely to lead one to distance themselves from or leave the situation, and hope "never to see them again". In the next chapter, I move on to talk about those social relationships which do ensue, following social interaction. Self-esteem (or lack of it) is an important motivating factor shaping who we do (or do not) want to associate with.

Now, however, I shall broaden the scope, encompassing forms of interaction which *foreground* these transactional or relational dimensions - encounters which are more extrinsically- or intrinsically-oriented in their overall goal. We now bring in music-making, explore its place on this spectrum and how it is related to speech.

Types of interaction

So far, I have focussed on spontaneous, spoken conversation, held between two (or sometimes three) unacquainted individuals. The question of this thesis, however, is the efficacy of *musical* interaction in social bonding and *group* formation processes. In this section, I show that we can think about music-making in the same vein as spoken interaction, as part of the same 'communicative toolkit' (Cross, 2012b: 26). Music and speech have different emphases though, making them more or

less suitable for different occasions - and possibly, as I shall argue later (see Chapter 4), differentially effective for different types of group formation.[45]

Our communicative toolkit: music and speech

As I noted in my overall Introduction, music is generally relegated to a different category of behaviour altogether. Most often, this comprises rehearsed, scripted performances by specialist musicians, which are then enjoyed and evaluated by a listening audience. However, here, I study music in a participatory field - focussing on the co-ordinated action involved in the process of performing, rather than on the "music itself". We shall look at ontogenetic and phylogenetic origins and cross-cultural examples of this activity, revealing deep-seated links to speech and language. These provide convergent evidence for a common 'communicative toolkit' which supports both music and speech; both of these are manifestations of a capacity for social, interactive engagement, in which we share some form of common ground.

Music and speech in ontogeny

In our infancy, we become attached to our primary caregiver; this is important not only to protect us and provide the resources we need, but also as the context in which we learn how to relate to all the others around us. The neurobiological architecture supporting this pair bonding process was explored in Chapter 1; here I address in more detail the form of this behaviour. It involves a type of interaction which has been described as proto-language, and also proto-music - it spans both categories, undermines a possible dichotomy between them, and shows how music and speech are linked from early in development.

When we speak to infants, we automatically adopt a much more exaggerated way of talking. We speak more slowly, with more obvious rhythmic emphasis, larger contours and sing-song melodic lines. We smile at them, and might rock them at the same pace. We adjust how we do this depending on their mood - if we need to wake them or attract their attention, we use steeper contours and a higher overall pitch; the opposite is the case if we are trying to soothe them or put them to sleep.[46] These are features which hold cross-culturally - see Mechthild Papoušek's (1996b) summary of findings from a number of studies, recruiting German, American and Chinese parent-infant dyads (also see Trehub et al., 2015, on cross-cultural parallels) - and which make infant-directed speech sound more "musical". Indeed, Papoušek argues that this is our 'earliest form of music education',

---

[45] I define speech here as the synthesis of syntax, semantics, and pragmatics. Language is a prominent part of this (along with its correlates in gesture, gaze and other non-verbal behaviour), and can be used for transactional and/or relational goals.

[46] Falk (2004) talks about the 'putting the baby down' hypothesis as a potential factor important in the adaptive phylogeny of this behaviour: it is a form of interaction which can maintain the infant's attention at a distance.

which is just as - if not more - important as the structured schooling in music we might receive later (Papoušek, 1996b: 90).[47]

This is a multimodal interactive context, involving verbal, tactile and vestibular stimulation - complementing the infants' 'transmodal' perception (Papoušek, 1996b: 90, provides a description; see e.g. Phillips-Silver & Trainor, 2005, for an empirical demonstration). Specifically, the patterns of sound display an inbuilt or honest emotional salience, which has the effect of modulating the mood and arousal of the infant in an apparently natural way. Hanuš Papoušek (1996a) discusses the basis for this in mammalian signalling, such as separation calls, drawing upon the work of Trevarthen, who describes how recognisable patterns in motion are linked to affect and expressed in sound via the 'Intrinsic Motive Pulse', i.e. 'the body-moving rhythmic and emotionally modulated system' (see Trevarthen, 1999: 160). Trainor et al. (2000) provides empirical support for this, showing the acoustic parallels between expression of specific emotions, exhibited in both adult speech and the infant-directed register.

Importantly, this is not just one-way expression; rather, it is part of a reciprocal exchange - developing our ability to engage in turn-taking patterns of spoken behaviour crucial in conversation (see above). The caregiver synchronises their own actions, for example, rocking the baby at the same rate as their own speech; they also are synchronised with the pre-verbal utterances and movements of the infant.[48] The caregiver is also responsive to the contribution of the infant to this proto-conversation, encouraging vocal matching and developing simple interactive games where the caregiver's response is predictably contingent upon the infant's input (see description in Papoušek, 1996b).

Particular phrases come to *mean* something categorical, such as approval or disapproval, requests or invitations, to which the infant learns an appropriate response, and also copies. The caregiver models these pre-linguistic phrases, clearly segmenting units using exaggerated prosody and rhythmic patterns to highlight phonological divisions. Mimicking this, the infant gradually learns how to segment the melodic stream using consonants, thus acquiring speech and an understanding of language (see e.g. Saffran et al., 1996, on statistical learning as a potential mechanism supporting this), developing their vocabulary through associations with the external environment (recruiting shared attention - see Franco, 2005). In this way, infants can begin to acquire complex

---

[47] Many studies have shown how newborn infants are good at discriminating auditory features such as melodic contours and rhythmic patterns - often with a sensitivity higher than that of adults (see e.g. Hannon & Trehub, 2005, and Winkler et al., 2009; overall review in Trehub & Hannon, 2006). They also recognise sounds that are familiar e.g. their native language, and their mother's voice - based on their intra-uterine acoustic experience (see e.g. Ockleford et al., 1988)

[48] As I mentioned above (see footnote 36, p. 62), this interactional synchrony is associated with the development of secure attachment bonds and the infant's acquisition of empathy.

communicative abilities before they can walk. Eventually, linguistic speech diverges from spontaneous singing, engagement directly with music or dance (see Brandt et al., 2012, for an overview of how mature, culture-specific forms of musical engagement and language diverge from a common precursor).[49] The roots of this, however, are in behaviour which undermines the distinction between music and speech. The deeper history of these interactive behaviours also point to something similar (Morley, 2013).

Music and speech in phylogeny

Our earliest vocalisations were probably extensions of broader emotional signalling mechanisms, exhibited by a whole range of species. Modern humans developed voluntary control of the articulatory and breathing systems which support this, but our vocalisations are still very much linked to our emotional experience and expressions (perhaps localised in the right hemisphere - see Morley, 2013). Vocalisation is also bound together with body movement - Iain Morley refers to all such behaviour as 'gesture', just either 'vocal' or 'corporeal'. He proposes that these are linked by a motor co-ordinator (which would explain why we gesture at the same time as our words). Due to perception-action links (i.e. our mirror neuron system), the emotions we express in these multimodal channels are also experienced by others through contagion. Our experience of both music and speech (alongside dance, the infant-directed register, among other things) is rooted in this interlinked system of emotion, rhythm, gesture and vocalisation. This system is a core part of all social intercourse. In particular:

> Musical experience relies upon systematic use of a gestural system, including vocal tonal gesture relying on rhythmic cyclical muscular control, which exists to allow the expression of emotional state and the understanding of emotional state in others. (Morley, 2013: 254)

Fitch (2012) suggests a slightly different account: the root 'proto-language' was a communication system which recruited our capacities for entrainment and hierarchical pattern perception; from this we acquired the ability to then abstract an understanding of musical and linguistic *syntax*. Morley (2012), in response to this, locates the common thread not in syntax (or other features of the musical or linguistic surface), but in the performance of vocal and corporeal gesture and emotional expression. Thus, just as holding a conversation is, at root, an interactive, joint activity that we share with others (and then use this basis to scaffold some perhaps more abstract form of common

---

[49] Cross (2012b) actually suggests that music-making specifically might have emerged in its own right in the context of our extended altriciality (i.e. our particularly extended period of development from birth until adulthood). The more "distinct" forms of music, resembling almost a different category of behaviour, might exist in order to continue into adulthood those same 'juvenile exploratory patterns of thought and behaviour' as exhibited in the infant-directed register, the music-language combination, that we use in infancy (Cross, 2012b: 25).

ground), engaging in music (or dance, for that matter) is the same sort of shared performance, with an emotional salience underneath the common cultural meaning.

Cross (2012b) comments, that, for these reasons, music probably did not either precede or follow language in an evolutionary timeframe, but rather:

> ... music and language are best conceived of as having co-evolved as components of a generalized human communicative toolkit. (Cross, 2012b: 26)

A more firm distinction between them, something we have in the separate categories given for "music" and "speech" or "language" is likely, he argues, to be culture-specific, and not due to any real biological distinction. A broader category of communicative interaction, encompassing a greater variety of possible practices, is more helpful. I turn next to some particular examples (from the "West" and elsewhere in the world) which demonstrate the validity of this more flexible approach.

Music and speech in practice

Improvised musical traditions like jazz are often thought of as conversational, with "jamming" sessions comprising informal environments in which musical ideas are shared and developed, in a relaxed social setting. Sawyer (2005) advocates an approach which focusses on the pragmatics of such interaction, rather than 'formal features' such as syntax.

> Musicians in an ensemble communicate with each other, and these interactional patterns replicate the essential interactional processes found in all human communication. As we listen to a performance, we are exposed to the distilled essence of human sociality. (Sawyer, 2005: 47)

He explicitly compares jamming to company brainstorming sessions, or even a family resolving a financial crisis. The pragmatic components common to all these contexts include interactional synchrony (and the process of entrainment), as well as a more general 'emergent' property of the interaction, which comes about through interpersonal collaboration (i.e. the idea that 'the whole is greater than the sum of the parts' - Sawyer, 2005: 47-49).

Others have started to investigate processes like gaze in the context of music-making. Vandemoortele et al. (2018) seem to be among the first to employ mobile eye-tracking in investigating gaze behaviour during the rehearsal and performance of a trio. There are no consistent patterns that emerge, however (perhaps because it is a scripted performance, rather than something improvised in the moment - the latter might have been more like a spontaneous conversation).

Healey et al. (2005) incorporate the concept of the 'F-formation' in their analysis of a small musical ensemble. This was coined by Kendon (building on the work of Goffman) to describe the spatial arrangements of people in conversation, a 'shared interaction space' in which they can see and

respond to each other movement's and gestures (see Kendon, 1990; and Healey et al., 2005: 2). In Healey and colleagues' study, they argue that something similar is going on in free improvisation among seven instrumentalists. The musicians' movements mirror in some way the passing of ideas around the group, e.g. by focussing the group's attention more or less on certain individuals' performance.

> Participants jointly determine when and where changes in the musical trajectory will occur. They also form a collective filter through which ideas are either elaborated or abandoned. (Healey et al., 2005: 9)

Maya Gratier (2008) expanded the parallel between music and language to include shared understandings between interactants. She talks about processes of 'grounding' common to speech and music - again in the context of jazz improvisation. Examples of common ground include the necessary cultural shared understandings (a language of sorts) that come from an awareness of certain melodic ideas, the repertory of a certain jazz tradition or school, as well as a specific personal history of musical interactions with one's co-performer. Grounding occurs in the course of performance when one participant introduces a certain familiar 'lick' or 'riff', the other signals their recognition (perhaps through joining in to complete it), and they develop it together. She also suggests that the process of entrainment, here forming the framework of the 'groove', forms another sort of common ground that they work from and develop. Even the drummer can contribute in this conversation: perhaps, by interjecting motifs within the regular rhythmic structure, they are doing something 'akin to the acknowledging head-nod and "uh-huh" performed by listeners in conversational contexts' (Gratier, 2008: 99). In this way, musical improvisation might be just like a spoken conversation, not just in its use of gaze and interactional synchrony, but also in common ground, even external references, presenting meaningful phrases which are recognised and then commented upon in the course of the dialogue.

Work in ethnomusicology has also provided a number of examples which show the interchangeability of musical versus spoken forms of interaction, supporting the argument that they can (and indeed should) be thought of as the same overall *type* of behaviour, with a common architecture. For example, the vocal genres of the Suyá, in Brazil (as documented by Seeger), include a whole range of styles - from different types of speech ('everyday', 'bad', 'angry'), each of which involves a particular manner, and occur in different contexts; to ceremonial 'recitatives' and 'invocations', narratives or anecdotes known as 'what the old people tell'; and more fixed-form, structured songs (see Seeger, 1987: 26).

In other contexts, the lament genre provides an interesting case study. In fieldwork conducted amongst Karelian refugees in Finland during the mid-1980s, Elizabeth Tolbert (1994) documented the

Finnish-Karelian *itkuvirsi*, a 'riveting performance of grief' which mixes together song with weeping and a ritualised language. This is an improvised form, a so-called 'work of sorrow' (Tolbert, 1994: 181) drawing upon a collective, historic Karelian identity while commenting upon their present plight. It comprises a form of communication traditionally understood to transcend this world, whereby the lamenter can reach out to the souls of their deceased kin in *Tuonela*, and by extension, demonstrate a longing for their own memories of their homeland (Tolbert, 2001). Tolbert provides a description of the lament:

> Alina begins her lament with the traditional muffled sobbing ... Shielding her eyes with her hands, she sways gently in a circle and starts to lament, employing the icons of crying, the special lament words, and the sigh-like melody to contact her mother in *Tuonela*. (Tolbert, 1994: 184)

In this way, the individual voice does not just enact the grief, but also refers to itself through sigh-like melodies and performed sobs. Thus:

> ... the lines between emotion and reference, language and music, and form and content, are blurred. The lament both references grief and performs it, interleaving reference to emotion and its performative expression. A strict opposition between emotional music and referential language is further undermined by the heightened and intermediary forms of crying, singing, and speaking. (Tolbert, 2001: 86)

Similarly, the 'sung-texted weeping tradition' of the Kaluli society from the rainforest region of Papua New Guinea, reported by Steven Feld (1990), breaks down the traditional boundaries of music, song, and speech. There are five different patterns of weeping, each with a distinct societal function; they range from more controlled performances of sorrow, mimicking birdsong in e.g. descending melodic phrases, to more "naturalistic" expressions of shock or distress, with ragged breathing, tears, and high or falsetto vocalisations (see Feld, 1982). Feld describes the variety of forms: 'some more recited, some more sung, some more cried, some more improvised, some more composed' (Feld, 1990: 252). The *sa-yalab* is one particularly interesting example: it involves conversational, spontaneous speech, as the performers collaborate in presenting distinct but interrelating layers of language (in the "lift-up-over sounding" aesthetic), alongside the descending melodic line associated with a birdcall, and sobs, tears, and a wobbling vibrato voice (see Feld, 1990).

These examples all undermine a traditional dichotomy of music and speech, exhibiting practices in which these are combined and performed together. Interestingly, music and speech tie together not only in their interactive potential and embodied processes involved in their performance, but also in their communication of culturally-shared meaning. This picture of communicative interaction spans the relational and transactional, incorporating shared understandings, ritual meanings or semantics, even perhaps a correct syntax.

Interaction goals: differentiating music and speech

We have good reason to think of both music and speech as part of the same communicative toolkit - a capacity for interaction that manifests itself in a variety of ways through ontogeny and across cultures, modes of interaction which have common phylogenetic roots. However, there are some differences between more musical and more linguistic forms of interaction. In this section, I shall present a framework by which they can be distinguished, along a continuum, rather than in separate categories. For this, I first think of the different ways in which music and language *mean* things, and then present the implications of this in terms of the different *goal* orientations of music and speech - i.e. emphasising the relational or the transactional.

Locating musical meaning

Signs can mean through iconic, indexical or symbolic relationships to their object (see footnote 31, p. 58, on Peirce's semiotic system). Language is primarily symbolic, with signs that are generally arbitrary in form, but meaningful through convention, or a collectively-agreed system of correspondence. Our vocabulary is built upon indexical foundations, however, as we learn it through association. In infancy, we point to things, aiming to share attention with others towards objects of interest. We gradually assign labels to things, splitting up the world around us into our culture's categories. Symbols then emerge as we develop the ability to evoke those categories at a distance, without those external referents actually present (see discussion in Tolbert, 2001).

Knight and Lewis (2017) go further and describe how the symbolic, and often metaphorical, use of language is in fact a form of deceptive signalling, dependent upon mutual trust within the community lest this reversal of costly signalling is misinterpreted. They use the example of play - a context in which: "the aggressive actions that follow are not to be mistaken for real" (Knight & Lewis, 2017: 443). This is significant for the evolution of language:

> While the vocal signals of young primates tend to be stimulus bound, inflexible, nonsymbolic, and limbically controlled, their playful bodily antics are strikingly imaginative, unpredictable, incipiently symbolic, and cognitively controlled - all suggesting a point of departure for the evolutionary emergence of language (Knight & Lewis, 2017: 443).

Thus, language has the capacity to mean - from direct reference to these 'fake' metaphors - being the unique exemplar of a symbolic cognition that is possessed only by human primates. For this reason, we can use speech for both relational and transactional purposes, as discussed above - both commenting on external objects, and modulating the internal, affiliative (or otherwise) frame of our conversation. However, music has slightly different capacities to mean, and, for this reason, tends towards a different goal orientation.

There are a few scholarly approaches which provide some helpful approaches to how music might have or acquire certain types of meaning. First, Trevarthen (1999) describes how musicality encapsulates and denotes aspects of our universal lived experience through being a direct expression of our 'Intrinsic Motive Pulse', which includes the ways that we move and feel emotion. Musical contours are expressive of movement, they are 'audible gestures', as he describes vividly:

> It seems we have to go to the source of experience in *acting*, the generative images of *moving*, to find a convincing account, to balance the reasoned explanation and give it authenticity; and we have to give a central role in our explanation to the *sense of time* in moving - the time it takes to step, jump, glide, hit, grasp, lift, throw, caress, or to think and talk - the measures and tensions of time that originate in the mind inwardly and become an *output*. The dynamic repetitive impulse that is cultivated and remembered in music is present in the way we have to move in our bodies no matter what we are doing, and this same impulse is both anticipated and reflected in all the senses, synchronously, as they seek, pick up and assimilate overlapping and transforming images of the effects of moving. (Trevarthen, 1999: 157-158)

Second, Tolbert (2001) built on this same line of thought in her account of musical meaning as a form of mimesis. She focusses on the musical (especially the lamenting) voice as an honest, non-arbitrary signal of emotional expression, which is rooted (as Morley's work also shows) in corporeal movements and gestures. Combined with voluntary control and access to memory, vocal musical expression is a medium by which iconic and indexical gestures can be displaced from their context and thus evoked "at a distance". This creates the infrastructure for symbolic thought: signals, forms of meaning or representation, could be arbitrary, while retaining the 'indissoluble bond between vocal emotion and vocal reference' which makes the voice more trustworthy.

> The musical voice is above all a meta-commentary on the conditions of symbolic thought and the possibility of culture, and a meta-commentary on the possibility of language; it reminds us that so-called arbitrary signs are anchored to intercorporeal representations of emotionally guaranteed, cultural truth. (Tolbert, 2001: 91-92)

Third, Cross and Woodruff (2009) present a framework which includes a greater variety of ways in which music might mean. Beyond this mimetic, 'motivational-intentional' dimension, music is 'socio-intentional', resembling speech in the ways in which it directs our attention, providing a framework of entrainment and joint action 'to impart a sense of communicative intent' (Cross & Woodruff, 2009: 87). It can also be 'culturally-enactive' in the shared, semantic meanings that music evokes (with a degree of arbitrariness) in a particular sociocultural context. Importantly, however, those low-level biological significances of the motivational-intentional, or mimetic dimension, 'hold the experience of meaning in music on a leash' (Cross & Woodruff, 2009: 84). In this way, under the right conditions, music is capable of some of the same sorts of meaning as is language (with possible

semantic significance, and syntactical conventions).[50] However, language has a much greater preponderance of symbolic meaning, and a much greater capacity for specific reference.

> Whereas in language it is usually possible to specify the subject of an utterance with some precision, this is almost never the case for music. Music appears to be a strangely malleable and flexible phenomenon. The meaning or significance of a musical behaviour or of a piece of music can rarely be pinned down unambiguously; music appears to be inherently ambiguous (Cross, 2005b: 30).

This is a continuum, where they do meet in the middle (and even overlap): music can occasionally refer to something specific, and language can be ambiguous, such as the 'significatively-freighted elusiveness of poetry' (Cross & Woodruff, 2009: 89). However, it is always possible for two different people to interpret the meaning of music in slightly (or very) different ways. In the words of Cross, it has a 'floating intentionality':

> In effect, one and the same piece of music can bear quite different meanings for performer and listener, or for two different listeners; it might even bear multiple disparate meanings for a single listener or participant at a particular time. Music has a sort of "floating intentionality" (the word "intentionality" here simply means "aboutness"); it can be thought of as gathering meaning from the contexts within which it happens and in turn contributing meaning to those contexts. (Cross, 2005b: 30)

This means that, in general, music and language each tend 'towards opposite poles' of the communication spectrum:

> ... language and music should properly be distinguished as tending towards opposite poles on a continuum of capacity for *specificity* of meaning (Cross & Woodruff, 2009: 89)

And, therefore, they each might prioritise different communicative *goals.*

Locating musical goals

We have so far have been comparing music to our use of language. However, when we think about the different components of speech as a whole, we might recall that the specific meaning (which is possible with words and certain e.g. pointing, or "hurry up", gestures) is not always salient for the overall *purpose* of the exchange. For example, when we ask "how are you?", but only for the purpose of beginning our conversation, the actual symbolic meaning of the words is redundant - the phrase is

---

[50] Examples of arbitrary semantic significance of music include the association between the sound of the opening phrase of Beethoven's Fifth Symphony and the idea of 'Fate knocking at the door'; or the Bolivian "tara" being used to "call the rain" (see Stobart, 1996, 1994); moreover, the syntactical correctness of particular harmonic conventions has been linked to language using neuro-imaging evidence collected by Koelsch et al. (2002).

no more than a ritual expression. The relational becomes more important - the phrase begins our encounter in just the same way as a wave, an approach or a hug. On the other hand, a hug is not sufficient if we really want to know the details of the health or happiness of the other. For this transactional, or instrumental, purpose, we need more precise meaning - encoded within the semantics and syntax of words (and supported by the pragmatics, too). Because of the lack of specific symbolic meaning which can be encoded and so shared through music-making, this activity thus tends towards different goals in comparison to speech; specifically, prioritising the former, relational, intentions over the latter, transactional, goals.

Cross (2013) notes that speech can be inexplicit in 'phatic communion', for the purpose of establishing affiliation; however, there is always the potential for conflict:

> ... this function can always be undermined by the potential for our utterances to be interpreted not as tokens of recognition of each other's communicative presence, but as definite statements about the world that are capable of being contested. (Cross, 2013: 419)

On the other hand, in music, 'the phatic or relational dimension is foregrounded':

> In music, we cannot formulate or convey semantically decomposable propositions. But music has the advantage over language in the relational domain in that music sets up and maintains its affiliative, relational, frame, without its affiliative qualities having to be continually re-negotiated, and the individual significances that participants may attribute to the ongoing musical interaction are not required to be made mutually manifest in order for the interaction to be sustained and to succeed. (Cross, 2013: 419)

Music generally has a goal of affiliation, structuring social relationships - it has an intrinsic orientation:

> ... music as participation has a function that is intrinsic to it - simple continuation of the joint activity - with no extrinsic goal in view. (Cross, forthcoming: 11)

On the other hand, while the relational is important in facilitating the phatic channel of interaction in speech, this is generally coupled with a goal to communicate something, to share attention and understanding about something external, i.e. speech is more extrinsically-oriented in goal.

> ... conversational interactions can express a function that can be thought of as proper to speech and that is extrinsic to the interaction: the organisation of joint action for a mutually explicit purpose. (Cross, forthcoming: 11)

We discussed above, in relation to these relational and transactional goals in speech, that both necessitate sharing common ground, but perhaps common ground of different sorts. In communicating information about the world, we go through a process of grounding the interaction,

ensuring that this common ground is maintained as we build upon it, incrementally, with each exchange. In a similar way, through the process of 'face-work', we aim to maintain a particular sort of relational common ground - protecting our own positive image in the eyes of the other, and protecting theirs at the same time. Thus, there might be a continuum of common ground types, depending on goal orientation of the current interaction context. Music is a flexible medium, which can exhibit a range of goal types.

Musical common ground

In some types of musical interaction, there are more specific, external meanings or conventions, which need to be shared for the interaction to be successful.[51] This requires some background common ground, in just the same way as shared cultural knowledge, or the ability to speak a language. This has been shown in how people perceive certain musical tonal patterns, just by listening. For example, in Castellano et al. (1984), the Indian listeners understood the music presented as according to the system of tonal hierarchies acquired through familiarity with that style. The Western listeners, on the other hand, did not have this pre-requisite common ground, and understood it at face value, rating tone hierarchies according to their frequency of occurrence.

This common ground is important in performance, where musicians share common expectations with the others involved. In Maya Gratier's (2008) analysis of jazz improvisation, the "content" of the contributions of each performer is situated within the 'community of practice', using certain melodic gestures or riffs that are known from previous performances, or knowledge of the style in general. In a completely different context, the Kaluli ritual weeping involves emulating the sound of the *muni* bird, a call that has significance not only from the external landscape, but from the cultural worldview and mythologies about the soul's transition to the afterlife (see Feld, 1982). As Tolbert puts it, commenting on the voice of the Karelian lamenter:

> Her voice...creates a bridge between the world of the living and the world of the dead, the past
> and the present, weaving individual lives into collective remembrances. (Tolbert, 1994: 194)

However, in other contexts, we might say that there is a lower threshold of external common ground which is required for a successful performance. The music is more oriented towards maintaining a certain type of relationship (or "internal" common ground), not necessitating such group-specific knowledge. This lies more within Cross' category of the socio-intentional, rather than the culturally-enactive dimension of musical meaning (see Cross & Woodruff, 2009). Music is often just about that 'communicative intent', comprising:

---

[51] Here, I am combining the meaning that music might convey to a particular group, and the particular rules or conventions by which it might be performed. These often inform each other, and rely upon collective, shared *understanding* about the "music itself", or the actions involved in producing it.

> ... performative actions and sound structures that could be interpreted as affording cues about shared intentionality that direct attention in interaction. (Cross & Woodruff, 2009: 86)

This is an important part of language as well - in elements such as interactional synchrony and prosody, for example; however, this phatic dimension is more salient in music. In Gratier's analysis, much of the common ground she talks about in jazz pertains to this level of shared understanding. She suggests:

> ... displays of mutual understanding between musicians are rooted in a collaboratively negotiated embodied phrasing through which repetition, mirroring and matching, punctuation, and completion and synchronisation might constitute the musical basis for grounding. (Gratier, 2008: 71)

The timing of performer's phrasing, the co-ordination of particular moments of synchronisation or mimicry moment-to-moment, all involve shared awareness and continual grounding, but about the performance frame or relational context which is being enacted.

In summary, music prioritises more relational goals, particularly those which are affiliative, above any sort of goal which requires participants to share more specific external common ground for the interaction to succeed. For this reason, music has particular efficacy in managing 'situations of social uncertainty'. These include the infant-caregiver relationship,[52] and also intra- as well as inter-group contact (see Cross & Woodruff, 2009). In the remainder of this thesis, I explore intra- and inter-group processes in more detail (Chapter 3), and then consider how different types of music-making - i.e. those which vary along this continuum of extrinsic- to intrinsic-oriented goals - might have different functions in this context (see Chapter 4).

An analysis of the processes involved in conversation is useful to understand all sorts of interactive communication, encompassing forms of both music-making and speech. Both are part of the same communicative toolkit, and necessitate grounding - both in the relational or transactional dimension (what we might call internal or external common ground). Speech, in its use of language and gesture with more specific referential meaning, prioritises the transactional goals; whereas, music-making, with its essential ambiguity and 'floating intentionality' makes relational goals more salient. Importantly, these are not distinct categories, but different goal orientations, along a continuum from the most extrinsically- to more intrinsically-oriented encounters. Therefore, different sorts of interaction (even different sorts which seem "musical" to us) might be more or less efficacious in

---

[52] In this contact, the indefiniteness of music's interpretation might also aid the development of domain-general intellectual flexibility, or the ability to apply the same thing to multiple situations, a 'metaphorizing capacity'. This property is exemplified by the flexibility by which music might mean different things in different situations (see Cross, 2005b: 37).

mediating social relationships, depending on what prior knowledge the participants share about the interaction, and what they deem necessary to make music in that context.

In the next chapter, I shall shift focus, and provide a review of the current research on social groups. After laying out this background, I can present my thesis about how these areas overlap: how different types of interaction create different types of social bonds or groups.

Chapter 3 Social groups

Simon felt a perilous necessity to speak; but to speak in assembly was a terrible thing to him.

"Maybe," he said hesitantly, "maybe there is a beast."

The assembly cried out savagely and Ralph stood up in amazement.

"You, Simon? You believe in this?"

"I don't know," said Simon. His heartbeats were choking him. "But..."

...

"What I mean is ... maybe it's only us." (Lord of the Flies, William Golding, p. 82)

In our modern society, we have seen the great power and influence of social groups - with devastating effects in the first half of the twentieth century, and wars continuing up until the present day (think of Israel and Palestine, and Syria and Turkey). However, bloodshed caused by the co-ordinated efforts of one group pitted against the other is nothing new - it has been a feature of human existence, throughout history (and probably prehistory).[53] On the other side of the coin, practitioners and researchers have identified the benefits of group activity, from sports to singing in choirs, seeking to publicise these effects in the face of funding cuts. Engaging in these communities, in co-ordinated team activity, can have important effects on wellbeing and therapeutic potential.

In Chapter 2, I explored the components of social interaction, showing how we can characterise types of engagement on a continuum between, on one hand, an extrinsic orientation towards more transactional goals, or, on the other, a foregrounding of relational goals - or an intrinsic orientation. All sorts of encounters - whether spoken, sung, or somewhere in the middle - fall along this spectrum. In general, musical interaction has an intrinsic focus, whereas, introducing language in speech makes the transactional dimension salient - speech has a more extrinsic goal focus. This is a helpful framework through which we can investigate how different types of music-making (in all of its various manifestations, varying along this axis) might form a sense of group identity.

In this chapter, I address the notion of the *group*. I start with a historical review of approaches to this topic - mainly within social psychology, but acknowledging the significant influence of sociology. There have been various standpoints - some emphasising the inter-individual nature of the group, and others the idea of group identity as an internal, psychological property. I trace the move from a focus on interdependence between individuals, to later ideas about social identity and self-categorisation.

---

[53] The prevalence of intergroup discrimination can be much closer to home - even in our own attitudes. The institutional manifestations of this are only recently being addressed, with new training programmes designed to overcome implicit biases that we might hold without knowing it.

The respective importance of interdependence versus categorisation processes is an ongoing debate. In the second half of this chapter, I shall outline a few ways in which researchers have attempted to evaluate the contribution of each process, or to synthesise them - whether as mechanisms operating in parallel, resulting in different types of group, or as ways in which the same group might become more inclusive (or restrictive) to outsiders. I then build upon the last suggestion, proposing that these balancing forces of inclusion (via interdependence) and exclusion (through categorisation) might involve interaction which is more intrinsically or extrinsically oriented (see above). In this chapter, I present just the outline of this idea; I flesh it out, with more examples, in Chapter 4. This then forms the basis for my experimental work (see Chapters 5 and 6).

A brief history of social groups

Living and working in social groups has been a context which has transformed our species - changing the selection pressures, and enabling us to build up a whole wealth of cultural norms. The ability we have to relate to other people, and communicate with them in various gestures and symbols, relies on our unique cognitive capacities which developed within this context, and which continue to make it possible (see Chapter 1).

A preference for those who are part of the same tribe, group, or family, is adaptive for the individual; distrust or even outright aggression towards those who are in neighbouring (perhaps competing) groups may also emerge.[54] We enact our social bonds with certain individuals (and withdraw such affiliative links with others) through our interaction with them, perhaps in making conversation or in collaborative music-making. Now that I have explored how to characterise these forms of interaction, I can present various approaches to the group as a social unit. This is a more or less separate domain of research. However, I outline this with a view to synthesising these topics, thinking about how different modes of interaction might shape the group formation process.

Research on social groups has a very long history, formed through the intersection of sociology (especially the work of Simmel, Durkheim and others) with psychology. The ideas which emerged have since formed the subject of a plethora of empirical, often laboratory-based social psychological studies, all operationalising "group-ish" feeling or in-group bias with a whole range of measures.[55] I begin this chapter by providing some necessary background, giving a rather whistle-stop tour of the main theories, as they were developed, in more or less chronological order.

---

[54] Although see Manson & Wrangham (1991) and replies, for discussion on this topic.
[55] These have included explicit preference ratings, attribution of certain stereotypic traits, or actual behaviour towards others, alongside more implicit measurements of the same judgments, using response times (as in the Implicit Association Test - see Greenwald et al., 1998), memory biases or the use of abstract language (see a review of how in-group favouritism and out-group derogation are measured, and compared, in Hewstone et al., 2002).

Individual and social: do groups exist?

An emerging field of social psychology, as it began to tackle the question of social groups, encountered key questions fairly quickly. Primarily: how should we think about psychology - something generally restricted to the domain of the individual - in relation to the social - phenomena which necessarily transcend the realm of individuals?

This tension is explored in a critique by Pepitone (1981) (where the same debates were apparently surfacing, even into the 1980s). On one hand, there was a general feeling amongst psychologists that 'the individual was the only reality' (Pepitone, 1981: 974). In a telling and oft-quoted remark from the 1920s, Allport claimed: "There is no psychology of groups which is not essentially and entirely a psychology of the individual" (quoted in Pepitone, 1981: 976). On the other hand, other ideas were surfacing: 'global, yet unseen, nonpalpable notions like the group mind', a collective consciousness, and other social phenomena (Pepitone, 1981: 974). These phenomena have a reality of their own - a "castle in the air" (Simmel & Hughes, 1949: 260) - drawn from, but irreducible to, the individuals that comprise them. In Durkheim's words, we have two 'states of consciousness', between our being as an individual and our part of society, which form an often conflicting 'double existence' (Durkheim, 2005: 43-44).[56] The relationship between the social and the merely individual was the sociologists' turf; however, an exploration of *social* psychology needed to engage with these ideas.[57]

However, psychological perspectives retained an 'individuocentric bias': a focus on individual (inter-individual, at most) processes (see Steiner's lecture, given in 1974 - 'Whatever Happened to the Group in Social Psychology'). The influence of behaviourism (from earlier in the twentieth century) meant that viable theories only concerned the observables, namely, our responses to any environmental stimuli (whether this is a reward, a threat, or another human). Psychologists emulated the methods of the "harder" empirical sciences: these "objective" natural sciences were felt to be somehow more reliable, or valuable; the methods successful there were automatically deemed appropriate here, too. However, Pepitone argued that this approach was 'inadequate to explain social behavior' [sic] (p. 972):

> ... it would appear that a major lesson from history is that inasmuch as social psychology is
> conceived as part of a natural-science oriented, general psychology of the individual, its theories
> cannot adequately deal with the influences on personality and social behavior that originate in
> the objective environment, including especially the social structures and normative systems in
> which individuals are embedded and psychologically subscribed. (Pepitone, 1981: 983)

---

[56] The idea that these two 'states' exist within and guide the actions of individuals has a few parallels with the social identity and self-categorisation approaches (these are explored later in this chapter).

[57] This became especially pertinent after the experience of war and economic depression in the 1930s and 40s raised many questions about our social behaviour and psychology: people started to think about how social organisations function, and how charismatic leaders could influence large crowds of thinking individuals.

Given that these social systems have real effects both on our subjective experience of the world and on the way we actually behave, it was important to develop social psychological theories which could take account of these phenomena in a more helpful way. Sociologists were more fruitful, working in the other direction. They acknowledged that:

> ... society cannot constitute itself unless it penetrates individual consciousnesses and fashions them "in its image and likeness"; so, ...it can be said with confidence that a number of our mental states, including some of the most essential, have a social origin. Here it is the whole that, to a large extent, constitutes the part; hence it is impossible to try to explain the whole without explaining the part, if only as an after-effect. (Durkheim, 2005: 35)

In this section, I shall review a number of attempts made by psychologists to develop a theory of the individual that resonated with the knowledge of some sort of social consciousness.

Group cohesion

One early approach to social groups was to understand how individuals calculate the potential costs and benefits of group membership. Group cohesion is a metric for this, signalling 'the degree to which the members of the group desire to remain in the group' (Cartwright, 1968: 91). The attractiveness of one's group is affected by properties of the other members, how similar they are to each other, the overall group goals as well as the effectiveness of the group in attaining these. A group might be more attractive if has a greater degree of 'co-operative interdependence', especially in the face of some (maybe shared) external threat. This group is compared to attractiveness of alternative groups, keeping in mind potential costs of leaving (or of moving from one group to another). The resulting group-level cohesion is thus associated with how concerned individuals are for their membership of that group, their motivation to participate with the others, and to contribute to the group's activities (see Cartwright, 1968).

Group interdependence

Calculations of group cohesion were fairly individualistic in scope. 'Functionalistic theories' (Thibaut & Kelley, 1959: 5) considered more the *relationships* between individuals; these drew on game theory to model how individuals might work together to achieve shared aims. Thibaut and Kelley's social interdependence hypothesis, a 'theory of interpersonal relations and group functioning' (Thibaut & Kelley, 1959: 1), is one example of such an approach. This still highlighted the benefits offered to the individual by the group:

> Because the existence of the group is based solely upon the participation and satisfaction of the individuals comprising it, the group functionalism becomes an individual functionalism. The

ultimate analysis then is in terms of the vicissitudes of individuals as they try out various adaptations to the problems confronting them. (Thibaut & Kelley, 1959: 5)

These 'various adaptations', however, involve *joint* behaviour of interdependent group members - behaviour which can be modelled using co-ordination games such as the Prisoner's Dilemma. Participants work towards shared goals, outcomes which comprise the 'operating consensus about a desirable state of a given task' (Thibaut & Kelley, 1959: 257). They adapt their choices to their expectations of other participants, mutually shaping the outcomes of the group as a whole. Working together often results in greater rewards than might be possible on one's own; but this depends upon mutual co-operation and trust. General, shared norms can also develop: a common framework about how members of the group "should" act across a variety of situations.

This resembles the definition of a social system, or 'collectivity', provided by sociologists Parsons and Shils (1951). This is, namely, a:

> ... social system having the three properties of collective goals, shared goals, and of being a single system of interaction with boundaries defined by incumbency in the roles constituting the system. (Parsons & Shils, 1951: 192)

In other words, individuals have their own, complementary roles, enabling them to work together in 'concerted action' (Parsons & Shils, 1951: 55). This theory goes beyond simple interdependence, however: individuals display *solidarity* in their common (even 'harmonious') values and expectations (Parsons & Shils, 1951: 193-4). In this way, there seems to be a group function or identity which is emergent from (or goes beyond) the individuals involved. It does not depend on the identity of the members, as long as the relationships and outcomes are maintained. In other words:

> ... if the group's resources enable it to withstand the loss of several members, without very dramatic changes in its structure or functioning or in the outcomes achieved by the remaining persons, we might decide to consider this collectivity as maintaining its identity even though there are minor fluctuations in the size and composition of the group. (Thibaut & Kelley, 1959: 192)

Group entitativity

In some way, then, through joint interaction towards a shared goal, a collection of individuals becomes a group, an entity, or some sort of "thing" in itself. Shaw (1976) provided a review of these various definitions of the group, synthesising them into two assumptions:

> If a group exists, ... members (1) are motivated to join the group (and hence expect that it will satisfy some of their needs), and (2) are aware of its existence, i.e., their perceptions are veridical. (Shaw, 1976: 11)

Shaw here traced a middle path between the individualism of contemporary psychology and the more holistic idea of the group as a real, collective consciousness. He argued that groups can be a real phenomenon, but which come about through interaction between individuals. This 'realness', or 'entitativity', is quantifiable; objects in the environment are the most "real", and groups vary in their 'entitativity' below this. This comes from Gestalt principles of perception - namely, the idea that the whole is emergent from, or more than the sum of, its parts (Shaw drew on Campbell's work in his discussion of this).[58] In particular, the common fate of the group's component features (i.e. individuals), and also their similarity and proximity to one another, were found to be reliable cues for the perceived 'entitativity' of a group. Shaw concluded:

> Since the only basis we have for attributing reality to any object derives from our perception of it, we must conclude (with Campbell) that a group is real to the extent that it is perceived as an entity. (Shaw, 1976: 14)

These psychological theories begin to resemble more the sociology of earlier decades. Simmel and Hughes (1949) had also emphasised the association, or 'togetherness' between individuals - a different level of reality, or thing (or even 'art'), which is distinct from the lives of those separate, interacting personalities. They even discard the notion of shared goals: the crucial part is 'the pure essence of association'; sociability is 'a legitimate end in itself' (Simmel & Hughes, 1949: 255, 259).[59] Moreover, a group's internal structure, or the component relationships which contribute to its 'entitativity' in a Gestalt fashion, had been modelled (albeit rather abstractly) in various ways - notably Granovetter's theory of weak and strong ties (see Freeman et al., 1992).

Intergroup relations

A convincing case for the argument that the group is a real entity in its own right - and not just a collection of individuals - came with Carolyn and Muzafer Sherifs' field studies investigating *inter*group relations. In providing a context for boys to take part in fun activities, with or against their peers, the researchers constructed a whole new social reality which provoked individuals to act in ways that were extraordinary - wildly different from their normal behaviour (and not in a way which could be justified as a response to just a new environmental stimulus). This behaviour stemmed from awareness not only of the boys' own group membership, but also of a different group of which they were not a part: the reality of group phenomena emerged in the relationship *between* groups.

---

[58] The higher-order system emerging from a collection of individual components, via Gestalt processes of perception, is exemplified by how a melody still sounds more or less the same regardless of whether it is the key of C or G (even though the notes used would be very different in each case). (This is the example given in a historical review of these developments by Turner et al., 1987: 12-13)

[59] This resembles Cross' definition of intrinsic-oriented interaction (see Chapter 2). As Simmel and Hughes put it, the motivation for interaction is: 'a feeling of the worth of association as such, a drive which presses toward this form of existence and often only later calls forth that objective content which carries the particular association along' (Simmel & Hughes, 1949: 255)

Sherif (1967) reported three experiments (the second had to be terminated due to excessive hostility), which were run in naturalistic settings - three outdoor sports camps - and recruited groups of 11-12-year-old boys who had no previous acquaintance. All the boys were from fairly affluent, middle class backgrounds, and were generally healthy and happy - there were no known underlying individual factors which might have affected their behaviour. First of all, the experimenters found that friendships formed spontaneously (due to common interests or similarity, when the boys were left to their own devices) were effectively nullified when the experimenters intervened only to divide the boys into two different groups. Second, after group members took part together in activities such as camping, cooking, or transporting canoes, the groups quickly developed their own internal structures, with clear leaders and norms such as in-jokes, preferred places, and even sanctions for those who did not fit in. These group activities were formulated in terms of interdependence:

> ... conditions that embody goals with common appeal value to the individuals that require interdependent activities for their attainment (Sherif, 1967: 80).

They observed that, as a result, the boys developed a sense of "we", as opposed to the "they" - the latter only on the basis of a mere awareness that another group existed.

Third, they brought the two groups together in a tournament of games. The groups competed in tasks such as a tug of war, baseball, and a treasure hunt. Despite being fairly normal games, these conditions alone created solidarity within the groups and a strong feeling of hostility between them: members of one group branded the others as "cheats"; the experimenters reported that one group set alight a banner belonging to the other; and the groups refused to spend time together even in pleasant, easy activities such as dinner or watching a film.

> If an outside observer had entered the situation at this point, with no information about preceding events, he could only have concluded on the basis of their behavior [sic] that these boys (who were the "cream of the crop" in their communities) were wicked, disturbed, and vicious bunches of youngsters. (Sherif, 1967: 85)

It was just the conditions of the intergroup tasks - an interesting reversal of the conditions of interdependence - which seemed to promote such a strong shift in the boys' attitudes and behaviour.

> The *sufficient condition* for the rise of hostile and aggressive deeds (including raids on each other's cabins with destruction of property) and for the standardization of social distance justified by derogatory images of the out-group was the existence of two groups competing for goals that only one group could attain, to the dismay and frustration of the other group. (Sherif, 1967: 85)

Finally, the researchers investigated what conditions might possibly reduce this hostility. Mere contact was not enough (eating together, for instance, turned into "garbage wars"). Rather, the conflict only eased through re-constructing conditions of positive interdependence, tasks with goals which were rewarding for both groups, and which they could only achieve through working together. These 'superordinate goals' were created through tasks such as finding the source of a leak in a pipe; together tugging a broken-down truck (which carried a delivery of their food); and all contributing towards the cost of a film.

> Joint efforts in situations such as these did not *immediately* dispel hostility. But gradually, the series of activities requiring interdependent action reduced conflict and hostility between the groups. (Sherif, 1967: 89)

This was effective only cumulatively; eventually, they did begin to form friendships with members of the other group, and could all sit together (they chose to take the same bus home).

In this case, the experience of being in a group seemed to come about through functional relationships, whether of co-operation or competition. Such conditions resulted in within-group preference and clear out-group derogation - as manifested in differential attitudes, stereotypes and behaviour, as well as a sense of loyalty to the in-group. Group-motivated behaviour such as hostility to the out-group did not stem from particularly aggressive individuals, but rather came from whole-group interests which were then owned and embodied by the individual members (who acted perhaps out of character). These intergroup dynamics were emergent; they exhibited a group-level reality that surpassed individual relations. We turn next to consider the subsequent phase of social group scholarship. This explored how such external, group relations could indeed have an individual, psychological impact.

Groups in the lab: social categorisation and identity

These demonstrations of group-level phenomena are corroborated by our own experience: the relationship between groups of fans of two rival football teams following a match, or the potential for conflict between neighbouring countries which are competing for the same natural resource. What was still unclear was how the subjective experience of group membership - on an individual level - could arise, following a concern for the group as a whole. Turner and colleagues called this 'the *psychological* reality of the group' (Turner et al., 1987: 24).

All of the work so far prioritised face-to-face contact as the way in which individuals both generated and enacted these group divisions. However, during the 1970s, when researchers studied the minimal conditions under which inter-group behaviour might emerge, they discovered that this might not be necessary.

The Minimal Group Paradigm (MGP)

Tajfel, Turner and colleagues set out to identify the minimal conditions of group formation which prompted individuals to display in-group bias. Although they expected some conditions of interdependence to be necessary, they actually found that participants showed a clear preference for other in-group members in their control condition. That is to say, individuals who were assigned a team identity (according to trivial criteria: classed as an 'under-estimator' or 'over-estimator', or a preference for paintings by Klee or Kandinsky) and then given the chance to reward other individuals who either belonged to their group or to the other group (designated only according to code numbers on their questionnaire, and with the awareness that this did not at all affect the rewards they would receive), then assigned more points to the in-group. Participants actually preferred to reward *relatively* more to the in-group than the out-group, even if this meant they awarded fewer points overall - the relative gain of the in-group above the out-group, rather than an absolute point total, was more important to them (see original experiment in Tajfel et al., 1971).

Interpreting the MGP: Social Identity Theory

These findings were interpreted as evidence that the subjective experience of belonging to a group, and the resulting in-group bias, arises from a process of mere categorisation. In the experiment, participants never met or interacted with any others, neither in- nor out-group members, so face-to-face interdependence could not have been a *necessary* part of the process. This did not deny the importance of interpersonal relationships or of intergroup contact (as we saw earlier). Rather, the findings of the MGP suggested that the group concept, the psychological state of group membership, comes about initially through categorisation: individuals identify themselves and others with a particular group or category. This formed the basis for Social Identity Theory (first presented in Tajfel & Turner, 1979).

A summary of the theory, presented by Hogg & Abrams (1988), emphasised the changes that occur within the individual, as a result of them being assigned to a particular group. They state:

> ... belonging to a group (of whatever size and distribution) is largely a *psychological* state which is quite distinct from that of being a unique and separate individual, and...it confers *social identity*, or a shared/collective representation of who one is and how one should behave. (Hogg & Abrams, 1988: 3, original emphasis)

We categorise people into groups as part of a general need to find order in our environment: perceiving regularities in this way is a process of uncertainty reduction, making the world a bit more predictable. We notice commonalities among people, helping us to form stereotypes about how we expect certain sorts of people to behave. This process has much broader implications - with real

consequences for how we think about ourselves and others. When we categorise ourselves as part of a particular group, we internalise this new identity as part of our self-concept (Brewer, 1991); we emphasise our own traits that align with the properties we have in common with other members (forming group stereotypes - see Haslam et al., 1999).

Our self-concept is comprised of several identities - some on a personal level and others social identities (see Brewer & Gardner, 1996). All of these contribute to how we think about ourselves, our self-evaluation, and thus impact our well-being, sense of self-worth, or (as is more common in the literature) our self-esteem. Our need for positive self-esteem in relation to our social identity - that associated with membership of a social group - motivates us to show our group's 'positive distinctiveness' (Hogg & Abrams, 1988: 23).[60] This becomes especially relevant in relation to other groups (as we saw in the Sherifs' work, above). Historically, there are many examples of large-scale social groups, categorisations based on nationality or religion, who are pitted against each other just on the basis of those differences, fighting for power or status over the other (Hogg and Abrams use the example of Protestants and Catholics in Northern Ireland). The sense of in-group preference, or inclination towards out-group derogation, therefore comes about through internalisation of this social identity as part of one's individual sense of self, combined with one's own need for "coming out on top".[61]

There has been some confusion over the role of self-esteem in this process, with rather mixed empirical findings. Hogg and Abrams (1990) identified three different hypotheses or 'corollaries', clarifying potential aspects of this process: inter-group discrimination results in higher self-esteem; self-esteem promotes inter-group discrimination; or mere category salience leads to higher self-esteem. They reviewed various studies, finding most support for the first of these corollaries (as did the more recent literature review by Rubin & Hewstone, 1998): most evidence indicated that successful in-group discrimination enhances self-esteem (but not vice versa).[62]

Interestingly, Crocker and Luhtanen (1990) distinguished the role of personal and more collective self-esteem in predicting these 'enhancement' motives, or in-group preference. Those with an

---

[60] Other motivations which might contribute include a need for coherence in our self-identity, a sense of meaning (which can come from upholding our social norms or stereotypes), self-efficacy or a motivation for power and control over others. These are self-*evaluative* motivations, rather than a drive for self-enhancement, and are likely to be more important day-to-day - although the latter might arise under conditions of threat (Hogg & Abrams, 1990). See also Amiot and Aubin (2013) for insights from self-determination theory: the effect of whether our identities are freely chosen or imposed upon us.

[61] Other strategies to attain positive group distinctiveness and self-esteem might instead include social mobility (changing group), if possible, or other sorts of social change (see Hogg & Abrams, 1988).

[62] These uncertainties about the precise role of self-esteem might be due to how it is measured. Often, empirical studies measure or manipulate *global* self-esteem. In contrast, Social Identity Theory implicates a more *specific* self-concept about a social identity which comes to the fore only in certain contexts (and so is much harder to measure; see Hogg & Abrams, 1990).

already high sense of *collective* self-esteem were more likely to discriminate in favour of their in-group, upholding this self-evaluation through preference for their own group identity. On the other hand, those with lower collective self-esteem do not feel the need to increase their self-evaluation; rather, they act consistently, and display less in-group bias. This provides some support for the second corollary: individual differences in self-esteem motivate in-group bias.[63]

Broadly, the social identity we acquire through being assigned to (or choosing) a social group becomes an important part of how we think about ourselves in relation to others. Although there are mixed findings, in general a sense of group identity, and a preference for those who belong to our group, stems from our personal wellbeing and our need to uphold a positive sense of self.

Self-Categorization Theory

Following the development of Social Identity Theory, Turner and colleagues investigated more the general 'social-cognitive basis' which underpins this sense of group membership. Their theory of Self-Categorisation (presented in Turner et al., 1987) formulated with more precision the way in which we align ourselves with certain types and combinations of social identities.[64]

> The fundamental idea is that group behaviour is the behaviour of individuals acting on the basis of a categorization of self and others at a social, more "inclusive" or "higher order" level of abstraction than that involved in the categorization of people as distinct, individual persons. (Turner et al., 1987: 2)

Turner et al. describe three levels on which we classify others: on a species level, as human; on the level of social groups, based on certain similarities and differences (e.g. male or female, working or middle class); and then as individuals, noticing the ways that we all are different, even within our own social categories (Turner et al., 1987: 45). Each of these layers of definition is a different facet of one's self-concept; they function independently, and can become more or less salient depending on the context (see also Brewer & Gardner, 1996). Our social identities operate on this medium level of abstraction: we do not always engage with others just as individuals in their own right, but rather as members of a category (Turner et al., 1987: 42).[65]

---

[63] Aberson et al. (2000) endorse most of these findings in their meta-analysis; however, they also reveal that individuals with low collective self-esteem do show some in-group preference, but only in indirect rather than explicit measures (in keeping with their need for self-consistency).

[64] Confusingly, they refer to Social Identity Theory as the 'social identity theory of intergroup behaviour' (Turner et al., 1987: 42), referring to the motivational basis for positive differentiation of one's own group above others) and their new Self-Categorisation Theory as 'the social identity theory of the group' (Turner et al., 1987: ix). Importantly, the realm covered by Social Identity Theory is *within* their broader model of Self-Categorisation.

[65] If one level of categorisation is activated, this inhibits any other categorisation based on other similarities or differences (due to a 'functional antagonism' between levels of categorisation - Turner et al., 1987: 49).

Working up from the bottom of this framework is often thought of as a process of 'depersonalization'. Although this term has negative connotations, it refers here to a purely perceptual phenomenon in the way we view all other people, and ourselves. Turner et al. draw on the work of Rosch and Bruner, who investigated how we develop hierarchical classifications for everything in our environment - from chairs, to trees, and now to people. Dividing people into social groups does not necessarily imply judgements of value; rather, these divisions merely come from trying to maximise intra-class similarities and inter-class differences to make the most sense out of the present surroundings (the meta-contrast principle - Turner et al., 1987: 46-47).

Of course, what seem to be purely perceptual processes do end up, as we know, having real (often negative) effects in the way we think about and treat others. When we are primed to think on a group level, we do not think of others as individuals, but just as members of a category - with all the associated attributes or stereotypes.[66] Importantly, this shapes our interaction with them, responding to them on the basis of these heuristics - the accounts of racial bias displayed by police in the United States, is a good (albeit often brutal) example of this.[67]

Category divides and bounded affiliation

We generally mimic people who we like, but not others who we have more negative feelings about (Likowski et al., 2008). Relatedly, we mimic people more during conversation when they belong to our in-group (e.g. on the basis of religious belief) than those in the out-group (Yabar et al., 2006). We are also more likely to laugh at jokes when we hear fellow university members laughing than when we think the laughter comes from members of an unpopular political party (Platow et al., 2005); and we only spontaneously mimic the emotional expressions of politicians when our views are aligned with theirs (Fino et al., 2019). Even when we are divided by only minimal criteria (like certain art preferences), we are more co-ordinated in performing a simple movement with an in-group than an out-group member (Miles et al., 2011). In general, then, phatic forms of interaction, simple and easy ways of affiliating with others, are already marked by the existence of category divides, or an awareness of having different social identities.[68] Kavanagh and Winkielman (2016) suggest that this intergroup mimicry effect is the result of social learning biases whereby we emulate only those who are reliable models, i.e. those from within the group.

---

[66] In Cambridge during the summer season, the streets are often full of tourist groups coming from overseas, especially China. On the basis of this experience, I have observed my own tendency to group individuals according to race, which involves a stereotyping of Chinese as tourists and white British as students or academics. This categorisation is a simplification of the demographic of the university town, and is unsuitable to be applied more generally in other University contexts (i.e. generally, Chinese people working in the library are not there as tourists).

[67] See, for example, the news report of a politician who was stopped by police for trivial reasons - and assumed it was because of the colour of his skin (see Revesz, writing in the Independent in July 2016).

[68] See Blocker & McIntosh (2017) for a comparison of how mimicry is shaped by different types of group boundaries - whether minimal, interdependent, socially consequential etc.

This, of course, has consequences for how we think about others, and how much we like or affiliate with them. Hess and Fischer note that:

> ... mimicry serves to foster affiliation but also crucially depends on an initial affiliative stance, that is, an initial openness to engage with the other (Hess & Fischer, 2013)

This is even shown on a neural level. When viewing photographs of certain 'extreme outgroups', we do not even activate the prefrontal cortex (central in our social cognition), instead processing the images of those people in the same way as we would objects; the insula and amygdala are activated - areas associated with our experience of disgust (Harris & Fiske, 2006). Even following just a minimal basis for categorisation (a dot estimation task), we show a larger neural response to in-group than out-group faces (Ratner & Amodio, 2013). Any mimicry of outgroups might emerge in fact through a motive of appeasement - in response to someone who is perceived as a threat (see fMRI study of this effect by Rauchbauer et al., 2015).

These group boundaries are also characterised by a limit of our capacity to empathise (see Fuchs, 2019). The motor resonance which lies behind this - enabling our understanding of others' pain, or other mental states - is least responsive when we observe people who we think belong to e.g. racial outgroups (among Canadians, these are the South Asian and Black individuals; Gutsell & Inzlicht, 2010). We are less likely to "feel their pain" (Avenanti et al., 2010; Xu et al., 2009), or experience sadness along with them (Gutsell & Inzlicht, 2012).

As a result of these actual processing differences, people feel less close or connected to those who are different from them (according to some criteria). It follows that we are more likely to help or trust those who are within our group, or who look the same as us in some way. In other words, we reinforce this group boundary by the way we interact and affiliate (or don't).

In summary, after a century of research into social groups, we can account for a broad range of phenomena - including our membership and sense of belonging to a larger-scale category, regardless of whether we have had personal experience of interacting with all of the other members, or none of them (or just a tiny fraction). The Social Identity and Self-Categorisation theories made this possible by highlighting the role of categorisation above a sense of interpersonal interdependence. Furthermore, they offered a way to re-consider our question of reconciling the domain of the social with that of the individual: here, we situate 'the social self-concept...in the mind of the individual' (Turner et al., 1987: 67). Social groups are not "out there" in society, but are represented on a psychological level, and mediate our individual choices and behaviour.

> ... society is in the individual as much as individuals are in society. (Turner & Oakes, 1986: 239-240)

However, this was not the end of the story. The respective roles of face-to-face interdependence and social categorisation (in the mind of the individual) in forming social groups (and the ensuing definition of what a social group is) are still the subject of debate. In the next section, I focus on more recent approaches to this question, how researchers have synthesised and combined these processes in various ways, before presenting my own ideas about how we might distinguish them - by thinking more about the types of social interaction which might be involved.

Interdependence and categorisation

These Social Categorisation and Identity theories did not replace older work into the role of interpersonal contact and functional group relationships. Rather, they provided new insights into possible psychological processes at work. However, there were competing claims about which process is primary, which is more important, or occurs first, in establishing group identity. In this section, I shall explore this dialogue. I start by outlining competing interpretations of the MGP; then I suggest how these different mechanisms might actually work in tandem and be complementary insights - they are two sides of the coin to consider when approaching questions of group formation and identity.

Interdependence and categorisation in minimal groups

Initially, the main criticism of the Social Identity interpretation of the MGP came from Rabbie, Horwitz and colleagues. They argued that participants still had a sense of interdependence, even in those minimal conditions: their task was to allocate rewards to others, thus they were aware that others would be doing the same, with them in mind.[69] They acknowledged that processes of categorisation do occur as we align ourselves with others who have some common attribute; however, they suggest that these categories only become meaningful units, or groups, with relationships of interdependence:

> A social group can be considered as a 'dynamic whole' or social system, characterized by the perceived interdependence among its members, whereas a social category can be defined as a collection of individuals who share at least one attribute in common. (Rabbie & Horwitz, 1988: 117)

We might categorise ourselves with those who are blonde, who share the same toothpaste, or who also play the cello. However, they argue, until there are some conditions which encourage us to act together, we will not be a meaningful unit. If, one day, others start to derogate the intelligence of

---

[69] The MGP actually involves both members of both groups having the opportunity to reward all other participants, regardless of group membership; however, other experimental variants have reversed these effects with clearer conditions of interdependence (favouring the in-group or the out-group accordingly - see Rabbie et al., 1987 (discussed in Rabbie & Horwitz, 1988).

blondes, we might perceive the need to defend ourselves, or demonstrate in some other way - acting together (even if not face-to-face) - the value of this social group. Identification is therefore with the outcome of the group goal, via perceived interdependence. The process of categorisation, by itself, is unlikely to evoke a sense of real belongingness or in-group preference:

> There is no evidence to date that the experimenter's *ad hoc* classifications of subjects into Blues and Greens or Kandinskis and Klees were internalized by them into their self-concepts. If so, their self-concepts must be extraordinarily malleable. (Rabbie & Horwitz, 1988: 120)

Turner and Bourhis (1996) responded to these criticisms by arguing that categorisation involves more than just noticing similarities. The act of categorisation means that we represent ourselves and others as belonging to this group or another; this is psychologically significant, and 'a relevant basis for [our] behaviour'. This collective awareness of group membership does go along with displays of interdependence; in fact, interdependence might be either a cause or an effect of such categorisation. But, he argues, the crucial process in group formation is an internal process of categorisation and social identification with that category.

Therefore, we have two possible accounts of the mechanism behind group formation - either as something emerging from real, instrumental, interdependence; or as the result of a process of categorisation and identification. The latter might come from such external interrelationships, but these are translated into something individually meaningful via representations (and motivations) occurring primarily on a cognitive level. More recently, other researchers have attempted to tease apart (or synthesise) these processes; I shall review this empirical work next.

A container for reciprocity

Like Rabbie and Horwitz, Yamagishi and colleagues argue that there is a 'residue' of interdependence in the MGP. The knowledge that we share group membership (or categorisation) with others is perhaps linked to an understanding that my actions will affect other group members, and that I have an obligation to them (as they do to me). This is not just in-group liking or preference, but an expectation or norm of reciprocity (stemming from a sense of outcome interdependence - see earlier discussion of Thibaut and Kelley). This is not inconsistent with a Social Identity approach; however, any categorisation process is inseparable from a root perception of interdependence - and this is a key part of group belonging.

> The defining feature that makes a group distinct from a simple aggregation is the existence of actual or imaginary interactions. No sociologist would dare to attach the term *group* to an aggregate of people who simply share some characteristics or category unless they believe at least that they are connected somehow with others and have the potential to influence each other. (Yamagishi & Kiyonari, 2000: 116)

Yamagishi and Kiyonari provided evidence for this by manipulating the conditions in which participants played Prisoner's Dilemma games - specifically, their knowledge of each other's group membership, i.e. whether or not they *both* knew they were in the same group. This is an important precursor for the influence of co-operative norms, which rely on a categorisation which is known to be shared:

> ... the mere fact that the player and the partner belonged to the same group was not sufficient to produce in-group favoritism [sic]: unless the player was aware that the partner knew they shared group membership, the partner's group membership had no effect on the level of his or her cooperation. (Yamagishi & Kiyonari, 2000: 119)

They also investigated these expectations more directly by comparing choices made in simultaneous versus sequential Prisoner's Dilemma games. In the former, participants were led to believe that them and their partner allocate rewards simultaneously; whereas, in the latter set-up, they thought they were the first player to allocate a reward (which the second player would see before deciding how much to reciprocate).[70] Choices in simultaneous games rely on a generic norm of co-operation, while those in sequential games on more direct reciprocity. They find that an in-group preference arises only in simultaneous games; this effect vanished when direct reciprocity became more relevant. Therefore, being categorised as part of a group comes along with general expectations of in-group reciprocity; as they put it,

> ... the social category is the container of the expectations of in-group reciprocity. (Yamagishi & Kiyonari, 2000: 127)

They go further, claiming that these are closely intertwined:

> ... the empty container itself can play no role in producing in-group favoritism in the minimal group situation unless it is filled with expectations of in-group reciprocity. (Yamagishi & Kiyonari, 2000: 127)

Our experience of being in a group (or our 'group heuristic' - see Yamagishi & Mifune, 2008: 7) necessarily incorporates an expectation of mutuality. This makes sense when we consider the beneficial effects of group living in our evolutionary history - we co-ordinate in hunting, gathering (and fighting for) other resources, and rely on others to protect us (as we do the same for them).[71]

Since these findings, others have also tried to discern whether categorisation or interdependence is more important in group formation. For example, Balliet et al. (2014) directly compared the predictive power of Social Identity (and Categorisation) theories with the hypotheses from this

---

[70] Actually, the participant was playing alone (with only a virtual "partner").
[71] See Cosmides & Tooby (1992) on domain-specific cognitive adaptation to these expectations of social exchange.

theory of 'bounded generalised reciprocity'. They conducted a meta-analysis of in-group preference observed in various economic games. In this case, their basis for distinguishing which mechanism is at play was whether the presence of out-group members has an effect, as compared to 'unclassified strangers'. The meta-contrast principle of Social Identity Theories implies the importance of a distinctive out-group for in-group preference to emerge (see also Hartstone & Augoustinos, 1995); whereas, the idea of the in-group as a context for reciprocity does not require this obvious comparison. They also compared the role of indirect versus direct reciprocity on in-group preference. Overall, they provide more support for the reciprocity account: the presence of a distinctive out-group does not seem important, and, again, the presence of a means for direct reciprocity undermines any effect of group membership.

Interdependence and categorisation in tandem

It seems likely, therefore, that there is a place for interdependence, probably *alongside* processes of categorisation. Stroebe et al. (2005) conducted a study where they manipulated participants' strength of identification (through the belief that they are either typical or not typical of their assigned group) and the relationships of outcome dependence (i.e. interdependence) they have with other participants (either working with other in-group, out-group members, with both - as in the original MGP, or none - in a control, no-dependence condition). As expected, according to the categorisation account, they showed a main effect of identification on in-group allocations in the control, no-dependence condition: the strong identification condition led to more within-group altruism. Simultaneously, they found support for an 'unbounded reciprocity' account: when participants are interdependent with out-group members (compared to interdependence with all participants and to the no-dependence condition), they displayed higher expectations for out-group reciprocity, and allocated bigger rewards across the group divide. However, the effects on in-group preference were stronger when the interdependence relationships were with other in-group members; this shows an integration of categorisation with a sense of *bounded* reciprocity within the group, both influencing participants' choices. In the light of this, they conclude:

> ... it can be fruitful to study ways of integrating social identity and reciprocity-based approaches to intergroup discrimination in the MGP, rather than testing the two approaches separately. (Stroebe et al., 2005: 844)

Other researchers echoed this call. Platow et al. (2012) showed that categorisation produces interdependence, and vice versa; they also criticised a 'propensity toward dichotomy', arguing that both interdependence and categorisation are likely to be important in group formation (and should not be considered separate "schools of thought"). In a similar vein, Durrheim et al. (2016) traced the changing roles of interdependence and categorisation, each emerging as important contributors to in-group preference according to current conditions.

Therefore, it seems important not to dichotomize interdependence and categorisation. Rather, we should explore which of the two processes might be more or less salient, and the different effects they each might have in structuring the social environment. The studies we move to next depart from the exclusive focus on economic games, and show how these mechanisms might work in tandem in slightly more natural settings.

Types of entitativity

Another recent body of work harks back to the notion of the group as a type of 'entity'. This is the idea that the cohesion, or "group-ish-ness", of social groups is measurable on a scale of 'entitativity' (a term that is drawn from Campbell's work - see above).[72] Research has suggested that we use certain cues in perceiving how far a collection of individuals can be considered 'a meaningful group, as one entity'; these are the Gestalt properties which underlie our perception of any objects' interrelationship (or combination into a whole) in terms of e.g. common fate, proximity or similarity (Rutchick et al., 2008). If we reconsider individuals' perceived interdependence, or their common categorisation, not as opposing mechanisms but rather as different *cues* for how individuals relate to one another, then different *types* of entities, or groups with different types of 'entitativity', might result.

Ip et al. (2006) ran a series of experiments, varying which cues corresponded to group membership (among cartoon creatures) and asking observers to rate the groups' 'entitativity'. They found higher ratings, both for groups that move together (with apparent interdependence), and for groups that all had the same body colour (perhaps indicating common category membership). Rutchick et al. (2008) investigated the possibility that these cues actually signalled a different way of thinking about or 'construing' groups. They primed participants with images of either dynamic interdependence (thinking about a hive of bees), or categorical similarities and differences (thinking of bee species), and then gave them a description of a social group. Under the first condition, participants were more likely to remember features of the members' *interaction*, whereas under the second, they remembered how the members were *similar*. Thus, these cues for entitativity both predict perceived group formation; however, they operate under (and contribute to) different ways of thinking about groups. They even claim that: 'any given group can (in principle) be construed categorically or dynamically', the crucial difference being only 'the mindset of the perceiver' (Rutchick et al., 2008: 905, 908).

Others have identified an analogous division in self-report data: we consider the everyday societies or social units we encounter (or are part of) to be *either* dynamic (interdependent) or categorical groups. Prentice et al. (1994) conducted a couple of questionnaire studies, each recruiting individuals

---

[72] Here, cohesion is used in a different sense to Cartwright's concept of group cohesion I discuss above.

involved in various university societies. They found two distinct characterisations of groups - some being 'common-bond groups' (based on interpersonal attraction) and others 'common-identity groups' (based on some common trait). As above, these are distinguished by the ways in which the individuals are related; in this case, either attachment among group members, or attachment to the external identity of the group.[73] They describe how these group types have different ways of operating, perhaps with different degrees of loyalty, longevity, and ways of responding to conflict.[74]

Therefore, categorisation and interdependence are not opposing mechanisms, but instead comprise different ways of representing groups (that we could potentially switch between),[75] as well as different bases (or cues) for interrelationships among individuals in order for them, as a collective, to be considered an 'entity' of one sort or another. Next, I shall look at research which continues our focus on categorisation and interdependence as two potential processes or cues underlying the formation or delineation of groups; instead, however, it considers how they act simultaneously, perhaps as opposing forces in our perception of where the group boundaries lie.

Shifting groups: de- and re-categorisation processes

A great deal of research has considered how personalising interaction (a form of interdependence, as we saw in Chapter 2) might reverse the 'depersonalization' processes involved in social categorisation. While, as we saw above, category divides limit our social relationships, affiliative interaction can also undermine those boundaries. When people are instructed to mimic others who belong to a different racial category, or a racial out-group, they then show less implicit prejudice than people who just watched without mimicking (Inzlicht et al., 2012). When we synchronise with others in an out-group, or in another team, we are also more likely to want to affiliate, or even just to be near them, than we were before (e.g. Reddish et al., 2016; Tunçgenç & Cohen, 2016); co-operation following synchronous movement reaches the same level as shown normally amongst in-group members (Cross et al., 2019) and seems to restructure our cognitive representations of those category divides (Good et al., 2017).[76] In this way, the psychological effects of social categorisation (forming distinct groups by a top-down process, resulting in intergroup discrimination in the way we behave towards others) might be moderated by conditions of interdependence, perhaps even re-categorising people into more inclusive groups.

---

[73] Hogg and Turner (1985) manipulate categorisation and group liking in a similar but much earlier study.
[74] Lickel et al. (2000) also classify a range of groups according to similar criteria, this time noting more specific differences between intimacy groups, task groups, social categories, and loose associations.
[75] Compare how, previously, only categorisation was the basis for the 'psychological group'.
[76] Even just listening to music can generate a similar 'empathic resonance'. Individuals (with high dispositional empathy to begin with) identified with, and therefore showed reduced levels of implicit prejudice for, racial outgroups - specifically Indian or West African people - after they had listened intently to Indian or West African popular music, i.e. when they had "allow[ed themselves] to be immersed in the music" (Clarke et al., 2015).

In order for these encounters to reduce intergroup bias, it is important for the new interdependent relationships to influence (from the bottom up) broader processes of categorisation - concerning the whole group, not just those select individuals. Within the research on so-called 'contact theories', there have been a number of approaches to this problem. For example, in their 'common in-group identity model', Gaertner and colleagues explore the possibility that individuals can change the cognitive representation of their in-group, expanding it to others via experiences of interaction with out-group members (see e.g. Gaertner et al., 1993). In this way, the out-group becomes assimilated within a broader in-group identity. Of course, possible problems with this, played out in the real world, include the dangers of assimilating minority ethnic groups within a more powerful group, and the threat to valued social identity that this might entail (see review in Hewstone et al., 2002). Conversely, models of crossed categorisation suggest that *increasing* the number of potential category divisions might reduce overall bias usually occurring across just one dimension; Mullen et al. (2001) provide evidence for the efficacy of this in certain conditions, but conclude that it is more just a 'redirection of bias'.

Thomas Pettigrew (1998) provides a more multi-faceted approach which includes a number of these separate models, but within a longitudinal framework. Initial contact must have the following conditions: equal group status, common goals, interdependence rather than competition, and the support of the authorities and wider society (rules originally presented in Allport's contact hypothesis) as well as 'friendship potential' (repeated contact in a variety of settings - Pettigrew, 1998: 80). This initial contact promotes decategorisation, more specific liking and reduced anxiety towards members of the out-group. Following this, the category divide needs to become salient again to promote generalisation towards other out-group members; recategorisation might occur, creating a unified group identity which is satisfactory for both groups concerned. A meta-analysis showed the effectiveness of this sort of contact situation across a range of out-group targets, successfully generalising to the group as a whole (Pettigrew & Tropp, 2006).

In this thesis, I expand upon this idea, bringing in insights about the nature of social interaction (explored in Chapter 2). Now that we have seen how processes of categorisation and interdependence might operate in tandem, each shaping our perception of where group boundaries lie, I shall turn to discuss how various forms of interaction might enact each of these processes - forming group boundaries through categorisation, but also re-structuring those relationships through affiliative interdependence. My proposal, which I set out in brief below (and expand upon in the next chapter), is that different components of interaction might realise *both* processes of interdependence and of categorisation, depending on the goal orientation of that interaction context. While Gaertner et al. (1990), Good et al. (2017) and others have investigated the process of forming and re-forming group boundaries through affiliative interaction, I go beyond the question of

just overcoming category divides, seeking how interaction contributes to group bonding and to group identification processes. I draw on the distinction between intrinsic- and extrinsic-oriented interaction goals as a helpful way of characterising different forms of social (including musical) engagement. These different emphases, shaping the way we interact with others, have implications for the interpersonal relations which result - either dividing us into categories, or instead bringing us together in conditions of interdependence.

The role of interaction - a proposal

When we make music, especially when our involvement is simple (like clapping along, or joining in with a well-known song), it is easy to engage with others who we do not know, and with whom we might not have anything else in common. The fewer background references we need to "get", or conventions we need to understand, the more likely it is that we will be able to join in. Music is an affiliative form of interaction; when we make music with others, we are more likely to like them, perhaps even overcoming any previous prejudices or reservations.[77]

When we hold a conversation, however, the threshold of external common ground we each need to access is much higher. In talking, we generate and maintain an affiliative context, a channel through which we can converse; however, for our interaction to be at all successful, and for this channel to remain open, we must at least share some common, cultural knowledge, especially a common language - and often, as well, the life experience which enables us to understand certain concepts or vocabulary.

We speak of "language barriers". It is this common ground - distinguishing who can take part from those who cannot - which forms the basis of important social categorisations. These requirements form the necessary "containers", inside of which reciprocity and affiliation is then possible (to use the terms of Yamagishi and colleagues). In a very similar way, musical practices which rely on extensive training, access to a score, or cultural knowledge of the correct way to interact, restrict who can take part to a certain, often quite select, few.

These forms of interaction vary along a continuum. Speech (and some forms of music-making) requires more external common ground; it is more extrinsic-focussed in goal. Simple music-making, however, often requires less of this external common ground; music foregrounds the phatic, becoming more intrinsically-oriented, e.g. its focus being primarily to co-ordinate with one's partner(s). I propose that, in moving along this continuum, making our interactions more extrinsically-focussed, raising the threshold of the external common ground or knowledge we need,

---

[77] See a slightly different framing of this idea in terms of 'The Biosociology of Solidarity' - i.e. how 'sustaining the established rhythm' contributes to a sense of solidarity, or belonging (Kalkhoff et al., 2011: 937).

we diminish the *potential* for interdependence, simultaneously providing a basis for social categorisation. Moving in the opposite direction, our interaction becomes more relational, more intrinsically-focussed in goal, with much less external common ground required to take part. These are conditions of interdependence which are much more inclusive, expanding the potential for affiliative group formation. Features of the interaction which were oriented towards transactional goals, or external common ground, cease to form category divides; instead, we have a new, wider basis for group identity, involving all who want to take part (rather than just those who share the same language of interaction).

In this way, the different components and goal orientations of social interaction contribute either to processes of categorisation, restricting relations of interdependence to those who are included within the in-group, or to conditions of affiliative interdependence. Our respective mechanisms of categorisation and interdependence thereby become more or less salient depending on the current *interaction* context. In my next chapter, I draw on a range of literature - from empirical studies on synchrony, prosody and empathy, examples from work in ethnomusicology, to discussions of group norms and conventions - to provide a comprehensive outline of this theory. This then forms the basis for my empirical work, in which I provide some initial, more direct evidence for these ideas.

Chapter 4 Musical groups

So the LORD scattered them abroad from thence upon the face of all the earth: and they left off to build the city. Therefore is the name of it called Babel; because the LORD did there confound the language of all the earth: and from thence did the LORD scatter them abroad upon the face of all the earth. (Genesis 11: 8-9, Authorized King James Version)

We exist in community - we have a social brain (Chapter 1) and interact with those around us in a variety of ways, with ease (Chapter 2). This leads to a sense of group identity: we affiliate selectively with known others or others with whom we share some common attribute (Chapter 3). The way we interact with others, and our feeling of belonging to a group, are tied together; this relationship is the subject of the present chapter. In his writing on sociolinguistic variation, Daniel Nettle (1999) shows that dialect is important in our perception of group boundaries, distinguishing those who we can trust from any outsiders, or possible free-riders. I build upon this work, suggesting that the way we interact is not just a social cue, but also performs those social relationships, maintaining groups and their boundaries.

So far, we have addressed the nature of our social interaction - the intricate performance of relational and transactional goals, and how these are each manifested (to a greater or lesser extent) in the way we talk to, or make music with, others. We have also considered the definition of a social group - viewed either as an interdependent collection of people, or as a broader social category. Here, I shall synthesise these areas of research to help us understand how our social interaction might relate to our group identity. My basic premise is as follows: on one hand, interdependent, relational interactions bring people together; while, on the other hand, the external common ground necessary for more extrinsically-focussed encounters forms the basis for categorisation, creating boundaries between who is included and who is not. In this chapter, I shall present this thesis in full.

Music-making is an activity that can bring many people together. The recent film *Bohemian Rhapsody*, a dramatic biopic tracing the life of Freddie Mercury, includes a scene recreating the performance of 'We Will Rock You', in a huge venue, for Live Aid. This depicted the emotional potency of a whole sea of people performing together the basic hand-clap rhythm and repeated lyrics which form the backbone of the song. This is social interaction on a large scale.[78] Everyone is focussing merely on the need to stay in time, in synchrony with the rest of the crowd. Nobody is communicating any information to others around them, beyond the act of moving and clapping together. We might say that this is interaction with an intrinsic-oriented goal - foregrounding the

---

[78] The huge scale of group interaction portrayed here (and as occurs often in mass sporting events) is somewhat unusual in our evolutionary history - it is only relatively recently that humans have gathered in such numbers.

phatic dimension of communication, with no need to understand or keep track of anything beyond the mere act of engaging and acting jointly with others. This is rewarding, and everybody feels like they belong - at least temporarily - in that social group.

We see this phenomenon often - among football supporters, or other massed choirs. But these occasions are not always so inclusive. We have all had the experience of being (it seems) the only one who does not know the words; while others around us are singing, we try to hum along and smile, but actually just feel slightly out-of-place and embarrassed. Only the best singers can arrive at a Hallelujah chorus 'sing-along' event without at least having a copy of the score. In these, and many other, contexts, we need to possess (or have access to) a higher threshold of external common ground or knowledge in order to take part, and thus feel we belong. The type of music is more extrinsic-oriented in its focus: we are not just moving together, but we are singing *something* together. The necessary common ground thus becomes the basis for social categorisation, forming a boundary around those involved in that interacting, bonded group. We need to meet certain prerequisites before we can feel we belong.

Here, I shall draw on various examples of these phenomena to provide some convergent evidence for this framework. In the first half, I explore the ways we interact and so *include*: how the phatic aspects of social interaction, and so the modes of interaction which prioritise these (being more intrinsic-focussed in goal), serve to bind people together into interdependent groups, increasing their liking and trust for one another. In the second part of this chapter, I look at the converse - how the way we interact *excludes* others. I discuss how those more transactional aspects of interaction necessitate external common ground (being more extrinsic in focus); it is these demands which form boundaries between social groups, and can become the basis for our in-group and out-group categorisations. Music-making can span the whole range, varying in goal focus, and thus in power to include and exclude. I give examples of both of these forces in action and highlight how my experimental work (reported in Chapter 6) shall provide some initial empirical evidence for these ideas.

Social bonding: intrinsic-oriented interaction

I start by considering the components of interaction through which we bond. In Chapter 2, I explored phatic communion: encounters which maintain the "channel" of our interaction more than communicating something concrete about the world. As we shall see, it is this dimension which plays an important part in social bonding - of dyads up to the largest groups.

A helpful starting point is to consider again the possible precursors of language - behaviours which lie behind (and maybe before) our ability to build up a complex symbolic code - and the role they played

in forming co-operative social groups, an important niche in the evolution of our species (see Chapter 1). In particular, we recall at this point Dunbar's notion of 'grooming-at-a-distance': in forming non-human and human primate social groups, interaction such as laughter, moving in time, and singing together are likely to be particularly effective - more so than just using language. He suggests that music-making (or an ancient practice perhaps resembling what we think of as music-making) is a ritual which provides the same emotional resonance as grooming, but on a larger scale:

> Imagine trying to co-ordinate the lives of 150 people a quarter of a million years ago out in the woodlands of Africa. Words alone are not enough. No one pays attention to carefully reasoned arguments. It is rousing speeches that get us going, that work us up to the fever pitch where we will take on the world at the drop of a hat, oblivious of the personal costs. Here, song and dance play an important part: they rouse the emotions and stimulate like nothing else the production of opiates to bring about states of elation and euphoria. (Dunbar, 1996: 146)

Here, we shall examine in more depth these forms of affiliative interaction. Even within a medium such as speech, the mere words we use do not comprise the whole interaction - rather, the *way* we speak, engage, and entrain with others forms the relational dimension of our interaction (including, of course, a degree of synchrony, and sometimes a great deal of laughing together). These are more intrinsically-oriented in goal, foregrounding the phatic, relational components rather than more transactional goals; they are the aspects of social interaction which are more simply and directly rewarding, reinforcing the relationship we have with our partner(s), and which can build inclusive, affiliative groups amongst any individuals who want to take part. As McNeill describes:

> Words and ideals matter and are always invoked; but keeping together in time arouses warm emotions of collective solidarity and erases personal frustrations as words, by themselves, cannot do. Large and complex human societies, in all probability, cannot long maintain themselves without such kinesthetic undergirding. Ideas and ideals are not enough. Feelings matter too, and feelings are inseparable from their gestural and muscular expression. (McNeill, 1995: 152)

In the following sections, I shall outline evidence which supports this hypothesis, addressing specifically those shared 'gestural and muscular expressions' of collective solidarity and experience. I turn first to work investigating synchrony, and then to the role of affect and prosody in generating a sense of shared experience, of intersubjectivity, amongst the participants.[79] These are important components of music-making; but - as we shall see in the second half of this chapter - they are not the whole story.

---

[79] The relationship between song style (e.g. vocal 'cohesion') and aspects of social structure (such as solidarity) had been explored a couple of decades earlier, in Lomax and colleagues' Cantometrics project (see Savage, 2018, for an overview).

Moving in time: synchrony and social bonding

Here, I shall explore findings about the effects of keeping in time, or moving in synchrony, with another person (or other people). The ease of engaging with one another, falling automatically into step, highlights how social interaction is our 'default mode' - not an unusual activity or special occasion (Hari et al., 2015). In Chapter 1, I considered the possible mechanisms behind this, and, in Chapter 2, how it appears in dialogue; here, I focus on another contemporary experience. William McNeill, a historian, has provided a range of examples of how synchrony is and has been central to *group* behaviour, from politics to religion. He starts with an account of his experience in the American military, specifically, the peculiar feeling of enjoyment in marching together - a strenuous exercise with no apparent purpose:

> Marching aimlessly about on the drill field, swaggering in conformity with prescribed military postures, conscious only of keeping in step so as to make the next move correctly and in time[,] somehow felt good. (McNeill, 1995: 2)

He describes how this so-called "muscular bonding" gave him a 'sense of pervasive well-being', 'a strange sense of personal enlargement', and even:

> ... a state of generalized emotional exaltation whose warmth was indubitable, without, however, having any definite external meaning or attachment. (McNeill, 1995: 2)

There is a whole myriad of empirical work which explores this phenomenon, unpicking its role in our social relationships, and our sense of group membership.

Being entrained

The experience of synchrony, moving or sounding together in time, is underpinned by the mechanism of *entrainment*, 'the interaction and consequent synchronisation of two or more rhythmic processes or oscillators', or the process by which two independent, self-sustaining components become coupled and "lock in" to a common periodicity (Clayton et al., 2005). Sometimes, this is 'asymmetric', as when a person is tapping their foot along to a recorded track; at other times, people move together, mutually adjusting their rate of movements to each other in real time (see Konvalinka et al., 2010; Clayton, 2012).

We seem to entrain unintentionally and automatically (Richardson et al., 2007). This is a key part of our ability to co-ordinate, whether in martial arts, hand-clapping, or joke-telling (Schmidt et al., 2011), or in ordinary conversation (see Chapter 2). Vesper et al. (2010) term it a 'coordination

smoother', as it makes our actions more predictable for the other (see also Vesper et al., 2011).[80] Similarly, Valdesolo et al. (2010) demonstrate that pairs who have rocked together perform better on a subsequent joint action 'labyrinth' task; they also feel closer or more connected to their partner. All of these effects lie within the category of the phatic or relational - they shape the *way* in which we relate to the other, maintaining the interaction channel, rather than involving or shaping our perception to anything outside the interaction itself.

Feeling close

Moving or drumming together with another person is a rewarding experience. It activates reward-processing regions of the brain and leads to prosocial or helping behaviour (e.g. picking up pencils dropped by accident; Kokal et al., 2011). It changes how we perceive the interaction, and our partner. Infants are more likely to look at characters that are moving in synchrony (than asynchrony; Fawcett & Tunçgenç, 2017). When we dance in time with someone, we look at them for longer (Woolhouse & Lai, 2014) and remember more about their appearance (Woolhouse et al., 2016). Even when we move in-phase (compared to anti-phase), we are more likely to remember both what we and our partner have said (rather than displaying our usual self-memory bias; Miles et al., 2010; see also Macrae et al., 2008).

We also like our partner more (see meta-analysis by Mogan et al., 2017). Just tapping in time with the experimenter (as compared to asynchronous tapping, or tapping alone) makes us more likely to rate them higher on explicit ratings of liking (Hove & Risen, 2009; this is even the case with virtual partners - Launay et al., 2014). Others who are just looking on notice a greater degree of rapport when we wave or walk in synchrony (Lakens & Stel, 2011). We are also more likely to trust others, contributing more in an economic game (Wiltermuth & Heath, 2009), and to help them (and more quickly - see Rabinowitch & Meltzoff, 2017; and meta-analysis by Rennung & Göritz, 2016). This effect arises from as early as 14 months (Cirelli et al., 2014a); these infants show directed prosociality to those who bounced in synchrony with them (and not to any other person; Cirelli et al., 2014b), but this helpfulness can transfer to others if we see that our partner is affiliated with someone else as well (Cirelli et al., 2016). The same effects of synchrony on liking can be seen in larger groups - but perhaps through 'distributed coordination', via a network of synchronised dyads, rather than the complete synchronisation of the whole group at once (see von Zimmermann et al., 2018).

Not only is synchrony a cue for our affiliation with others, but it also affects how we think about ourselves in relation to them. In particular, we feel like we are somehow more similar, or that our

---

[80] Mari Reiss Jones (2016) uses this predictive, anticipatory dimension of entrainment in her multi-levelled model of 'dynamic attending' to music - we predict where the next beat (and the next strong beat) will fall, and thus modulate our attention and expectations accordingly.

identities have got closer to (or even merged with) them. In a classic study, Botvinick and Cohen (1998) demonstrated that after brushing subjects' fingers in synchrony with the brushing of a rubber hand placed in front of them, they are then more likely to distort their perception of where their hand is, acting as if the rubber hand is actually their own. This effect does not emerge when the brushing is asynchronous. This sort of self-other merging seems to be experienced not only with rubber hands, but with other people with whom we have moved in synchrony. McNeill quotes A. R. Radcliffe-Brown, who describes a ritual conducted by Adaman islanders:

> "As the dancer loses himself in the dance, as he becomes absorbed in the unified community, he reaches a state of elation in which he feels himself filled with energy or force immediately beyond his ordinary state, and so finds himself able to perform prodigies of exertion." (Radcliffe-Brown, 1922; in McNeill, 1995: 8)

He links this with Judith Hanna's idea of "boundary loss", and provides his own description of this phenomenon as:

> … a blurring of self-awareness and the heightening of fellow-feeling with all who share in the dance. (McNeill, 1995: 8)

In experimental settings, tapping in synchrony results in higher ratings of closeness, or felt similarity to one's partner (in dyads of 8-year-olds; Rabinowitch & Knafo-Noam, 2015); dancing in synchrony results in a more interdependent self-construal (Reddish et al., 2013); and this perceived similarity is likely to be the mediator of greater helping (Valdesolo & DeSteno, 2011).[81] In fact, moving in synchrony can inhibit our own affective self-regulation, as we become absorbed in this group experience (Galbusera et al., 2019).[82]

Showing group membership

In this project, I am investigating not only interpersonal liking, but the formation of social groups (although these do, of course, often go hand in hand). Good et al. (2017) manipulated the degree of synchrony or asynchrony according to 'minimal' group boundaries (being assigned to either the red or blue team). Synchrony within these groups or across these team divides changed subjects' cognitive representation of the group accordingly - as either individual teams or a more inclusive group membership, respectively - with the corresponding in-group prosocial behaviour exhibited afterwards (in a public goods game). Therefore, synchrony with others affects perceived group

---

[81] These studies often use the 'Inclusion of the Other in the Self' scale, where subjects indicate their experience of self-other merging according to felt similarity with one picture from a series of more or less overlapping circles (Aron et al., 1992).

[82] This effect seems to come at the cost of self-monitoring and agency; the most successful joint action might actually recruit a mixture of self-other merging and distinction processes (see e.g. Huberth et al., 2019, and Fairhurst et al., preprint).

boundaries; this might even result in you feeling the protection afforded by this group membership (Fessler & Holbrook, 2014).

This matches our experience outside of the lab: religious rituals often involve moving in synchrony, with resulting feelings of in-group trust and prosociality (as shown in an economic game, where participants could contribute to a 'public investment'; Fischer et al., 2013). In what seems to be a real-life version of the study by Good et al., Lucas and colleagues (2011) document how separate groups of musicians, who together participate in the Afro-Brazilian Congado ritual, display their differential community identities not just through banners and uniforms but also in '*rhythmic difference*' (original emphasis). That is, they try to resist entraining with other groups (which becomes a harder task the closer they get), avoiding eye contact and maintaining their own sense of beat as far as they can.

In these ways, synchrony bonds us together with all who are moving and playing in time. However, our experience of music is often richer than this; there are other components which also contribute to our mutual affiliation.

Shared emotional experience: prosody and social bonding

Forms of emotional signalling are displayed across the animal kingdom; predator alarm calls, for example, can convey more information than we might realise on first hearing, through subtle acoustic variations (Seyfarth & Cheney, 2003).[83] In our way of interacting with infants, we also modulate how high or fast we sing, or the shape of our melodic line, according to their emotional state (and whether or not we want them to sleep); here, we are not just conveying a signal, but are in fact trying to change their own state of mind, or to share a feeling with them. In this section, I shall discuss how we share our emotional experiences through forms of resonance and mimicry, and how those mechanisms are in play when we listen to music (in a presentational field), and in promoting feelings of empathy after we have *interacted* musically (in a participatory sense).

In his account of how music forms community, William Benzon (2002) includes anecdotes of feeling moved by the emotional experience of another (even a stranger), conveyed through song (e.g. see his example of watching the tearful performance of Bette Midler on *The Tonight Show* - Benzon, 2002: 2-3). He describes how a process of coupling and resonance might underpin this, due to common experiences and expressions of musical emotion:

> ... for individuals sharing a common musical culture, there is a strong and systematic similarity
> between the tonal flow of music and its neurophysiological substrates that allows a tight coupling

---

[83] These contact calls are another manifestation of Dunbar's 'grooming-at-a-distance' (Dunbar, 1996).

between the brains of those individuals. While participating in the music those individuals constitute a community of sympathy. (Benzon, 2002: 44)

Freeman (2000) gives a similar account of music and social bonding: when we imitate one another, in making music together, we undergo a shared experience - bridging 'the solipsistic gulf' (Freeman, 2000: 420) in order to create a basis for joint understanding and trust. Of course, as we have seen, moving in synchrony is a potent form of shared experience; sharing an emotional experience can also be powerful in bonding people together.

Intersubjectivity

We have our own subjective experience of emotion within our private worlds, drawing on prior experience and expectations which only we have. However, these experiences can merge in what Rabinowitch et al. (2012a) describe as 'a complex entanglement between individual players entailing a fluid sharing of intentions, emotions and cognitive processes' (Rabinowitch et al., 2012a: 111), or *inter*subjectivity. At a neural level, this likely comes about through the activation of mirror neurons in our motor cortex (see Chapter 1), causing us to simulate others' actions and thus expressions of emotion (which are always tied to certain bodily manifestations).

We mimic others' movements automatically (Chartrand & Bargh, 1999); being mimicked causes us to focus more upon the other (i.e. developing a more interpersonal self-construal; Ashton-James et al., 2007); and also enhances our rapport with others (e.g. Lakin & Chartrand, 2003; van Baaren et al., 2004; Lakin et al., 2003). We also experience emotional contagion, whereby the emotional state of just one individual can "spread" to others in their vicinity (Spoor & Kelly, 2004; see Barsade, 2002, for demonstration of this process in the context of a group discussion). Some scholars have argued that the same processes are in play even when we just listen, passively, to music. Emotional contagion is one component of Patrick Juslin's model of musical emotion:

> ... an emotion is induced by a piece of music because the listener perceives the emotional expression of the music, and then 'mimics' this expression internally. (Juslin, 2013: 241)

There appears to be a 'common code' by which emotional expression is conveyed through speech and musical stimuli (Juslin & Laukka, 2003). This is not a purely acoustic phenomenon. Overy and Molnar-Szakacs (2009), in their model of Shared Affective Motion Experience, describe how musical perception is rooted in an embodied experience, whereby we simulate the actions which underlie the production of those sounds and emotional signals. Musical emotion is thus communicated via the same mirroring mechanisms which lie behind our capacity for emotional contagion, and, ultimately, for our Theory of Mind and empathy. As listeners, we have 'a shared *affective motion experience*' with the performer; even in just listening in isolation, the music gives 'a sense of the

presence of another person, their actions and their affective state' (Overy & Molnar-Szakacs, 2009: 494; see also Molnar-Szakacs & Overy, 2006):

> … the model suggests that when we hear music, we hear the presence (or agency) of another person, whose actions we can interpret, imitate, and predict. (Overy & Molnar-Szakacs, 2009: 495)

This again emphasises the experience of interacting directly with others as a more basic form of musical engagement - one which helps us to engage with real people, face-to-face.[84]

Music and empathy

To investigate how *participatory* music-making might result in a sense of shared emotional experience, Tal-Chen Rabinowitch (2012b) developed a number of musical games for groups of primary school-age children. This was a programme that was run over the course of a school year, and involved group interaction: children were given tasks - trying to be as rhythmically co-ordinated as possible, to improvise together, imitate each other's ideas, compose a piece with others according to a common theme, or to convey a particular emotion through music for others to guess. Children who took part in these games had higher scores in emotional empathy (tested using self-report measures, as well as non-verbal 'matched faces' and memory tasks) than those who took part in a control programme (of equivalent but non-musical games), or those who had no extra activities.

Empathy involves both an understanding and experience of another's emotional state; it depends upon the same mechanisms of emotional resonance and mimicry, but with additional cognitive appraisal and self-other distinction (see Chapter 1, and also distinction between emotional and cognitive empathy - Shamay-Tsoory et al., 2009). The capacity for this depends upon individual differences; some people are higher in dispositional empathy, whereas others struggle to understand or feel what others might be feeling. Musical interaction is likely to enhance this capacity through certain 'empathy-promoting' components such as motor resonance, imitation, and flexibility to act with others and, of course, to entrain and synchronise with them. The way that music conveys a sense of musical emotion in a simultaneously 'honest' and ambiguous way enables people to share (or just think that they are sharing) a common experience (see below on floating intentionality). This is all enacted within a context which is 'disinterested', i.e. involving predominantly a focus on the interaction rather than being occupied with any functional, external outcome - something similar to the idea that the interaction is intrinsically-oriented in goal (Rabinowitch et al., 2012b). These components of music-making rest on the capacity for shared intentionality (see work by Tomasello

---

[84] See also Trevarthen's account of musical intersubjectivity, discussed in Chapter 2 above.

discussed in Chapter 1), and contribute to this sense of intersubjectivity, and even merged subjectivity:

> ... where one subject may regard another participating subject almost as himself, to the point that one may experience another's sensations as one's own (Rabinowitch et al., 2012a, p. 111).

This experience lies behind our capacity for empathic understanding. We have already encountered this phenomenon in relation to synchrony - we feel self-other merging after moving or tapping together in time. However, Rabinowitch and colleagues (2012a) argue that this also comes about in combining ours and others' contributions to the musical texture - we merge these into a single unit:

> ... it is likely that when playing music together, as one gets absorbed in the music, it becomes harder and harder to discern the sources of the different sounds, to the point that one can no longer clearly tell whether the sounds being played were one's own or another's. (Rabinowitch et al., 2012a: 117)

In these ways, therefore, participants in music-making bond to (and perhaps merge agencies with) each other through a common experience, both in the very act of making music together and also in the emotional bond generated from the evocative context of collective musical prosody.

Summary: music and social group formation

By moving together in time and affect, with a shared aim to maintain a smooth, co-ordinated interaction, people signal to themselves and to others that they belong to that group, they identify with one another, and thus will come to each other's aid at a time of need. Music-making provides a context in which these phatic components come to the forefront (Cross, 2013); therefore, interacting musically is an efficacious means of bonding.

Group singing encouraged more co-operation in a prisoner's dilemma game and led to higher ratings of trust for other participants, compared to just listening to music, reading poetry or watching a film together (Anshel & Kipper, 1988); similarly, group singing resulted in higher co-operation than following art or competitive games (Good & Russo, 2016). Pearce et al. (2015) showed that participating in a choir led to faster increases in self-other closeness (as indexed by the Inclusion of the Other in the Self, or IOS, scale, see Chapter 5) and positive affect than other creative group activities, terming this the 'ice-breaker effect'. Moreover, social bonding can even occur in a much larger group context than possible with other activities - up to a choir of hundreds (Weinstein et al., 2016). These effects might implicate the same neurobiological mechanisms of social bonding, i.e. leading to higher levels of oxytocin compared to chatting together (Kreutz, 2014; see Chapter 1). Kirschner and Tomasello (2010) tested these effects of music-making with pairs of children; those who took part in a musical game, moving together with a pulse and singing a repetitive melody as led

by an experimenter, then helped their partner to pick up dropped marbles and worked together in a game more often than did those who took part in the spoken version of the group activity.

In these contexts, participants engage with one another, focussing more on the relational dimension of interaction than what the interaction might mean or achieve externally. In its emotional resonance, music appears to be an honest signal (with a biological basis in costly signalling present elsewhere in the animal kingdom); in the action of making music together, we engage our deep-seated capacity for shared intentionality, sharing our attention and intentions, 'to impart a sense of communicative intent' (Cross & Woodruff, 2009: 87). Aside from other more specific cultural associations (the implications of which we shall come to), music embodies no precise meaning: it is ambiguous, polysemic, with a 'semantic open-ness' which means that it can be interpreted simultaneously in many different ways.

> Engagement with music thus affords access to multiple and simultaneously available layers of meaning, allowing participants in a musical behaviour to interpret the significance of the music individually and independently while collectively affording to participants a sense that the music embodies an honest signal. (Cross & Woodruff, 2009: 87)

We encountered this 'floating intentionality' in the discussion in Chapter 2, where we saw how it afforded music-making more of an intrinsic focus in its overall goal. In other words, the main thing we achieve when we make music is the action of doing music together. This means that it is a lot easier to engage musically with any number of other people because there is no need to share external common ground or understanding through the course of the interaction. We might say it has a lower threshold for necessary common ground (or even that the common ground it does require is more internally-focussed). As a result, it minimises the potential for conflict, and instead develops our capacity for social flexibility (and even a 'metaphorizing' capacity to apply our knowledge across intellectual domains; see footnote 52, p. 87). This is very helpful in bonding people together in social groups, and perhaps especially so for those 'situations of social uncertainty' in various life transition rituals, between the mother-infant dyad, maintaining within-group cohesion, or managing positive inter-group relationships (Cross & Woodruff, 2009). In fact, this gives it potential evolutionary importance:

> We can suggest that "something like music" is likely to have played a role in enabling our ancestors to get on with each other: to form, maintain and re-form stable yet flexible groups or cultures. (Cross, 2012b: 25)

Therefore, music has great potential in forming social groups by interdependence (see Chapter 3). It is likely to be more effective than other media, such as speech, because it is intrinsically-oriented - foregrounding the relational goals. Thus, anyone can easily participate.

In my first experiment (see Chapter 6), I test this hypothesis by manipulating whether or not certain external goals are present and shared in the interaction group - all within a broadly musical context. I compare these music conditions with simple, non-musical group tasks. I hypothesise an effect of goals only within the non-musical context; group music-making enhances group bonding regardless of goal condition. This would indicate that group bonding comes about through the privileging of the phatic components of interaction (the synchrony and the shared emotional resonance of making music together), rather than a shared, external group objective.

However, music-making does not always result in group bonding. In their experimental work on music and empathy (see above), Rabinowitch et al. (2012a) note that personal conflict, competitiveness, unbalanced musical skills, or a general unwillingness to co-operate can all undermine the benefits of the musical games. In order to create the right conditions of musical engagement, the children need 'proper guidance and attention' (Rabinowitch et al., 2012a: 115). In fact, attributes of music more generally (even just as something to listen to) can evoke real social divisions, whether on a political scale or just between fans of different musical styles (Lonsdale & North, 2009).

In the second half of this chapter, I continue my focus on music-making as a form of interaction, but consider the conditions in which it might emphasise group divisions. To do this, I move from considering its more relational goals to those which might make it more extrinsic in focus; and simultaneously from considering group formation through interdependence to that resulting from categorisation.

Group boundaries: extrinsic-oriented interaction

Making music does not always bring people together. Pearce et al. (2016a) show that more 'competitive' singing, between university subgroups, resulted in higher in-group closeness; shifting to a more 'co-operative' context (singing together in a circle), reduced this discrepancy, and led to the students feeling closer to all who were involved. On village dancing, McNeill describes an analogous process:

> ... dance could and did become a way in which all sorts of new groups could define themselves, both by differentiation from within existing communities and by allowing marginalized persons or complete outsiders to coalesce into new, more or less coherent groups. All that was needed to achieve this was to exclude some persons while admitting others to the dance. (McNeill, 1995: 52)

Here, I shall explore how music-making demarcates category boundaries: how does our way of interacting musically form the basis for group divides, including some but excluding others? I propose

that a key element is the external common ground necessary for more extrinsically-oriented interaction.

Speech comprises a mixture of the relational and the transactional: while some gestures are used for the purpose of starting or continuing the interaction itself, others convey and share a new understanding or attitude about something else in the world. The latter (predominantly language, but also other referential media) is a large part of speech - thus we say speech is more extrinsic in focus. Because of this, we need to share more in common before we even start to speak - we need to share knowledge of that language, sometimes even of a specialised vocabulary; as we talk, we continue to make sure we share understanding. There is a higher threshold of common ground, excluding people who do not keep up, or who do not possess certain pre-requisites.

Music is slightly harder to pin down. A large part of it, as we discussed, is for the purpose of maintaining the interaction; it is more intrinsic in focus. The concept of floating intentionality alludes to the lack of external common ground which needs to be shared; music thus has a lower threshold, including more people in the activity. However, not everyone can always be involved in music-making. Sometimes, external components limit the scope of the interaction - participants might need access to a song-sheet, a score, or have knowledge of the performance conventions - and exclude those who do not share the necessary external common ground (in perhaps the same way as in speech).

Here, we shall focus on the basis of these divisions, and how they shape the formation and boundaries of social groups. Our focus moves beyond just face-to-face interdependence in bonding together co-acting individuals, to include our tendency for categorisation: assigning individuals to separate groups according to certain criteria. My argument is that the components of interaction which make it more extrinsically-oriented are the basis for such category divides. In other words, while some conditions of interaction include, and bring people together, other conditions exclude and enact the boundaries that come between certain groups (maybe splitting one group into subgroups, or separate factions). Music is multi-faceted, and occurs in as many ways as there are cultures in the world. I suggest that changing the musical goals - making music more or less intrinsic in focus - can act as a force for inclusion, or exclusion.

I begin by looking at how interaction more generally can form the basis for categorisation. We address how this occurs in language groups, and consider the role of common ground in distinguishing who is involved. This is relevant not just in our everyday speech, but more generally in our shared norms and expectations for how we relate; these considerations of course play a role in our musical practices. I outline a few case studies which illustrate this, before concluding this chapter with a preview of the experiments I have devised to test these hypotheses.

Constructing boundaries: languages and conventions

The work on Social Identity and Self-Categorisation (see Chapter 3), describes how people form categories (and categorise themselves in this social landscape), leading to intergroup differentiation, perhaps motivated by a need for positive self-esteem (in comparison with an out-group). Whereas theories of group interdependence focus on the interpersonal bonds between group members, social identity and categorisation approaches seem to privilege a discussion of how group *boundaries* are perceived and shape our behaviour. Here, I shall discuss how these boundaries have been thought about, criticised, and defined in relation to shared social systems, such as language or other conventions.[85]

Defining group boundaries: insights from anthropology

The idea of group boundaries has had a rather troubled history in the discipline of anthropology, given the potential for scholarship to validate the claims of one social group at the expense of another, perhaps supporting nationalist or other divisive forces. Moreover, we now live at a time where traditional societies or groups are being dispersed into multi-cultural cities, while millions of refugees are forced to move even across continents in search of a new homes and communities. However, as Bashkow (2004) argues, it is not an adequate response to simply dispense with the idea of group boundaries altogether: arguably, group and category divides become meaningful as they are crossed. Even so-called 'imaginary communities' exist only due to the presence of a boundary between those who belong and those who do not (Anderson, 2006). Bashkow draws on the earlier work of Boas (and his students) to devise a definition of boundaries which is still relevant (perhaps essential) in the contemporary world. These boundaries are (a) permeable, (b) pluralized, and (c) different depending on whether they are drawn by analysts or the people under investigation. Boundaries are plural: individuals can draw boundaries on the basis of ecology, social relationships, type of livelihood, or shared history or artistic styles, i.e. there are all manner of ways to categorise people. While we generally consider what is 'foreign' to lie outside of our cultural remit, an outside observer would be justified including this 'zone of the foreign' as part of our social group - even its perceived foreign-ness is part of a single way of viewing the world. Overall, Bashkow argues:

> ... we need to move beyond the notion that cultural boundaries are motivated by sharedness,
> whether it is conceived of in objective terms (shared language, ancestry, territory, social habits,
> or other traits) or in subjective terms (shared feelings of belonging). What we need to appreciate

---

[85] When we explore literature on the evolution of sociality, kinship emerges as an important factor. Theories of inclusive fitness (the idea that you co-operate with your relations because they also play a role in the continuation of your genetic family line) were among the first accounts of group prosociality. Animals rely on tags, or 'greenbeards' to recognise kin from non-kin for this purpose. In our human communities, family groups are often important subunits (see e.g. Roberts & Dunbar, 2011); here, however, I am focussing on group formation amongst both kin and non-kin, which can emerge through interdependence or categorisation according to number of possible dimensions.

is that boundaries can be productively defined in terms of a relationship of mutual comprehension. (Bashkow, 2004: 452)

This common understanding of the world, what is considered meaningful to us, is a helpful and productive basis for understanding the boundaries between groups of people. We can find a similar focus in Dunbar's work on friendship. He suggests that there are seven key dimensions which are likely to be shared amongst our affiliates: dialect, place of origin, education, interests, sense of humour and worldview.

Taken together, they constitute the set of beliefs and rituals that remind us who we are, where we come from and why we form a single community with a common set of values and convictions. (Dunbar, 2018: 45)

He links this explicitly to the feeling of 'Us-versus-Them' which arises with inter-category divides; sharing this common ground in how we view the world also helps us relate to certain others, and makes us more confident that they will understand us if we try to engage and interact with them:

... being a member of your community means that you know how they will see the world, how trustworthy they are likely to be, whether they will appreciate the same jokes as you do, and even how elliptical (and hence efficient) you can be in conversation. They will understand your metaphors, for example, without you having to laboriously explain them. There is also the implicit knowledge that someone in the extended network will be wagging their finger at them if they default on their obligations to you. (Dunbar, 2018: 45)

Children expect that individuals with a common group membership share common knowledge, specifically of cultural phenomena such as songs. However, they do not share more generic factual knowledge, nor do they necessarily like the same songs, by virtue of this group membership (Soley, 2019). In this way, we can make helpful distinctions between different social categories on the basis of shared knowledge or understanding, a shared system of meaning - in short, some sort of common ground in how we relate to each other, and to the world. This still allows for a 'release from proximity' (Gamble, 1998) - social groups do not depend on face-to-face interdependence, but rather can exist foremost in the minds and expectations of individuals as they approach and engage with even unfamiliar others.[86]

---

[86] Groups are thus not limited to e.g. geographic region; this also allows for the existence of Anderson's (2006) 'imagined communities'.

Defining group boundaries: social norms

One way in which we share meaning within our social group is in a common understanding of norms or conventions in the ways we are expected to act. Bicchieri (2006) call norms 'the grammar of society':

> ... like a collection of linguistic rules that are implicit in a language and define it, social norms are implicit in the operations of a society and make it what it is. Like a grammar, a system of norms specifies what is acceptable and what is not in a social group. (Bicchieri, 2006: ix)

From a young age, children have a sense of the right or wrong way of doing an (otherwise arbitrary) action in a game, and correct people (or characters) who do not conform (e.g. Rakoczy et al., 2008). These 'regularities in attitudes and behavior' are often an explicit part of what distinguishes one social group from another (Hogg & Reid, 2006: 7). Hogg and Reid draw on the Social Identity approach to describe how particular normative expectations (what they call 'shared cognitive representations') are the basis for our out-group and in-group stereotypes: the prototypical members of each are required to embody and communicate these norms to the greatest extent (Hogg & Reid, 2006: 10). This does not just include conventions in the way we dress or the types of food we eat (although these are salient in our everyday experience). But, they are also bound up in our way of interacting with one another:

> Norms are *shared* patterns of thought, feeling, and behavior, and in groups, what people do and say communicates information *about norms* and is itself configured *by norms* and by normative concerns (Hogg & Reid, 2006: 8, original emphasis)

The language we use, and the other gestures we deploy in speech, are all defined by these shared conventions. Lewis (1969) describes language in terms of general co-ordination problems; in these games, two participants share a goal, but, in order to achieve it, they need to make a move which is co-ordinated or aligned with the other's action. In order for participants to be successful, they rely on their own expectations about what the other will do, or what they think they know. In the same way, we use language in a way which is dependent upon our shared expectations and common knowledge (see also Lewis, 1975).

Defining group boundaries: language groups

Nettle and Dunbar suggest that language is a basis for group identification, in a way analogous to kinship. According to models of inclusive fitness (see footnote 85 above, p. 124), individuals are more likely to co-operate with others who exhibit 'tags' that demonstrate their genetic relatedness: our relatives do not just continue our family line, but are also more likely to reciprocate our help in the future. As groups expanded in population size and density, a need to co-operate with non-kin,

beyond those who we know or have encountered before, necessitated *linguistic* markers or tags. Dialect is a particularly hard-to-fake tag, a marker of group membership, by which we know that someone belongs to our community and we can trust them to reciprocate our help - it is, as Nettle puts it, 'an unfalsifiable indicator of group membership' (Nettle, 1999: 214). It is harder for potential free-riders to move from group to group: our way of speaking to one another develops over a long period of time engaging with and living in that community. In fact:

> Someone who speaks in the same way as you do, using similar words with the same accent, almost certainly grew up near you, and at least in the context of pre-industrial societies, is likely to be a relative. (Dunbar, 1996: 168)

Even if we cannot distinguish outsiders from insiders on the basis of appearance, just the sound of their voice is an honest signal. Nettle summarises this as the following strategy:

> ... sound like those you wish to cooperate with, and cooperate with those that sound like you (Nettle, 1999: 222)

In this way, language, dialect, or even just accent is a basis for categorisation through homophily: it is a way in which another looks and seems similar, so we treat them as part of our in-group (see also Cohen & Haun, 2013).[87] We show this in the way we use our speech: when Welsh speakers have a difficult interaction with an English person, they broaden their accent (or even speak in Welsh; see Bourhis & Giles, 1977); nationalist movements are often linked to linguistic purism; and we change the way we speak depending on who we are talking to, or trying to affiliate with.

> It is easy to see how accommodation to group members and discommodation from non-members, if sustained and not balanced by positive intergroup contact, would ultimately lead to the existence of separate languages. (Nettle & Dunbar, 1997: 94)

And, in fact:

> ... it seems that access to cooperation can depend on the use of the right linguistic markers. (Nettle & Dunbar, 1997: 95)

Nettle and Dunbar (1997) provide evidence for this using a computer simulation, demonstrating that the presence of social markers or tags, distinguishing group members from free-riders, aids the evolution of co-operation.

---

[87] This is likely to refer to the outer layers of our social network - those people who we are not personally familiar with, but who are similar enough to us that we consider them likely to have come from the same tribe or community (or, in more specific types of similarity, the same club or team - e.g. Dunbar's 'one-dimensional clubs': Dunbar, 2018: 45)

However, sharing a language has broader implications than just perceived similarity. As discussed above, an adequate basis for group boundaries (thus for processes of categorisation) goes beyond just seeming similar, or sharing certain attributes. Rather, it denotes shared meaning, the ability or potential to engage with others, who are "on the same page". In combining Social Identity Theory with a study of ethnolinguistics, Giles and Johnson (1981) discuss why shared language is an important part of ethnic identity. They allude to a similar source of common ground or meaning in their collective identity:

> The use of an ethnic speech style is...a reminder of a shared past, of a shared solidarity in the presence of a shared destiny in the future (Giles & Johnson, 1981: 205)

It still seems, though, that the language itself is still just an association: it acts as a cue for membership, a signal of one's identity, and an attribute that is valued, in a process of 'psycholinguistic distinctiveness' (Giles & Johnson, 1981: 217). However, one potential role for language that they mention is as 'the ideal medium for facilitating intragroup cohesion' (Giles & Johnson, 1981: 241). They also note that speech should be considered:

> ... as a dependent variable reflecting the interethnic situation and changes in it, and simultaneously as an independent variable defining, redefining and maintaining interethnic relations (Giles & Johnson, 1981: 243)

It is the role of interaction itself, the potential to engage with others and share a common experience with them, in showing but also *enacting* group boundaries which interests us here.

For Dunbar, the potential to interact using a common language means that we have a way to find out about others in the outer layers of our network, about their relationships, and also manage our own reputation in the eyes of others (compare Goffman's work on face) - 'language evolved to allow us to gossip' (Dunbar, 1996: 79). Aside from reputational concerns, having the potential to engage with others is also the potential to affiliate at all. We saw earlier how affiliation is bounded within the in-group: not sharing a common language is a potent basis for categorisation as "other". If we cannot talk to someone, we often cannot engage at all - at least in a context where speech is conventional, or needed. Malinowski puts this in stronger terms:

> The stranger who cannot speak the language is to all savage tribesmen a natural enemy. (Malinowski, 1923: 477)

However, when we compare different interaction media - namely, more musical or more linguistic ways of relating - we see that this potential for interaction varies, according to (a) our continuum from a more intrinsic to an extrinsic interaction focus, respectively, and, as a result, (b) a different threshold of necessary (external) common ground. This has implications for group formation. While

language has a higher threshold of necessary common ground (both before and during our interaction), a likely basis for categorisation, music, by virtue of its floating intentionality, has a much lower threshold of common ground, and so is an easier way of bring people together in relations of interdependence. Next, I shall explore in more depth the role of common ground in this process, forming the basis for group boundaries.

Constructing boundaries: common ground

While affiliative interaction can form interdependent groups, our interaction might depend on common ground (whether possession of necessary prerequisite knowledge or access to the grounding process). This distinguishes participants from outsiders and creates a boundary around the interacting group, thus forming the basis for categorisation. Here, I shall give some examples of this process. I consider conversation groups, in particular, which illustrate helpfully the role of common ground on a smaller scale. After this, I can apply the same principles but in a slightly different context, thinking about how forms of *musical* common ground might form and re-form group boundaries.

Conversation groups

Holding a conversation requires common ground at the outset, and then continuous checking, ensuring that all participants are up-to-date, before everyone can proceed together. We signal that we have understood by nodding, repeating phrases back to our partner(s), or asking for clarification if something does not make sense. If we are not able to contribute in this way, we do not share in the encounter, feel much less involved, and then understand less of what the others are talking about (Schober & Clark, 1989).

In speech, we allude to complex, external references and subtle meanings, through both language and gesture. In order to follow all of this, we need direct engagement and perhaps face-to-face contact with our partner(s). Kendon described how this occurs within 'F-formations', a type of focussed encounter or 'joint transactional space' (Kendon, 1990: 211); or, in other words:

> ... whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access (Kendon, 1990: 209)

Participants need to be able to see and hear each other well, continually shifting their posture to adapt to each other, and to signal changes in frame e.g. moving from a greeting to the main part of the conversation, changing topic, or ending the conversation (see Kendon, 1990). These spaces are generally closed to outsiders, who cannot take part without some ritualised initiation into the conversation (otherwise, they would be rudely interrupting). As a result of these requirements, only

four or five people can take part in a single conversation group, before it generally breaks apart with some individuals starting their own strand (Dunbar et al., 1995). Where there are more than five people present, conversations generally only occur between 'collaborative pairs', or with individual dominant speakers addressing the whole group (Fay et al., 2000). These sorts of boundaries are shown more explicitly in the context of games, where participants are more sharply divided from the rest of the world by specific rules of play (and by the expectation of 'fun' - Goffman, 1972: 17).

> ... games place a "frame" around a spate of immediate events, determining the type of "sense"
> that will be accorded everything within the frame. (Goffman, 1972: 20, quoting Bateson)

Goffman writes about 'rules of irrelevance': certain actions are given particular significance, but others are ignored as irrelevant.[88] Participants playing a game of Snap need only focus on the top card as it is turned, their task to notice a match, and then to react more quickly than their partner when the time comes; however, when we play Cheat, we attend closely to many more behavioural cues, in an effort to discern whether our competitors are deceiving us. These expectations provide a frame within which particular actions are interpreted: 'an engaging activity acts as a boundary around the participants, sealing them off from many potential worlds of meaning and action', and thus becomes a 'world-building activity' (Goffman, 1972: 24-25). All participants need to know this in order to be a part of the interacting, interdependent group: all others are excluded on the basis of this knowledge (which is thus a real basis for categorisation of who is in and who is not).

Groupthink

A perhaps stronger manifestation of this need for common ground (a requirement to be part of the group) is in people's preference for consensus or conformity. This was originally shown in Asch's (1956) line paradigm: participants feel the need to conform, to agree with others' judgements, even when this means making decisions which they would consider clearly incorrect in any other context. In the context of group discussions, this same pressure gives rise to the phenomenon of 'groupthink':

> ... a mode of thinking that people engage in when they are deeply involved in a cohesive in-
> group, when the members' strivings for unanimity override their motivation to realistically
> appraise alternative courses of action. (Janis, 1982: 9)

Irving Janis (1982) considers this a distinctly negative influence, and associates it with a number of 'historic fiascoes' such as the Bay of Pigs invasion and the escalation of the Vietnam War (Janis, 1982: 174). In these cases, group decision-making was characterised by an overestimation of the morality and invulnerability of the group, a collective ignorance and censorship of any different opinions, warnings or counter-evidence, and pressure on all members to conform to the official, apparent

---

[88] Compare the principles of relevance, discussed in Chapter 2.

majority view. Stasser and Titus (1985) found something similar, using their 'biased sampling model': members of a decision-making group are more likely to discuss information they all share, and do not bring up or pool information which specific individuals hold but others do not. This biases their eventual judgement according to the consensus; they come to a decision which is by no means the most balanced or informed.

These phenomena are contingent to some extent on the cultural context, or norms of the group. Postmes and colleagues (2001) found that such consensus-focussed norms could be counteracted by a group-level preference for deviance and critical thinking. Moreover, in a meta-analysis, Bond and Smith (1996) find that the degree of conformity shown by participants in Asch's classic line paradigm varied according to external cultural context. Higher levels of conformity might be expected during the 1950s (the era of "McCarthyism" in America), whereas they have since dropped; also, levels are higher in collectivist than in individualist cultures. In general, though, a need to conform to certain requirements, whether in explicit judgement or even in a more general norm of behaviour (as discussed earlier), is often a feature of social groups. Especially in the context of these discussion or decision-making groups, acting or speaking in such a way that thwarts others' expectations is more likely to lead to conflict. We should recall that the floating intentionality inherent in music-making is particularly useful in avoiding conflict (especially in contexts involving some social uncertainty).

Breaking frame

The interaction breakdown that occurs when participants lose their common ground is another helpful indication of the role of this shared framework in defining the boundaries of successful, interacting groups. When we are talking to one or two others and have a momentary misunderstanding, we can go through processes of clarification and repair which can set things right before we can carry on as normal (see Chapter 2). In his discussion of games, however, Goffman (1972) talks about what happens when our different understandings of reality (within and outside the game-play) start to conflict with one another; even if we are obliged to continue our game, it might be difficult to maintain the barrier between the internal common ground and rules, and our external norms and understandings. He describes how 'flooding out' can occur (Goffman, 1972: 50). Individuals cannot help but act against the current rules or expectations for the game (perhaps laughing uncontrollably). This creates a new, distracting focus for any others, who might all have to work to reconfigure the identity of the group, ceasing the game for the time being (Goffman, 1972: 53).

While these examples generally refer to the role of external common ground (i.e. in shared understanding of our current context), more relational goals can comprise frames - which might, in certain circumstances, be broken. We act to save face, protect our image and that of our partner: a

form of relational or internal common ground (see Chapter 2). Losing this, or breaking this sort of frame, also has implications again in the boundaries we draw between ourselves and others. We feel ashamed or embarrassment; our smooth interaction is threatened by individual 'flustering', and broken in 'paroxysms of laughter'. As noted earlier, our 'little social system' then 'collapses', and we 'hurriedly try to assume a new set of roles' (Goffman, 1956: 267).

This, therefore, also involves the unexpected process of attaining a *new* common ground. We might try to distance ourselves from the previous encounter altogether and engage with others, or perhaps attempt to retain our self-esteem in some way. Of course, our motivation for self-esteem and the social groups or categories which we identify with are explicitly linked in Social Identity Theory (Chapter 3). We shall see the impact of interaction success and embarrassment on our group identity (or lack thereof) in the experimental work I report in Chapter 6.

Overall, then, sharing some sort of common ground is a criterion for belonging to the group; it is a basis for the boundaries between positive, interacting groups, and a basis for categorisation. How are these boundaries formed in music-making? In line with our need for conformity, a matching of our expectations, Maher et al. (2013) found that listening to music which sounds unconventional or deviates from what we expect cause participants to evaluate an out-group (supporters of another sports team) negatively, or opt to help them less (in this case, deciding how much money to give to help the traveller community). Here we shall focus on how our shared norms, or languages, our common ground in how we perform, affect the way we interact and affiliate with others, and thus draw group boundaries in practice.

Constructing boundaries: musical identities

As I outlined earlier, music-making foregrounds the phatic dimension of interaction, and thus is efficacious for affiliative, non-conflictual interaction, bonding individuals into groups by forming relationships of interdependence. We know from everyday experience, however, that musical performance (both in its presentational and participatory guises) adopts culturally-defined, often very specific forms, with its own traditions, conventions or norms and extra-musical meanings. According to Fritz (2013), any given musical system would 'dock in' to a pool of universal attributes, but in specific combinations according to the unique cultural style or musical language.[89]

These conventions generally necessitate access to (and understanding of) certain notation, awareness of the external religious significance, certain permitted or unpermitted actions, and understanding of who can play what, when, and in what way.[90] The notion of what is or is not

---

[89] Although Fritz generally limits his discussion to features of the musical *sound* in his model, these observations are still useful when considering modes of music-*making* within particular cultural contexts.
[90] See Turino's discussion of various core or elaboration roles within participatory music - Turino, 2008.

"correct" within a particular context of music-making adds certain requirements to just a willingness to take part: all performers need to hold to some sort of shared framework. Benzon calls this a 'community of sympathy'; across these divisions, he claims, individuals are 'incapable of mediating secure coupling' (Benzon, 2002: 44).

Moreover, there is a continuum in how far musical systems can exclude in this way: while some are more extrinsically-oriented in goal focus, conforming to an almost linguistic grammar (as in the specified forms, subjects and counter-subjects in fugue-writing, or in the expectations we hold for the harmonic progressions in a Bach chorale - see work by Koelsch et al., 2002) - with a primary aim often to perform *something* correctly - other forms of music-making are much more intrinsically-focussed, where the main goal of the activity is to move and clap together. While the latter binds people together through shared affect and entrainment, the former provides more bases for exclusion.[91]

Overall, I argue here that these conventions - specifically, the common ground thus required to take part in music-making - form the basis for categorisation, creating boundaries around those groups of interacting individuals, and excluding others. In this section, I shall provide several specific examples of how these processes occur in this way, using a range of musical examples. Music-making does not just form affiliative groups (through interdependence), but also creates the awareness of boundaries, of separate social identities (by categorisation).

Songs around the campfire

My first example of these two process in play is provided by a case study of egalitarian hunter-gatherers, the Ju/'hoansi (!Kung Bushman) society of Southern Africa. In her exploration of the potential social (and subsistence) roles of fire, Wiessner (2014) compares the content of conversations during the day with those held in the evening (which is extended using firelight). While 'day talk' focusses on economic and social concerns (including social regulation, i.e. gossip), interaction during the long evenings (i.e. 'night talk') is more likely to include 'bursts of song', dancing, ritual ceremonies and 'enthralling stories'. These are forms of interaction which veer towards the intrinsic-oriented end of the spectrum, and are thus more primarily affiliative than the daytime exchanges:

> Stories told by firelight put listeners on the same emotional wavelength, elicited understanding, trust, and sympathy. (Wiessner, 2014: 14033)

---

[91] Knight and Lewis (2017) use similar ideas in their hypothesis concerning the emergence of language: 'words and grammar are means of navigating within a shared virtual world. Singing, dancing, and other forms of communal ritual are necessary to join people together in such ideal or imagined worlds.' They argue: 'language will not even begin to evolve unless ritual action has already begun to establish intensified levels of community-wide trust in association with a shared virtual domain.'

Wiessner even describes how, in their song, they included all who belonged to the community - from long-gone ancestors to those who had just travelled away: these evening activities 'expanded the virtual social universe':

> ... people went to sleep with absent kin filling their thoughts; not infrequently they left for visits shortly after. (Wiessner, 2014: 14030)

At the same time, these stories and rituals enacted and reinforced their cultural identity through telling of shared history. So, although there is less scope for conflict, with more inclusive interaction in a time of bonding and shared experience, taking part is still more or less restricted to those who belong to that wider community, those who recognise the legendary narratives and who live according to the same broad system of meaning and significance.

The performance of society

My next couple of examples also come from egalitarian hunter-gatherer communities. First is Seeger's (1987) account of music-making amongst the Suyá, an indigenous community in Brazil. Song is central in enacting their social relationships. While their ritual music-making can be highly participatory, a 'euphoria of community participation' taking place in the public plaza, there are specific conventions about who can participate in each ceremony, according to relationships of kinship and gender roles. Thus:

> ... performances established and re-established important relationships between groups and among individuals in very concrete ways (Seeger, 1987: 78)

Seeger describes how sometimes sung rituals even revealed or clarified otherwise ambiguous or secret relationships. He used the term 'structuration' to describe this performative function of music; it thus bonds people together, but in culturally-prescribed ways (Seeger, 1987: 139-140). This function is exemplified clearly in my next example: the hocketed polyphonic style of song, exhibited amongst the BaYaka and Mbendjele Pygmy people in Northern Congo. This performance involves close co-ordination; there is no hierarchy, rather everyone can be accommodated equally - but within certain conventions.

> To contribute appropriately one must not drown out one's neighbours, or sing the same melody as they do. Listening is as important as singing. If too many sing in unison, participants immediately and instinctively diverge by choosing alternative melodic modules to maintain the polyphony. (Lewis, 2014: 87-88)

This 'grammar' of interaction is something learnt implicitly:

Each participant's life-long musical apprenticeship has ensured that this musical deep structure is so effectively inculcated that each singer knows how variations can be executed and when to integrate them into the song. (Lewis, 2013: 60)

This does not just apply to specifically 'musical' behaviour, but structures all sorts of interactions amongst community members; the skills that they acquire in learning to perform together in this way are applicable (and indeed essential) to their everyday dealings and activities:

... the instinctive way that singers avoid unison has economic implications. In an egalitarian society, daily hunting and gathering activities are intuitively co-ordinated without the need for anyone to tell others what to do. If too many do the same thing, there may be nothing to eat, so being musically primed to do something different but complementary to others improves the chances that the camp will eat well. Similarly, knowing a sufficient range of melodic modules and when to insert them into the song structurally resembles the way environmental knowledge is employed to identify and extract resources from the forest efficiently. Musical participation in spirit plays is the main avenue through which BaYaka learn these unspoken grammars of daily interaction (Lewis, 2014: 88)

Again, this specific musical language is recognisable beyond that immediate community. While it restricts access to those who have grown up, immersed in that system, it creates at the same time a wider group membership, amongst other Pygmy people who share that specific 'socio-aesthetic'. Lewis recounts a comment made by one member of the BaYaka people on hearing the musical language of the Mbuti, who lived some distance away:

"They must be BaYaka since they sing just like us!" (Lewis, 2013: 53)

Interestingly, the spirit plays are often musical contexts in which these same processes of inclusion and exclusion operate at a small scale, even within a single community. Single-gender performances, often female displays of 'reverse dominance', emphasise their unity within that group, while asserting their difference - often hostility - to those men who cannot be involved. On one hand, during *Ngoku*, 'the united body of the singing women dances arm-in-arm up and down the central area of the camp'. Lewis describes how 'they speak as "Woman" to the men' and work to assert themselves, and embarrass their husbands in the process. On other occasions, such as *Sho* or *Ejengi*, men form a strong, united (and male-only) voice in return (Lewis, 2013: 58).

Team spirit

Football supporters provide another case of these processes in action. Fans sing their well-known team songs when they are all together for a sporting event; this also links them to a wider network of fans through the common knowledge of their team's chants - providing just the potential for

affiliative interaction with other supporters (now facilitated via the internet, for instance, in the exchange of comments in response to videos and recordings on Youtube). An interesting case study of this is the 'Carsi', a particularly avid fan network supporting the Beşiktaş Gymnastics Club in Turkey. The supporters are united not just through football, but also have a collective, class-based identity and a motivation to confront authority. In his exploration of this group, Kytö (2011) documents the strong solidarity, the 'feelings of togetherness' which come from the singing of well-known, simple songs, as loudly as possible (Kytö, 2011: 82). Often, a large number of people take part, creating an 'impenetrable wall of sound'. The sense of community, more than support for the team, might be an important motivation to belong to a group like this - especially for economic immigrants moving from rural towns into the larger cities:

> Supporting a team was a way to belong to a community and ease the feelings of loneliness and
> being a stranger. (Kytö, 2011: 81)

However, this sense of community depends not just on those interpersonal, affiliative relationships generated and reinforced through song, but also on a common awareness of the repertory, which is essential before anyone can participate and show that they belong. Again, through this shared language, they can show their support even from afar; by chanting the songs in a pub, it is as if they are part of the activity in the stadium. Music's 'floating intentionality' (as discussed above) makes it particularly suitable for this context - given the uncertainties involved in staking one's identity on a team that might or might not be victorious in the match. Supporters can even sing together in order to help their team, somehow:

> In such a world of latent ambiguity, singing and chanting are perfect tools to argue that influence
> can be created; for their veracity can never be objectively tested. (Armstrong & Young, 1999:
> 180)

Of course, football support is as much about excluding (even attacking) others who do not belong, as creating a sense of inclusive community within the group. Armstrong and Young characterise it almost as a form of ritualised warfare (something which has been clearly demonstrated in the stand-off between the England and New Zealand rugby teams, where the latter performed their famous 'haka'):[92]

> Football is thus about social differentiation. It is about us against them, and their defeat. It denies
> egalitarian ideals, and revels in our superiority, which it sings and dances on its way to success. It
> denies the Christian ethic that would turn the other cheek, and rather reemphasizes danger,

---

[92] See article published on the BBC website - https://www.bbc.co.uk/sport/rugby-union/50229807 - accessed 06/12/2019. This harks back to the 'coalition signalling' function of music, suggested twenty years ago by Hagen & Bryant (see Introduction to this thesis).

victory and domination in battles against some clearly identified "other". (Armstrong & Young, 1999: 179)

The musical standards and practices of rival football teams in Sydney form a case in point. On one hand, the Western Sydney ultras, the Red and Black Bloc (or RBB), perform their 'call to arms' as follows:

After chanting the words of the last verse "stand together and fight as one", fans embrace each other, making long lines as if they were a single organism, and jump to the left and to the right, screaming "terererere" (Knijnik, 2018: 951)

While, as a team, they emphasise their inclusion of all different ages and nationalities, this is still a bounded group. Their distinct identity rests upon both the prevalence of a working class, 'Westie accent', and the enhanced quality of their musical contributions. They contrast themselves with 'The Cove', the supporter group of the rival Sydney Football Club, who do not have 'the authenticity of real fans'. One RBB fan recounted:

"Mate, when I used to go to SFC games, 50 guys with acne on their face would turn up in the cove. The only drum they would bring was a wheelie bin that they would smash" (quoted in Knijnik, 2018: 952)

On the other hand, 'The Cove' has their corresponding traditions, with an emphasis on singing together, and crucially singing the right, or authorised, songs - what Collinson documents as the 'sing or f*** off' ideology (Collinson, 2009: 18). Although often the repertoire is comprised of easy, well-known melodies, many fans would argue that:

... an authentic fan would take the time to learn the songs, difficult or otherwise...simple songs were an unnecessary and unwanted concession to less committed supporters. (Collinson, 2009: 19)

In all of these examples, music-making is a powerful bonding medium, foregrounding those phatic components, creating a context of interaction which is more emotionally salient than referentially specific. However, this resource is used within bounds specified by the society, whether norms on *who* is bonded (and against who), or the conventions as to *how* exactly people interact within this space; this restricts participation to those who know these rules (even if only implicitly). These rules - the common socio-aesthetic or types of societal meanings - can be abstracted beyond the immediate interaction space to incorporate others who share this history, i.e. it is a means of categorisation, sometimes without any face-to-face contact at all.

Summary: inclusion and exclusion

Some components of music-making bind people together: anybody who moves together in time, and attends to each other's expressive gestures, is going to affiliate to one another (just as non-human primates do in grooming one another). However, other components of music-making impose some restrictions: in order to be involved, participants must share a common knowledge of the conventions, and they must understand the significance of certain gestures. This common ground divides those who can take part from others who are excluded: it forms the basis for categorisation.

Therefore, we can align our forms of interaction alongside our mechanisms of group formation. Forms of music-making which are easier to take part in, due to being more intrinsic-oriented in goal, are particularly efficacious in group formation by interdependence. However, music-making which requires more specific knowledge, both a priori and in interpreting events as they unfold (as in the continuous grounding required in speech), occurs only in bounded, interacting groups (or even dyads, in the case of some more intense conversations) - forming separate social groups by categorisation.

Victor Turner's (1969) distinction between 'communitas' and 'structure' is a similar conceptualisation of these different social relationships. While the latter forms cognitive classifications which provide a sense of order in the social environment, 'communitas' emerges where those boundaries are broken, in 'reclassifications', especially in those times or rituals of 'liminality':

> ... communitas emerges where social structure is not. (Turner, 1969: 371)

These ritual, liminal occasions often include members of out-groups, excluded or marginalised members of the community; they are a time when the common humanity of the group is emphasised, and those who are usually dominant are temporarily belittled. Of course, one needs those boundaries in order to break them - these two processes work together in tandem:

> ... each individual's life experience contains alternating exposure to structure and communitas, and to states and transitions. (Turner, 1969: 361)

While there is already some evidence of the role of music-making in social bonding (especially in its prominent role in rituals, or other times of 'social uncertainty'), there is much less on which components of music-making might form the basis for boundaries, or category divides. My experimental work includes two experiments which explicitly test the role of external musical goals (i.e. elements which create an extrinsic goal focus) on identification with a specific team membership, an attribute which divides people into separate groups or categories. I also run video analysis on the difference between music-making with and without this external goal. In the next chapter, I expand on how I have operationalised this overall theoretical framework into specific

experimental conditions to test my predictions. I then report on the results of this empirical work in Chapter 6.

Chapter 5 Methodology

Here, we reach the centre of this project. The purpose of this chapter is to build a bridge between the theoretical framework (outlined in Chapters 1 to 4) and my empirical research paradigm (reported in Chapter 6). First, I shall outline the main predictions arising from my theory, and, second, the methodology I developed to test these.

Overall, I propose a role of musical interaction in establishing group identity - both in bringing people together into inclusive, affiliative, interdependent groups; and in dividing people into distinct categories, according to features of or expectations about the musical context. These ideas are rooted in the literature on social psychology - namely, understandings of group formation as emerging from interdependence and theories of Social Identity and Categorisation. I also seek to characterise musical interaction. Previous literature highlights the value of holistic, behavioural analysis; I also draw on the conceptual distinction between more relational and transactional goals to communicate something within the encounter. In this chapter, I shall demonstrate how we can bring music into the laboratory - varying to what extent our musical interactions are more intrinsically- or extrinsically-oriented in goal, and comparing music with similar but non-musical activities - in order to test the antecedents of group formation, through both face-to-face social bonding and a sense of shared team identity.

I begin by outlining the rationale for my empirical work, explaining the aims of each of my experiments, and of my video analysis project. Then, I shall justify the various interaction conditions I use to manipulate certain variables within these experiments. I finish with explanations of the main outcome variables I used to measure the resulting social relationships, distinguishing the degree of interdependent group formation from a sense of team identification. This provides a research paradigm by which to convert some predictions emerging from my theory into clearly testable hypotheses.

Testing the theory: an empirical project

I devised a series of experiments to investigate the following research questions:

(1) Is musical interaction more effective than a non-musical mode of interaction at bonding participants together into an interdependent group?

(2) What is the effect of having, and sharing a goal within the group on this bonding process? Does sharing or not sharing the goal have a greater effect in the non-musical than the music

interaction condition, where differing goals and intentions can co-exist in a context of floating intentionality?

(3) What is the effect of external interaction goals on participants' sense of team identity? Does a team-related goal lead to preference for one's team members, and identification with that team more generally?

(4) Are there dissociable mechanisms leading to group bonding and team identification? Does bonding relate more to the intrinsically-oriented mode of interaction, and identification from success in distinctive, external team goals?

In the section below, I shall outline in more depth how each of my experiments operationalised and tested the relevant hypotheses. I stress that this is just the beginning of a much larger project: I aim to provide initial insights which support and further inform my theory, rather than conclusive answers at this stage. In general, this project will shed light on the role and efficacy of music in group identity, and the mechanisms which might lie behind this. This work also presents a new, music-oriented research paradigm: I think of this as an empirical resource - a way to manipulate social interaction in order to explore further research questions in group formation and social engagement.

Testing group formation - experiment 1

My first experiment investigated the antecedents to group formation or group bonding - establishing relationships of interdependence amongst individuals, through forms of face-to-face interaction. I varied two parameters of the interaction. The first I call the interaction *mode* - whether the group activity is *musical* or *non-musical* in its character and attributes. The second is the presence of an external goal, shared or non-shared within the group - which I introduced in both the musical and the non-musical activities. I then used a variety of outcome measures to test the extent to which the interacting individuals have become (or feel that they are) a "group".

Interaction mode

I noted before how music 'foregrounds the phatic'; it is intrinsically-oriented, emphasising more relational than transactional interaction goals. As discussed in Chapter 4, various forms of affiliative interaction (perhaps singing or moving in synchrony) generate higher levels of liking, co-operation, and the sense of belonging to the same group. I therefore expected that a more *musical* mode of interaction would promote more group bonding and cohesion, facilitating affiliative relationships among a number of individuals, than would a *non-musical* interaction task. I tested this hypothesis by comparing very simple musical and non-musical activities. The musical task emphasised direct interpersonal response and joint action; the non-musical interaction comprised a visual collage task, where participants worked together to move shapes into a picture. Aside from this, I aimed to make

these group activities as similar as possible - testing just the effect of the *musical* mode of engagement, as far as possible.

Interaction goals

The original literature on social groups emphasises interdependence *towards a shared goal* (Chapter 3). However, it is not always the case that interacting individuals share (or indeed have) an external goal. In foregrounding the phatic, music can accommodate many simultaneous, separate understandings or goals in a context of 'floating intentionality'; it does not even need an external goal - music-making can be sustained merely as participants share the intention to keep the interaction going; they do not need any mutual understanding greater than this.

Some of the interactive activities in this experiment included an external focus (they were more extrinsically-oriented): participants were given a goal to depict something through their interaction (either visually, or through music). The visual conditions always included a goal - it was difficult to operationalise a non-musical, no-goal task. However, some participants in the musical interaction condition had no additional goal (while others did have a goal introduced): music is an interaction medium which could be purely relational. I tested whether adding a goal affected the resulting social relationships.

Additionally, I did not just compare goal with no-goal activities, but also investigated the effect of whether or not those goals were *shared* within the interacting group (where, importantly, all participants expected to be working together). This is somewhat similar to the task used in Curioni et al. (2019), where participants had to synchronise in drawing different shapes. I hypothesise elsewhere (see experiments 2 and 3) that sharing an external focus, such as an interaction goal, might enact processes of *categorisation* more than interdependent group bonding. Here, it is important to explore whether sharing an interaction goal (or not) affects the process by which individuals build relationships, face-to-face, forming a sense of group identity from the bottom up. I also test whether the effect of goal-sharing occurs alongside, is itself affected by, or is just dwarfed by the effect of interaction *mode*. It makes sense, for example, that musical interactions are generally more affiliative, by virtue of their overriding characteristics - so sharing or not sharing some external goal might not have an effect (music has 'floating intentionality'). Not sharing a goal in a visual task is likely to be more disruptive, and more difficult for participants to work alongside each other (even perhaps resulting in a degree of frustration at others' seemingly inexplicable actions).

Summary: conditions and hypotheses

In Table 1, I summarise the main (broadly 2x2) design, the five different conditions, and the predicted degree of group formation. I expected music to be more effective than the non-musical task, perhaps

even more so when there is no external goal (i.e. when music is the most intrinsically-oriented; this is exploratory, however - even musical interaction might benefit from at least some shared focus). I then expected the goal conditions to have an effect on the level of bonding achieved through the visual task.

Table 1 *Experiment 1 design and hypothesised degree of group formation*

	Shared goal	Non-shared goal	No goal
Musical task	High		Highest?
Non-musical task	Medium	Low	

Note: there was no non-musical, no-goal task in this experiment (the shaded box).

Creating team divisions - experiment 2

Although musical interaction affords some degree of 'floating intentionality', some forms of music-making incorporate some reference, meaning, or convention which can be (and sometimes need to be) shared as part of the interaction. For example, one might require general knowledge of the structure of a performance, awareness of particular stylistic conventions, or access to a score in reading through a string quartet. These meanings or goals, although intimately linked to the process of interaction itself, are external: they give the interaction a focus outside the immediate relationship; they make it more extrinsically-oriented. Similarly, we categorise others on the basis of common, external attributes, perhaps signalling team identity, ethnic background, kinship or language group - this shapes who we interact with. Perhaps, musical interaction *contributes* to the formation of these categories: common ground, or interaction goals, might highlight shared traits, maybe even restricting the interaction to those who hold the requisite knowledge, awareness or understanding.

In my second experiment, I introduced the idea that participants belong to two different, supposedly competing teams - comprising a broad membership from across their cohort. I distinguished the teams from the interaction *groups* - these face-to-face gatherings involved individuals across both teams, requiring them to interact both with those who shared their team identity and with others who belonged to the opposing team. Again, I varied attributes of this group interaction: whether participants had an external interaction goal or not (a goal directly corresponding to their team identity). I then measured both how bonded they felt to their *group* and, whether they felt more connected to individuals who shared their *team* membership (more than to others who did not; both within and outside the experimental context). This provided a way of testing whether adding an interaction goal, a shared focus specifically among the *team* members, reinforces bonding on the basis of their team membership (i.e. bonding only to each other, not extended to other-team members, even when they participated within the same interaction group).

Goals and teams

In this experiment, all participants interact with each other in a *musical* mode, varied according to the presence or absence of a team-related goal. In one condition, participants had no extra goal for their group interaction; in the other, participants were instructed to work together to depict something through their musical interaction.

At the outset of the experiment, all were randomly assigned to either the "elephant" or the "bird" team. Where they were given an interaction goal, this was to depict a herd of elephants (if in the elephant team), or a flock of birds (for the "birds"). Because the groups in which they interacted comprised a mixture of team members, the goals were correspondingly not shared within the groups, but only shared amongst those in the group who also belonged to the same team.

I tested the effect of interaction goals both on general group bonding, and on specific team-related affiliation. These goals formed external properties of the interaction, a type of common ground which is shared amongst members of the same team. Possessing a goal might emphasise the division between these two teams - a division which thus becomes more apparent as they interact. We see something similar (albeit on a much more complex, larger scale) in cases where inter-group boundaries are marked by difference in language use (Giles & Johnson, 1981), in music listening preferences (Lonsdale & North, 2009), and in different conventions of music-making (see above on the BaYaka musical 'grammar'). I aimed to model this process, using musical depiction goals as a very simple means of varying the musical common ground or conventions, according to team identity.

Goals and groups

This experiment also provided another opportunity to investigate the effect of musical goals (an external focus, thus a more extrinsically-oriented interaction) on more general group bonding. I tested whether having no additional goal (i.e. interaction which privileged the phatic) enhanced mixed-team group bonding compared to interaction tasks with an added outside focus.

In this case, however, adding a goal did not just change the mode of interaction; the goals aligned with team identity, and so possibly drew attention to differences in team membership within the interacting groups. Therefore, goals would not just change the mode of interaction (reducing the more *relational* focus) and thus diminish group bonding (as expected in experiment 1), but they might also emphasise the differences in team identity, making whole-group bonding even less likely.

Summary: conditions and hypotheses

I present my hypotheses for this experiment - relating both to group formation, and to team identification - in Table 2. There are just two conditions - both involving musical interaction between individuals who belong to two different teams.

Table 2 *Experiment 2 design and hypotheses*

Conditions > v Outcome	No-goal activity	Non-shared goals (according to team membership)
Group bonding (across teams)	High	Low
Team-related affiliation	Low	High

Social relations and team identities - experiment 3

The purpose of my third experiment was to make a clearer distinction between interpersonal bonding and a sense of team identity. In experiment 2, I measured affiliation to others on the basis of team membership; mostly, these others were also involved in the interaction - this made it harder to discern whether preference for one's own team occurred as a result of personal identification, or through face-to-face interdependence. In experiment 3, I aimed to distinguish as far as possible the affiliation between interaction pairs (through interdependence), and the sense of team membership that goes beyond the immediate interaction context (through self-categorisation).

I used the same interaction conditions as in experiment 2, with the same goal/no-goal manipulation. This time, participants interacted in pairs; both partners shared the same team identity. Then, I measured direct interpersonal affiliation towards one's interaction partner; and their sense of team identification, implicitly, using a memory test.

Interaction goals and interpersonal relations

By comparing the same depiction goal tasks (some elephant-, some bird-related) to the no-goal musical interaction, I could test again whether interaction *mode* shapes the degree of interpersonal affiliation between the interaction partners. As above, if music is efficacious at group bonding because it has a more intrinsic goal focus, privileging the phatic over external goals, then I predicted that a no-goal interaction task would lead to higher affiliation between interaction partners.

Interaction goals and team identity

With these same conditions, I could test whether having an external goal reinforces a sense of team identity - and one that generalises beyond the interaction context. This time, interaction pairs shared

either a bird- or elephant-related goal, corresponding to their team membership. I investigated whether adding this external focus emphasised team identification, in comparison to a no-goal interaction task. Moreover, following a pilot, and the results of experiment 2 (see Chapter 6), I predicted that the elephant-related goal would be more difficult to achieve (than the bird goal); I expected that a less successful goal would be less effective at reinforcing the related team identity (given that our social identities are tied to our self-esteem - see Chapter 3).

Summary: conditions and hypotheses

This experiment again measured the antecedents to interpersonal bonding (through interdependence), and the impact of an external focus on team (or category) identification. This was largely exploratory, but some initial hypotheses are presented in Table 3 below.

Table 3 *Experiment 3 design and hypotheses*

Conditions > v Outcome	Interaction goal		No-goal musical task
	Bird goal	**Elephant goal**	
Interpersonal bonding	Low		High
Team identity	High	Low	Low

Characterising musical interaction - video analysis

If the goal focus of social interaction (i.e. whether it is more intrinsic - as in no-goal conditions - or more extrinsic in orientation - with goals added) is a crucial factor shaping both the degree of affiliation between interacting individuals, and a resulting sense of team identity, then it is important to investigate the behavioural properties of these interaction conditions. This was the purpose of this final part of the present empirical project. I filmed case studies, where pairs engaged in one interaction task for five minutes. Three of these pairs were selected - one for each of the no-goal, bird- or elephant-related goal conditions. These conditions varied only in their goal focus - my interaction tasks were exactly the same, but for the presence or absence of an external depiction task. These were analysed in detail, focussing on inter-individual synchrony and eye gaze patterns.

Synchrony has been linked to co-ordination, affiliation and merged subjectivity. Additionally, conversation partners look at one another to ensure successful grounding, making sure things have been understood at regular points in the exchange. If our interaction conditions exhibit different patterns of such checking behaviour, or a different relationship between the interaction "content" and participants' gaze behaviour, then this might signal differences in the main focus of the interaction - perhaps a different type of common ground between our conditions.

Overall, a close analysis of these controlled forms of musical interaction offers insights into the nature of social interaction - specifically between forms which are more focussed on external goals and others which privilege the relational, phatic dimensions. If this distinction forms the basis for a continuum between music and speech, then the first behavioural analysis of this sort (found in Chapter 6) has quite extensive significance.

I shall now move on to discuss my experimental set-up. Specifically, I describe the activities I devised - the conditions under which my participants interacted with one another, and thus the way in which I turned my theoretical understanding of music-making into an operational reality.

Lab-based musicking

I have used Christopher Small's term, as it emphasises how "music" should always be a verb - it is something which is performed, in many different contexts, guided by the various conventions and expectations of a particular social group or community. In a laboratory context, we face a trade-off between trying to model this complex reality - in other words, maximising ecological validity - and maintaining experimental control (as we try to manipulate only certain parameters of the encounter; see D'Ausilio et al., 2015, for an overview of this issue).

Many similar studies, investigating the effects of music on social bonding, have taken place in an everyday musical setting, often a choir rehearsal; researchers measure certain indicators of social bonding, whether in self-reported 'closeness', level of oxytocin, or pain threshold, before and after a given amount of rehearsal time. In Kreutz (2014), this involved warm-up activities, time of song, and rehearsal points (and is compared with time spent 'chatting' with a few individuals about positive life events); Weinstein et al. (2016) also made use of an existing scheme where small local choirs gather once a year to form a 'megachoir' in order to investigate the effectiveness of music-making in bonding smaller or much larger groups. Others have created their own activity groups, recruiting professional singing coaches (alongside craft and/or creative writing teachers) to form a choir especially for research purposes (e.g. Schellenberg et al., 2015; Pearce et al., 2015; Pearce et al., 2016b). Even here, though, the activities are still very broadly defined. Pearce et al. (2016a) constructed a semi-naturalistic study in the lab, creating contexts of co-operative and competitive singing (which already occurred in the activities of the university 'cliques' or fraternities who were recruited), to measure how music might enhance or undermine these pre-existing boundaries. These participants were already part of friendship groups before the study, which perhaps helped facilitate the activity. Other studies involved laboratory singing (such as Wiltermuth & Heath, 2009; or Anshel & Kippur, 1988), but one can imagine that adults are likely to feel uncomfortable, especially if they

have not had musical training. Children are usually less inhibited - as shown in Kirschner and Tomasello's work (2010).

Here, I needed an accessible musical task for adults, one which ideally models the enjoyment of simple music-making, while doing justice to the richness of music as it occurs "in the wild" - even within a somewhat artificial setting, where participants are asked to interact with strangers. I devised a musical activity involving aspects of music-making which occur across cultures, recruiting perhaps "natural", even universal, capacities; the task did not need any specialist training, or even a great deal of self-confidence, to participate. At the same time, my experiments required some control, manipulating the task to investigate exactly which components contribute to specific social relationships. In particular, I devised a way to manipulate the intrinsic or extrinsic goal orientation of the interaction - adding or taking away an external goal to an otherwise highly flexible task, foregrounding the phatic as far as possible. In other words, I tried to develop musical activities which were more or less musical, or speech-like - varying just their position along our continuum of interaction goals (Chapter 2). In addition, I developed a task which could be considered a *non-musical* version of the same activity. For this, I needed to decide which components comprise more "musical" aspects, and selectively remove them from the task. In this section, I shall detail exactly how I created these conditions.

A basic musical interaction

Music-making is a culturally-situated activity, rooted in universal capacities to relate and tune to others, but shaped by specific learned customs or conventions - whether ritual expectations, a harmonic language, or even a social worldview enacted through musical performance. If the universal attributes make it an accessible way to bond with others, it is then the need for knowledge of the particular language which divides or categorises those who know it and so belong, from those who do not. In order to test this theory empirically, I needed a simple baseline musical condition, modelling and recruiting those universal components. Then, I could investigate the effect of adding external goals - making the interaction more extrinsic in focus, with a form of common ground that is shared amongst only some participants - on limiting or structuring those social relationships. Here, I shall enumerate the characteristics of our most *basic* musical task - the baseline condition.

A social encounter: participatory and improvised

In this project, I focus on social interaction and relationships - thus on 'participatory' rather than 'presentational' forms of music (Turino, 2008; see Introduction). The main aim of the musical task was relational - participants were to interact, not focussing just on their individual contributions, or on the musical instrument they used. I instructed participants to work together (rather than following any leader), as a group (or a pair), and to co-ordinate with each other in making music. As is

often the case with forms of participatory music, there was no start or end of a "piece", no specific item to recreate, but instead just the opportunity to interact and to improvise together, responding to one another moment-by-moment. There were no other guidelines, rules or scripts to follow (compare the external goal conditions below).[93] They were not even encouraged to synchronise, or keep time; I expected that this interactive, affiliative context would provide a space where this would emerge spontaneously (as it does in everyday spoken conversation).

Basic, accessible musicking: a universal capacity

The musical activity was one in which all participants could engage equally well, regardless of their musical background. It was a novel task where no-one would have any expertise; participants all approached it from the same level, despite potential differences in exposure, musical tradition, training or preferences.

Participants were given a simple musical instrument to interact with (a kazoo); this acted as a "mediator" of some kind (as we shall explore below), but it only required the player to vocalise, or hum, into one end for it to sound (no specialised hand movements, or "technique" was required). Thus, it was simple and easy to use, not offering any particular interest which would distract participants away from their primary task to interact and engage with their partner.

This instrument changes the sound quality, but otherwise mirrors the pitch and contour of the participants' vocal or sung input. Our ability to sing is a deep-seated ability, perhaps our most basic (and most ancient) musical instrument (other than body percussion). Ian Morley suggests that our vocal tract is our 'principal tonal sound-producing apparatus' - it is a:

> *... biological* instrument, possessed by all of us, which constitutes the principal tonal sound-producer in many musical traditions. (Morley, 2013: 131)

There is evidence that a lower larynx position (creating more resonance space, greater control of tongue movement and lower formant frequencies), and the increase in innervation necessary for more control of pitch and intensity, were already emerging from the time of *Homo ergaster* (at roughly the same time as our shift to upright bipedal posture). The capacities for extended utterance duration came after this, probably around 600 000 years ago, in *Homo heidelbergensis.* We know too, of course, that prosody and contour is important from very early in ontogeny, in the communicative and affective potential of infant-directed speech (see synthesis of all of these strands

---

[93] This is in keeping with the work exploring jazz (and other) improvisation (see, for example, MacDonald et al., 2012, and Wilson & MacDonald, 2016). My experimental conditions perhaps explore further the roots of these tasks: I have devised an activity which is likely to differ from anything my participants have encountered before (compare Canonne & Aucouturier, 2015, who show the effect of familiarity on implicit models which form a basis for collective, free improvisation).

of evidence in Morley, 2013). Our ability to sing seems to be the root of our music-making capacities; I recruited this as a primary part of my experimental music-making task.

Instrumental affordances: extension and distance

If vocalising is the root of human musical capacity, to which instruments 'constitute an accessory' (Morley, 2013: 131), why use a musical instrument at all? First, it is important to note that extending musical capacities using tools is still a very natural thing to do. The goal-directed actions involved in manual tool-making, and our ability to learn these actions through imitation, have a large functional (and probably neural, or structural) overlap with our capacity for language acquisition and intentional spoken communication (Stout & Chaminade, 2012). Even more so in music-making: this does not just involve sung gestures or melodies, but a holistic, embodied experience in which we entrain in movement, and align our affective experience with our interaction partners. All sorts of gestures are used for various purposes in music-making (see overview in Jensenius et al., 2010); some have advocated an embodied understanding of music in general (even when music is defined in presentational, often purely sonic terms). Musical tools then act as 'extensions of the human body' in the production (or even search and retrieval) of these musical works (see Leman, 2008: 137). Considering the participatory field of music-making, if our vocal mode is so strongly tied to motor, gestural affordances (perhaps even with a common time-keeping mechanism; see Morley, 2013), then it makes sense that we can interact, musically, through our gestures, even via instruments or tools - without this mediation substantially changing the nature of the musical activity.

Second, the use of musical instruments offers practical advantages in the context of an experimental set-up. Participants were restricted to interaction via this channel, precluding the use of speech and emphasising the relational dimensions of prosody and synchrony, rather than any external references or goals - the primary orientation of this basic musical task (see above). This instrument could be played hands-free: although participants were instructed to interact through or using the instrument, they could move their hands freely (as is normal during speech). Perhaps most importantly, the use of an instrument as a mediator provided some distance between the individual and their contribution - this facilitated participation by reducing the amount of embarrassment they might feel at having to interact or perform with a stranger, in a new setting (and sometimes with a video camera pointing at them). Burrows (1987) describes how musical instruments can act as a mask - 'depersonalizing' the actor through 'timbral standardization' across performers:

> ... an instrument replaces the performer's own sonic face, the voice, with a proved impersonal sound used only for making music. (Burrows, 1987: 122)

Through this 'cloaking' effect (Turino, 2009: 100), my participants could be less self-conscious, as their voices became more uniform (or, you might say, as silly as each other).[94] Again, this focussed the task on the relational, the task to interact with one another - as it is these dimensions which make our task predominantly musical (in a participatory sense), rather than anything else.

Changing the music: manipulating interaction goals

If the defining feature of music is a form of interaction which foregrounds the phatic, establishing and maintaining a context which is primarily for relational purposes, then we have described above a quintessential, musical condition. We have stripped away any cultural specificities, conventions, or references, only requiring the ability to vocalise - a universal capacity which likely forms the root of our human musical expression. However, the key interaction manipulation within this project was the goal orientation of the task - whether it was primarily intrinsic, or more extrinsic in focus. Here, I shall outline how I introduced an external goal focus, "changing the music" condition so that it became slightly less relational in focus, instead directing some attention towards an external interaction goal.

A depiction goal

In speech, we aim to share some sort of common ground, knowledge or awareness about something in the world. This often characterises musical behaviours too, although generally to a lesser extent: certain melodies might be iconic, or even symbolic; musical structures can be pre-defined according to convention, maybe a ritual purpose; musicians often need access to and understanding of some sort of score to guide their interaction or performance. In other words, we all need to read off the same song-sheet, as the saying goes. To model this external common ground in a controlled, systematic and fairly simple way, within the same musical interaction task described above, I decided to introduce the goal to *depict* something.

Participants were instructed either to depict a flock of birds or a herd of elephants - they were told to make music which was inspired by this image (a picture was also shown on the instruction card). This required them to hold in mind this external reference, and anticipate that their partner(s) also shared this aim (probably with a shared selection of sounds or behaviours commonly attributed to these animals, drawing on their common knowledge or experience of them in the world). This becomes an interaction goal, and a form of common ground - something external they share and aim for - on top of just the goal to interact.

---

[94] This 'cloaking' function emerges in Turino's participatory 'field' of music-making, which is characterised by a 'wall of sound' - in contrast to the expectation of 'clarity' in a presentational context (I assumed the latter would be the default understanding of music-making amongst my participants; see discussion in Turino, 2009).

Goals and groups

The purpose of this project is to ascertain how this change in interaction focus or common ground might enact or structure our social relationships differently. As I outlined above, the purpose of my first experiment was to test whether having - and sharing - these interaction goals affected the degree of group bonding individuals experienced following the activity. Participants always expected that others in their group would have the same goal;[95] however, sometimes the goals were not in fact shared within the group - would they notice this, and might it affect group formation? Moreover, both experiments 1 and 3 provided the opportunity to test whether just having a shared goal or not affects the degree of affiliation - does this goal manipulation change the interaction mode substantially enough to have a real effect on how much they relate to, or like, each other as a result?

There is also some reason to think that having different goals or common ground might create group boundaries by emphasising a process of *categorisation*. While team identity is usually signified in experiments by some perceptual attribute (such as team colour etc.), having a depiction goal that corresponds to, or is semantically related to, one's team identity (either the elephant or the bird team) aligns with our real-life social experience. Language is one signifier of social identity: it is not just a perceptual difference, but one which actually shapes or limits our potential for interaction with others - we have to share common ground to associate with them.

The role of music-making practices in these processes of group formation (and division) has been demonstrated in some case studies (see Giles et al., 2009, for overview and examples). Musical listening preferences (or personal playlists) provide an identifying "badge" for young people, aligning themselves with certain social categories in the eyes of their peers (North & Hargreaves, 1999). Again, *making* music provides even more potent examples of this - signalling national identity, membership of some ethnic class or group - especially so in the experience of migration or diaspora (e.g. Baily, 2005; Baily & Collyer, 2006). Crucially, people show that they belong through being able to participate in the music; necessarily, those who do not share this common ground are excluded (knowledge of and participation in a team's football chants is a clear example of this - see e.g. Collinson, 2009). In this much-simplified empirical model, introducing an external form of common ground - here a musical depiction goal - might provide a way of emphasising one's social identity, or one's sense of belonging to a particular team.

A non-musical task: creating a matched control condition

Although the difference in common ground or goal orientation of musical interaction forms the main part of this project, it was necessary to investigate interaction *mode*, more broadly speaking. In my

---

[95] They were led to believe that the task was investigating "group memory" (see method in Chapter 6).

first experiment, I compared a musical with a non-musical task, measuring how far music is more efficacious for processes of interdependence. In this section, I discuss how I designed this task; in other words, how I defined what made a simple, interaction task more or less "musical".

The interaction engine

The fundamentals behind all sorts of interaction - regardless of mode, goal focus, or anything else - are encapsulated in Levinson's idea of our 'interaction engine':

> … a set of cognitive abilities and behavioral dispositions that synergistically work together to endow human face-to-face interaction with certain special qualities. (Levinson, 2006: 44)

The idea is that human social interaction has a number of universal attributes - with just superficial differences across cultures (or, here, across modes of interaction). These attributes include some awareness of common ground, communicative intentions (or Theory of Mind); a general structure of turn-taking or reciprocity (although the importance of this in *musical* communication is an open question); close co-ordination in timing; in short, a participatory, multi-modal joint activity. These are all (with the possible exception of turn-taking) present in my musical task. I needed a non-musical task which involved matched levels of co-ordination, potential for entrainment and synchrony of movement, some sort of intention-reading - but just differing in "musicality".

Musical and non-musical manifestations

My definition of the musical is somewhat broad, encompassing many forms of multimodal, communicative activity, but especially those on the intrinsic end of our spectrum: it foregrounds the relational goals, or the phatic dimension. I have already described how I can make this slightly more "speech-like" by introducing an external depiction goal. However, I needed to change the modality more significantly in testing, in the first instance, the efficacy of more musical forms of interaction for group bonding more generally. In order to do this, I drew on descriptions of the 'statistical universals' of music-making, ensuring that a musical task included these to some extent, while the non-musical task did not. I shall outline these features in brief here.

These musical 'design features' might include: a discrete set of pitches, some sort of scale system that can be used to create all sorts of different melodic or harmonic combinations, an isochronic pulse, focus on contour and the transposability of melodies, all within a performative or ritual context (see complete list in Fitch, 2006). Similarly, Trehub et al. (2015) identifies the presence of specific melodic units or phrases, the predominance of unequal intervals, especially pentatonic scales, as well as a periodic pulse and the social occasion of music-making, as statistical universals (see a similar list in Stevens & Byron, 2016). While Bispham (2009) also highlights the role of pulse

and pitch contour, he situates these within a broader social context or framework: this provides the overriding *motivation* which defines musical activities - namely, 'an intrinsic motivation to share convergent intersubjective endstates' (Bispham, 2009: 41). My musical condition comprised only this broad social framework (of course, using an instrument that emphasises melodic contour). However, within this interactive, improvised and free context, participants did in fact then incorporate the more superficial features - playing in synchrony, repeating certain notes, even sometimes using what sounded like a pentatonic scale.

My non-musical task also provided a social context for free interaction that was not scripted in any way. Participants had to work together, co-ordinating with the others in their group. Again, no speech was allowed. In contrast, though, they had no instrument by which they could include any sort of melodic contour or prosody to engage with their partner(s). Without this dimension, facilitated by the musical instrument, we needed some other task to scaffold participants' interaction, a context in which they had to work together to some end. From very early in our lives, we engage with others around us by sharing *visual* attention, pointing to objects and directing others to attend to something alongside us. Therefore, a visual task seemed to be a simple context in which participants could interact and affiliate, in a non-musical way.

Participants in this non-musical condition were to arrange shapes in a collage, forming one picture as a group. This was just as engaging as the musical task (both were run in pilot sessions); it was very easy to take part in, involved participation and contribution from all group members, who had to work *together*, simultaneously, in just the same way as in the music task. They had to work with a great variety of shapes, with the potential to arrange them in many different ways (matching as far as possible the generativity and freedom of the musical activity).[96] Just as tool-use is a normal and natural feature of music-making, moving objects is a core part of all sorts of everyday interaction (see Streeck et al., 2011). Any non-musical task had to include some sort of external goal: while music is characterised by an unusually relational goal focus, other forms of interaction are generally more extrinsically-oriented, i.e. achieving common ground on something external is a crucial part of the encounter. Thus, the visual shape task always involved the depiction goal - depicting elephants or birds (as pictured on the instructions), a task which was either shared or not shared within the group.

Summary: experimental conditions

To complete the first half of this chapter, I present again the place of these interaction conditions in my experimental designs. Tables 4-7 below provide a summary of each.

---

[96] See Merker et al. (2015) on generativity.

Table 4 *Summary of conditions in Experiment 1*

Variables tested	Testing the effect of interaction <u>goals</u> on group formation		
Testing the effect of interaction <u>mode</u> on group formation	**Musical** group interaction, with depiction goal **shared** within group	**Musical** group interaction, but depiction goal **not shared** within group	**Musical** group interaction with **no** additional goal
	Visual group collage, with depiction goal **shared** within group	**Visual** group collage, but depiction goal **not shared** within group	

Table 5 *Summary of conditions and outcomes in Experiment 2*

Interaction conditions	Musical group interaction with **goals** corresponding to team identity (mixed teams, mixed goals within group)	Musical group interaction with **no** additional goals (mixed teams within group)
Interaction outcomes		
Group bonding across teams		Affiliation to individuals on basis of team membership; team identification in general

Table 6 *Summary of conditions and outcomes in Experiment 3*

Interaction conditions	Musical dyadic interaction with shared **goal** corresponding to team identity (same-team pair)	Musical dyadic interaction with **no** additional goals (same-team pair)
Interaction outcomes		
Affiliation towards interaction partner		Team identification in general

Table 7 *Summary of conditions and behavioural measures in video analysis project*

Interaction conditions	Musical dyadic interaction with shared **goal** corresponding to team identity (same-team pair)	Musical dyadic interaction with **no** additional goals (same-team pair)
Behavioural observations	Degree of vocal synchrony Incidence of mutual gaze Relationship between gaze and vocal behaviour	

Groups and teams in the lab

Now that I have presented my experimental conditions, I turn to my dependent, or outcome, variables. I have drawn from the literature (and, in some instances, devised myself) indices which tap into a sense of group or collective belonging, within the social landscape I constructed in the confines of the lab. Important here is the distinction I make between a feeling of bonding, or belonging, to the face-to-face group formed in the immediate interaction context through the experience of *interdependence*; and the feeling of belonging or identification to a much wider team membership, a more individual process of self- and other-*categorisation* which might be reinforced by certain features or attributes of that interaction context. To investigate these processes separately, I needed different outcome measures; in this section, I shall discuss each measure in turn - in particular, why I expected it to tap into one process more than the other. I start with those I used to test interpersonal, interdependent group formation.

The evidence for group formation

In experiment 1, and also to some extent in experiments 2 and 3, I tested how far participants felt that a group had formed as a result of the time they spent interacting together (musically or otherwise). This process of group formation is often equated with that of 'social bonding', defined as:

> ... the psychological experience of increased social closeness, reflected in prosocial behaviors
> (Tarr et al., 2014: 1)

However, it is also important to distinguish between bonding towards individuals in the same group, and the sense that all of the individuals involved have been bonded into one 'entity' (prosocial behaviour might stem both from individual relationships, and from being in a group context, e.g. Kirschner & Tomasello, 2010).[97] Pearce et al. (2016b) investigate something like this distinction by measuring specific individual relationships which have been formed (e.g. by writing down names of others), as well as a sense of collective identity (the latter has some overlap with my construct of team identification - explored later). Here, I consider both of these dimensions - the personal relationships, and the general collective bonding - as important facets of group formation. I include both types of measure, but ensure they are clearly distinguished. Instead of measuring prosociality, I probe participants' *feelings* about the social relationships they have formed; I assume that this experience of closeness or connection is a precursor to (or at least underlies) any behavioural manifestation of in-group preference. I shall now go through each measure in turn, providing evidence which supports its validity for this purpose.

---

[97] The question of the relative importance or sequence of these bonding processes is a subject for future investigation (furthering the work of Pearce et al., 2016b).

Trust - the definition of a group

In his work on social networks, Robin Dunbar considers that what defines the core group size of 150 is the perception and expectation of trustworthiness for all other individuals in that group. He says:

> What defines this group of people seems to be recognition of a mutual, reciprocal relationship of trust and obligation, combined with a willingness to act prosocially. (Dunbar, 2018: 34)

This acts on the same two levels - trust for specific individuals, and a general feeling about the trustworthiness of the group as a whole. Dunbar characterises individual friendships as 'implicit social contracts' (Dunbar, 2018: 37, 41); time spent interacting with a specific other results in greater emotional connection and more trust in them (a commitment that does not rely so much on immediate reciprocity - see Sutcliffe et al., 2012). There are important neurochemical correlates of these bonded relationships too - feelings of trust are linked to the release of oxytocin, the neuropeptide which plays a large part in selective attachment relationships (Kosfeld et al., 2005).[98] I incorporated a measure of trust towards specific others in a questionnaire following the group interaction: participants were shown photos of others (who either belonged to the group, or who did not), and had to make forced-choice decisions between them as to who they would trust *more* for certain scenarios (adapted from Johnson-George & Swap's 1982 Specific Interpersonal Trust scale). These judgements are thus personal evaluations of how much they trust a specific other (or who they trust more).

However, trust is important within the much broader group context - we need to trust those who we do not know so well; and this is perhaps the criterion for who belongs to our wider social group, and who does not. Rotter (1971) highlights this in his conceptualisation of trust - it is something more general:

> Interpersonal trust is defined here as an expectancy held by an individual or a group that the word, promise, verbal, or written statement of another individual or group can be relied on. (Rotter, 1971: 444)

Importantly, this is a 'generalised expectancy' - derived from individual experience, but applied across the board, even generating trust in our society as a whole.

> The entire fabric of our day-to-day living, of our social order, rests on trust - buying gasoline, paying taxes, going to the dentist, flying to a convention - almost all of our decisions involve

---

[98] Interestingly, oxytocin levels also increase following group singing (Kreutz, 2014).

trusting someone else. The more complex the society, the greater the dependence on others. If trust weakens, the social order collapses. (Rotter, 1971: 443)[99]

Therefore, we need to include another layer of trust - not just towards specific individuals, but to 'classes of significant others' (Rotter, 1971: 446), in other words, to the experimental group as a whole. I included this very simply with self-report Likert scales where participants had to rate how far they would 'rely on' the others, how far they have 'confidence' in the group, and how far they think they are 'generally trustworthy'.

Importantly, this feeling of trust, or expectation of goodwill on the part of others, should be distinguished from a sense of commitment - a different emotion, but one which also emerges within a group context (and which exacerbates the boundary between the in-group and any outsiders). Here, we deal with trust, as it is the 'booster rocket' - enabling relationships with others who we might not know so well (as in the case of our experiments; Yamagishi & Yamagishi, 1994: 138). However, we need trust less when other conditions or loyalties *oblige* us to co-operate with specific others via our commitment to them: perhaps due to the threat of punishment, or loss of reputation, preventing us from seeking new relationships with others outside this closed network (Yamagishi & Yamagishi, 1994).

Feeling close: self-other overlap

Tarr et al.'s (2014) definition of bonding includes the experience of 'closeness'; in their review, they highlight the role of self-other overlap in generating this feeling. This is elsewhere described as the sensation that the boundaries between us and another have been blurred, probably due to our mirroring mechanisms - the strong connections between perception and action in our neural responses to observing others' actions. When we move with another, watching them and doing the same thing ourselves, it starts to become difficult to work out who has done what (for example, which of us played the wrong note in our piano duet).

This sense of overlap, or closeness, has been measured using a pictorial scale called the 'Inclusion of the Other in the Self scale' (the IOS scale). This is a single-item Likert scale, originally devised by Aron et al. (1992), where the participant and their target (usually another person, with whom they have interacted in some way) are depicted as two circles which are more or less overlapping. Participants have to decide which picture best represents or describes their relationship. Aron et al. state that this was 'intended to tap directly people's sense of interpersonal interconnectedness' (Aron et al.,

---

[99] Yamagishi and Yamagishi have since added more nuance to this definition - suggesting that 'confidence' is an index of social competence, whereas we trust in another's goodwill in their dealings with us (Yamagishi & Yamagishi, 1994: 131).

1992: 597); it corresponds to feelings that attributes of the other have become one's own, incorporated into the self (and vice versa) (see validation of this scale by Gächter et al., 2015).

Again, in my case, I needed to apply this sense of 'overlapping selves' and 'closeness' on a group scale (not just towards individuals). A few others have modified the IOS scale such that the target is a group as a whole. One such attempt is by Swann et al. (2009), who use it as a measure of 'identity fusion' - the idea that one's personal identity has been fused with their sense of belonging and contributing to a group:

> ... the self-other barrier is blurred and the group comes to be regarded as functionally equivalent
> with the personal self. (Swann et al., 2009: 995)

Although they consider predominantly how people act on behalf of their groups following a process of self-categorisation (something we shall explore more below), often with extreme behaviour, they also state that the concept of identity fusion, and the IOS scale used to test it, can be applied in investigating how far an individual bonds with their 'relational groups' (presumably similar to the face-to-face group bonding we are concerned with here; Swann et al., 2009: 995).

They highlight again how this pictorial scale is useful in tapping into just a feeling of closeness, not even one which is easily accessible or understood; this IOS scale is

> ... a straightforward index of degree of alignment with the group that can draw on conscious as
> well as nonconscious material. (Swann et al., 2009: 997)

Thus, although it is a self-report scale, it perhaps draws on more implicit feelings of closeness or a general sense of overlap - feelings which correspond to a sense of group formation.[100]

Connectedness

Other self-report items I presented to my participants perhaps show the results, or manifestations, of such feelings of trust or closeness. The first of these is a rating of interpersonal 'connectedness': a 7-point Likert scale in response to the question 'How connected did you feel with the other participants in your group during the activity?' This is a fairly common measurement of group bonding or group formation. Wiltermuth and Heath (2009) found that their participants feel more 'connected' following synchronous than asynchronous walking. Participants in singing classes form more individual relationships or 'social ties' - specific people to whom they feel 'connected' (in a yes/no question) - more quickly than those who took part in craft or creative writing (this occurred alongside a general increase in group-level IOS scores; Pearce et al., 2016b).

---

[100] See also how it is used by e.g. Pearce et al. (2016b) as an indicator of 'collective bonding to the class as a single, non-individuated entity.' (Pearce et al., 2016b: 506)

While these uses of 'connectedness' concern specific relationships, Pearce et al., (2016a) combine general scores of connectedness with the IOS ratings to form an overall index of 'closeness' - this time directed to the group of four as a whole, not towards other individuals. Weinstein et al. (2016) also show higher levels of 'connectedness' felt towards all others in a large group of a couple of hundred people. Pearce (2014) also relates this same concept to larger-scale group structures, discussing how the overall, wider social network of weaker ties might be maintained through occasional social gatherings or encounters (like large-scale music-making).

Both of these levels are present in a definition of the term given by Lee and Robbins (1995). They speak of a general 'sense of connectedness' amongst all others, being "human among humans" (Lee & Robbins, 1995: 233), the feeling that one can participate in, belongs, and feels related to others in society; ratings on this scale also seem to predict the ability to identify with *specific* others, such as partners or friends.

Similarity

The feeling of being similar to another is semantically related to the idea of closeness or self-other overlap; however, this dimension is specifically directed towards other individuals, concerning specific interpersonal relationships rather than a sense of collective belonging. The question I used in experiment 3 (considering participants' relationship with their single interaction partner) was a 7-point Likert scale - 'How similar do you feel to your partner?' - taken from Rabinowitch & Knafo-Noam (2015). They discuss how similarity is strongly linked to the mechanisms supporting smooth interaction - the tendency to mimic others, for example (Chartrand & Bargh, 1999) - and enhancing affiliation between participants. In their design, the general question about similarity is presented alongside other questions about one's perception of the other person (similarity in appearance, character, hobbies etc.) - it is the overall sense that the other is "Like Me".[101] Moreover, Valdesolo and DeSteno (2011) found that similarity is an important mediating factor between synchrony and affiliation, compassion or prosocial behaviour.

Likeability

Liking is another measure concerned with one's direct relationship with another - an indicator of individual affiliation. I have used the 9-point Likert scale 'How likeable is your partner?', taken from Hove and Risen (2009), for measuring interpersonal bonding in experiment 3. Hove and Risen linked the concept of liking, or affiliation, to social cohesion and self-other closeness, perhaps reinforced by mechanisms such as action-perception resonances and mirroring, towards their specific tapping partner in the study.

---

[101] Compare this with the discussion of imitation in Chapter 1.

Memory for group members

Memory for others is another measure of attention or bonding towards specific individuals. Woolhouse et al. (2016), in their 'silent disco' study, used memory of a certain feature (whether other participants had a cat logo on their sash) as an indicator of social bonding. In their literature review, they draw attention to the role of memory in forming attachment relationships - individuals need both the motivation to approach another and engage with them, and then the formation of social memories - specifically for that individual - so they can recognise the other when they encounter them again.

I incorporated memory tests about particular features of others in experiment 1, but with the purpose of assessing what participants can remember about *all* others in their interaction group. This accommodates the idea of individual bonding, but within a broader context of group interaction and cohesion. Specifically, I ask participants how many members of their group (a) were wearing glasses, (b) have blonde hair, and (c) were wearing a pink wristband.[102]

Entitativity

One final measure, used in assessing participants' perceptions of the whole-group interaction context in both experiments 1 and 2, was the index of 'entitativity'. Rutchick and colleagues (2008) define this as 'the extent to which a set of people is perceived as a meaningful group', or an 'entity'. This measure comprises three separate pictorial scales, one assessing overall 'entitativity' (a set of circles which are more or less overlapping), and then the others each testing the importance of certain cues, or antecedents, for this entitativity; namely - similarity (to what extent the group members share something, perhaps appearance or traits, in common, have a 'common essence' or homogeneity); and action and interaction (a sense of interdependence, working together towards a common goal, perhaps with some sort of group organisation). I included all three of these scales in my questionnaire.

This item spans both the idea that groups are formed through interdependence (the suggestion is that entitativity which emerges from a higher rating of action and interaction indicates an interdependent group), but also categorisation (entitativity instead relating to higher perceptions of homogeneity or similarity; see Rutchick et al., 2008). I shall next go on to discuss this process of categorisation in more depth - first, the way in which I used the experiment set-up to create possible

---

[102] Certain limitations of (a) and (b) should be noted: pictures of the other group individuals had been included earlier in the questionnaire (in the forced-choice trust measure, for example) - some participants might have used this rather than their memory to inform their answers; also, it is difficult to assess what 'blonde' hair is (how dark can it be to still be considered blonde).

social identities, and second the way I measured how they were established in participants' self-concept following goal-focussed interaction.

Summary of measures

I summarise the items measuring group formation in Table 8 below. I split them up according to whether they measure interpersonal bonding (towards specific individuals within the group), or a broader sense of whole-group cohesion.

Table 8 *Summary of group bonding measures*

Whole-group bonding	Interpersonal bonding
General trustworthiness of the group	Specific interpersonal trust
Inclusion of the other (group) in the self	Similarity
Connectedness	Likeability
Entitativity (through similarity or interaction)	Memory for individual attributes

Team formation: creating and measuring identities

The line between forming a group through interdependence and face-to-face personal bonding, and a sense of team, collective identity, is somewhat blurred - of course, face-to-face bonding can lead to collective identity, and a prior sense of team membership makes individual affiliation easier (compared to relations with others in the out-group). I alluded to this by including a measure of entitativity which encompassed both the similarity and the action-interaction criteria as possible antecedents of each group 'type' (as either 'categorically-' or 'dynamically-construed' - see Rutchick et al., 2008). In my first experiment, I also decided to include an extra 7-point Likert item taken from Wiltermuth and Heath (2009) - 'How far did you feel you were on the same team with other participants in the task?' - despite not including any reference to 'teams' in the experiment set-up. In these ways, I acknowledged the fluidity of these processes; it is difficult to separate one from the other, as they feed into each other, and are deeply associated within the mind of each individual participant.

Studies often merge these two concepts together. For example, Good and Russo (2016) measured the impact of various group activities (singing, art, or competitive games) on co-operation with that group (through a prisoner's dilemma game); but, the singing and art conditions involved making up, and using, words describing positive things about living in their home city - this could well comprise and reinforce a collective identity that the participants shared. I wanted to measure a sense of collective identity *separately* from the more local sort of group bonding, as far as was possible. This is important for the aims of my project: it seems that music-making can establish group category

162

divides, as well as forming group bonds, in slightly different ways - I attempted to devise experiments which clearly enacted and tested this distinction.

In this section, I shall outline how I introduced the idea of opposing team identities at the outset of my experimental procedure; and then, how I investigated the role of musical goals in reinforcing that sense of identity amongst the participants.

Team competition: experiment set-up

In both experiments 2 and 3 (and also in the conditions for my video analysis), I divided participants into two teams. I aimed to create the basis from which a sense of team identification might then emerge as a result of certain properties of the interaction. All participants within a given experimental session (ten for experiment two, and four in experiment three) were divided into two equal, opposing teams - creating a clear sense of "us and them", an in-group against one out-group (Hartstone & Augoustinos, 1995).

Experiments conducted elsewhere have used all sorts of criteria to distinguish two groups in this way: existing membership of a university (Levine et al., 2002) or a social club (Pearce et al., 2016a), or assignment at the start of the experiment to phis or gammas (Locksley et al., 1980), or reds or blues (Miles et al., 2011 - supposedly on the basis of 'art preference'). Participants' team identity is often then displayed through some colour or symbol, or other distinguishing, perceptual, characteristic: an outward sign (e.g. a Crucifix or 'Got God' bracelet) of being a Christian (Yabar et al., 2006), the presence of a university logo (Reddish et al., 2016), or wearing a coloured sticker or t-shirt (Miles et al., 2011; Good et al., 2017, and others). My procedure has been influenced by the "minimal" group formation stage of studies such as Gaertner et al. (1990), Good et al. (2017), and Tunçgenç & Cohen (2016).[103] In these, participants are assigned to one of two colour identities, then go to a given team location, where they might complete an activity together (e.g. drawing a team 'flag', thinking of a group name, maybe together discussing the 'winter survival problem' or telling others something about themselves). Given that I aimed to test the effect of group interaction, my team formation stage just included establishing separate identities and creating the sense of a competition between the wider team memberships (beyond those present in the experiment) - before they engage in within- or across-team interaction (in experiments 3 and 2, respectively).

First, I created separate team identities by instructing half of the participants that they were in the "elephant" team, and the other half the "bird" team. I established a *perceptual* distinction by giving each individual a badge to wear, clearly showing either a bird or an elephant team symbol. In experiment 2, participants started the experiment seated in a circle with other team members (with

---

[103] These groups are generally not very "minimal".

the other team in another circle, elsewhere in the room). Similarly, in experiment 3, same-team pairs were seated next to each other, and then were sent to do their dyadic interaction in adjacent rooms. Thus, they were always aware of their own team membership, as well as the presence of an out-group.

Second, I instructed participants that they were part of a large-scale competition - involving elephant and bird team members from across the faculty. The experimental tasks (the interaction, and, in experiment 3, the memory test) were posed as opportunities to gain points for their team (video cameras were set up to reinforce this, supposedly recording their behaviour for this purpose). Brewer (1979) notes greater in-group preference in conditions of competition between the two teams; of course, Sherif's (1967) study also highlighted this as one instrumental component in forming team or group identity. I used these procedures in my studies as ways of introducing a clear sense of team membership before I tested the effect of group interaction goals on the social categorisation process (see above on how these goals resembled this perceptual team distinction).

Measuring team identification: in-group preference

The most common way in which social identity, or team formation, has been measured in the past, is to establish whether there is a greater preference for in-group members, or a particular discrepancy between the treatment of those belonging to the in-group compared to others who are part of the out-group. This is an important difference, compared with indicators of group formation: whereas the latter measure bonding, or affiliation, inclusion within the group, the experience of team identity always necessarily includes a distinction between who is in and who is not.

There is a whole range of indicators of this sort of preference, from facial recognition (Bernstein et al., 2007), amount of behavioural mimicry (Yabar et al., 2006), to distinct neural responses while viewing others (Ratner & Amodio, 2013) or more active prosocial behaviour (Stürmer et al., 2006). The original 'minimal group' studies, however, assess participants' strategies in reward allocation: Tajfel et al. (1971) use a point allocation matrix, and find that participants award more points to in-group members, maximising the difference in point allocation rather than making sure they award as many points as they can. Locksley et al. (1980) use a simpler reward or 'chip' allocation measure, where participants just have to decide how many of their 100 chips they can allot to in-group or out-group members (phis or gammas); Hetherington et al. (2014) include a similar resource allocation element in their design. In general, these sorts of decisions are made individually - participants just see whether another is part of their team or not from a code number (e.g. in Hogg & Turner, 1985; or Turner et al., 1983), or a coloured background, and make the choice purely on this basis. With the exception of Hetherington et al.'s study, they do not see the identity of the target individual;

sometimes they just see unfamiliar faces, which are counterbalanced for team affiliation across participants.

I incorporated a measure of this nature in my third experiment: I provided pictures of individuals who were unfamiliar to my participants, alongside a symbol which indicated their team affiliation.[104] These individuals were shown in pairs - one associated with the elephant, and the other with the bird team - and participants had to decide how much to reward each, from a limited resource for each pair. They were instructed:

> Below are pictures of other members of the bird and the elephant teams who have taken part in the experiment already. Their team is shown by the symbol. Imagine that you have 4 pots of £5 in £1 coins. ... Decide how many £1 coins to allocate to each participant and write a number in each box. This must total £5 on each line.

This was thus a forced-choice task, where participants have to choose who to give more or less to, just on the basis of team identity.

I did not use just a reward allocation paradigm to test team formation in this way. In experiment 2, I also adapted my interpersonal trust forced-choice measure (see above) to test how far team affiliation shapes individual evaluation or judgements of trustworthiness: participants had to choose between more unfamiliar individuals (this time, they were from their cohort, but not involved in their experimental session), who had been assigned a team identity, with the relevant symbol displayed next to their picture. Moreover, in experiment 2, I included several different self-report ratings of connectedness: participants had to rate, in turn, how connected they felt to members of their own team, and members of the other team, thinking only of those involved in their mixed-team group interaction activity. These three dimensions - reward allocation, trust, and connectedness - all provided potential indicators of in-group preference, individual or group-level bonding *on the basis of* team membership.

Self-other overlap - with the collective

The measure of self-other overlap described above also does not just extend to general group-level bonding; more commonly, in fact, it is applied to a much larger-scale category identity. Swann and colleagues' concept of identity fusion is explicitly linked to ideas of Social Identity and Self-Categorisation, as individuals 'fuse' their personal identity with that of their e.g. religious or national group, and therefore feel they - even as an individual - can act on behalf of this wider category (Swann et al., 2009). The same 5-point pictorial Likert scale can be directed towards the team as a

---

[104] This measure was actually limited in validity as some participants claimed to recognise the photographed individuals.

whole - not just the other individuals present in the experiment session (in my study, I make this clear e.g. by presenting the questionnaire item next to the particular team symbol). Other researchers have used the IOS scale to measure a sense of in-group identity. Tropp and Wright (2001) discuss the process of self-categorisation as follows:

> When individuals categorize themselves as group members, the ingroup becomes included in the self and individuals recognize the characteristics of the ingroup as representing part of themselves. (Tropp & Wright, 2001: 586)

Also, Schubert and Otten (2002) further explored how the IOS scale could be adapted to capture inter-group relations more fully - measuring not only the overlap of self and one's in-group, but also how we relate to our out-groups, and how overlapping we feel our whole in-group is to the relevant out-group.[105]

Team-related memory: implicit identification?

Finally, I devised a more exploratory measurement of individual identification with a team that did not involve rating other individuals (especially other individuals who were involved in the interaction task). Having an implicit measure of the salience of one's social or team identity - that could not be confounded by having interacted with specific individuals more than others - was important in distinguishing group bonding from a sense of team identity.

Implicit measures have been useful in various investigations of social cognition, uncovering attitudes or stereotypes which might not be so readily observable, especially through self-report. Tasks involving word completion or reaction times might demonstrate such 'traces of past experience', being affected by factors or memories of which a participant might not even be consciously aware (Greenwald & Banaji, 1995). An individual's memory is a key part of this process. Certain tests of racial prejudice (notably the Implicit Association Test) use a priming procedure, and measure how far certain concepts are more or less congruent with certain categories of individual (using reaction times) in the memory representations, expectations (or stereotypes) held by the participant (see Fazio & Olson, 2003, for a review). Therefore, it makes sense that establishing a certain team identity in the lab might shape just how one recalls semantically-related concepts or words in a subsequent memory test.

Tests of recall or recognition memory are generally affected by the learning context (see e.g. Smith & Vela, 2001; or, earlier, Smith, 1986). In a perhaps similar way, I supposed that if the musical interaction task had instilled more of a sense of team identity in the participant, then they might be

---

[105] A 'group' in this context refers to what I generally call a 'team': a sense of social identity, beyond those present in an interdependent, face-to-face interaction.

more prone to remember words related to their team in a subsequent written recall test, perhaps through a process of priming, or a memory bias effect.

Summary of measures

Table 9 below lists the indices I used to measure different aspects of team-related affiliation or preference - relating to specific individuals with whom they had been directly involved, other individuals who shared a team membership, or a more general, wider sense of identification with the team. In general, these tap into one's own affiliation with a team, but sometimes this also results in a distinction between those who are in, and who are out - a process of exclusion (as opposed to the bonding, or inclusion, measured in the section on group formation above). This is something that needs to be investigated further, beyond this initial project.

Table 9 *Summary of team-related affiliation or identification measures*

Target > v Experiment	Known individuals	Unfamiliar team members	Team identity
2	Connectedness	Trust	IOS
3		Reward allocation	Memory & IOS

Controls, feelings, and other considerations

I shall conclude this chapter with a few other considerations which were important for this project. The first are the normal dimensions that should be controlled - degree of musical training, prior knowledge of others in the experiment, and mood or affect. As I discuss in depth above, my musical interaction tasks were designed to be accessible regardless of musical background or expertise, and so should "level out" any participant differences in this domain; however, I measured this in one experiment (and kept it roughly constant in the others), just to ensure this did not affect the results in any way.

Positive mood or affect might well affect processes of bonding or team identification (although it is unclear whether this mediates the bonding effect, or just correlates with other mechanisms, such as endorphin release). Weinstein et al. (2016) found that positive affect correlated with other indicators of group formation (e.g. IOS). Moreover, affect contagion is one outcome of shared group interaction or experience (see e.g. Barsade, 2002). Therefore, I used the Positive and Negative Affect Schedule (PANAS) - a short self-report measure (Kercher, 1992; reviewed by Mackinnon et al., 1999) - both

before and after the experiment to control for any confounding effects of mood, or to measure where increased positive mood might occur in tandem with other group bonding effects.[106]

Beyond these normal control variables, I expected that the motivation for a participant to identify with a social class or group, in other words, to adopt that team identity as part of their self-concept (and so feel affiliation towards others who share that attribute - see discussion of self-esteem and social groups in Chapter 3) would be shaped by other social emotions. These are not necessarily indicators of bonding, but are other feelings that arise from the experience of the interaction task. In particular, I expected that the experience of *embarrassment* would have a significant effect - perhaps making people less likely to want any sort of association as a result (see discussion of this, and the motivation to 'save face', in Chapter 2). In my experiments, I included self-report questions about participants' perception of the joint or group successfulness of the task, how difficult they found it, and whether they think they individually performed well. These items give an indication of how well they felt they presented themselves, how well they worked together, but also whether they *individually* might have been responsible for the success of the task. This was exploratory, I expected that these might shape how far participants felt they could identify with - or needed to distance themselves from - the social context created through musical interaction. In experiment 3, I introduced a further item which measured more directly how embarrassed they felt during the task (from Sabini et al., 2000). I did not consider this a confounding variable. Rather, avoiding potential embarrassment, and the task of impression management, is a central part of managing social interactions, and a factor shaping our sense of group identity. Including this in our investigation into social, musical interaction is necessary to begin to understand how music-making functions in social bonding (and in creating social distance or category divides) in the light of these other interpersonal factors.

In this way, I have developed a thorough research paradigm to explore the role of music-making in forming social groups, and in distinguishing social categories; and to differentiate particularly "musical" forms of interaction from others. I shall now report on each of my experiments in turn, and then present the results of my video analysis.

---

[106] This is a tricky issue of causality: I am measuring the effect of musical interaction on group bonding effects; however, this might be closely linked to the affective experience, especially given the rewarding nature of social interaction. Compare results and discussion of experiment 1 with those of experiment 3 for more discussion of this: musical interaction did not correspond to changes in self-reported mood (in experiment 1), however, baseline mood scores correlated with the subsequent success of the task (in experiment 3).

Chapter 6 - Experimental groups

Here, I bring in my empirical work: new data to shed light on our central question, that is, the role of music-making in forming and demarcating group identity. I present the method, results, and brief discussion for each experiment in turn, followed by my video analysis. In my conclusions, I shall discuss in greater depth the insights offered by this work, and the possible direction of future research.

Experiment 1 - music in group formation

The aim of this first experiment was to investigate the conditions of group formation: contrasting a "musical" interaction (very simple, instrumental music-making) with a "non-musical" (visual) interaction mode; comparing the effect of shared with conflicting goals, and also the effect of having or not having an interaction goal at all - on subsequent group bonding. The following were the hypothesised effects:

(a) A higher level of group bonding following "musical" interaction than after a "non-musical", visual, interaction task

(b) A higher level of group bonding following interaction with shared goals than after a task involving conflicting goals (see Chapter 3 for discussion of interdependence in group formation)

(c) An interaction - the effect of sharing goals is more important in the non-musical condition than in the musical condition (where non-shared goals can co-exist in a context of floating intentionality)

Method

Participants

A total of 40 undergraduate students (20 male and 20 female) took part in this experiment. They were aged 18-21 years (mean 18.58; two participants did not provide their age). All participants had a high level of musical training (they were recruited through the music faculty). They all gave informed consent to take part in the experiment, which was reviewed by the Faculty of Music Research Ethics committee.

Design

This comprised a between-groups two-way 2 x 2 design, varying the mode of interaction (music/visual), whether or not groups shared a goal (shared, not shared), with a no-goal control

condition (see Table 10 below). All participants interacted in same-sex groups of four. Participants were split equally between each condition, with one male and one female group in each.

Table 10 *Conditions in experiment 1*

Music & shared goal	Visual & shared goal
Music & non-shared goal	Visual & non-shared goal
Music & no goal	

Note: n = 8 in each condition.

Each participant assigned to one of the goal conditions was given a task to complete with others in their interaction group - to depict either elephants or birds. In the shared goal condition, all group members were given the same animal to depict; in the non-shared goal condition, two were given the elephant, and two the bird goal. Participants assigned to the no-goal condition were not given a depiction task; their only aim was to interact with one another. The distribution of tasks is shown in Table 11.

Table 11 *Number of participants assigned to each condition and to each specific goal type*

Mode of interaction	Shared goal?	Task assigned	No. participants	
Music	Shared	Bird	4 (1 group)	
		Elephant	4 (1 group)	
	Not shared	Bird	4	(2 groups)
		Elephant	4	
	No goal	No goal	8 (2 groups)	
Visual	Shared	Bird	4 (1 group)	
		Elephant	4 (1 group)	
	Not shared	Bird	4	(2 groups)
		Elephant	4	

In line with previous literature, a range of outcome measures were used to test the level of group formation; these were administered via a post-interaction questionnaire. These included in-group trust (towards specific others and the group in general), memory for group member attributes, ratings of connectedness, closeness to the group (in an IOS scale), and perception of being on the same team as other group members. Participants were given a free written response task to describe their experience of the interaction. (See details of the measures in Appendix 1.1.) The short form of the PANAS mood questionnaire was administered pre- and post- interaction task (Kercher, 1992; reviewed by Mackinnon et al., 1999). Participants were also asked to rate (and detail, if necessary)

how much they knew the group members (7-point scale), and the designated out-group members (for the trust measure - see below) prior to the experiment.

Procedure and apparatus

Pictures of all participants were taken, with consent, the previous day. These were used in the specific trust measure.

Participants began the experiment seated individually, where they signed the consent form, completed the pre-task PANAS and then read their individual instructions. Pink wristbands were distributed at this point to two members of each group.

Participants were led to believe that this was an experiment in group memory and non-verbal communication, and thus should be completed in silence; where a depiction task was given, the instructions implied that all group members would share that goal. Participants were not sat near any other interaction group members, ensuring that they were unaware if goal tasks varied within the group (as was the case in the non-shared goal conditions). They did not know who was in their group before the interaction task began. All participants were then instructed to go to one of ten different locations (signposted, and marked by letters A-J), located in four different rooms. There was one group at each location.

The groups assigned a musical task could interact using only a simple vocalising instrument (the kazoo); this was chosen as it is easy to play, not distracting from the interaction with other group members, and encourages face-to-face engagement. Groups with a musical goal were instructed to:

> interact with each other, and improvise together, using only these kazoos, to perform music inspired by a herd of elephants(/flock of birds), as shown in the following picture ... Work together to depict this scene as accurately as you can.

Two groups were given a shared musical goal: all four members of one group were given the elephant task; all members of the other group were given the bird task. Two groups had non-shared goals: in each group, two individuals had the elephant, and the other two had the bird goal (they were not aware of their differing goals). Another two groups were assigned the no-goal condition; they were instructed:

> interact with each other, and improvise together, using only these kazoos

Another four groups were assigned the non-musical, visual, task. Each group was given an identical set of cardboard shapes and instructed to:

arrange these to create an elephant(/bird), as shown in the following picture ... Work together to reproduce it as accurately as you can.

The shared and non-shared goal conditions were operationalised as above (two groups with shared, and two groups with non-shared goal). (See pictures used in Appendix 1.1.)

All groups interacted for 10 minutes, working only within (and not between) their assigned groups for the duration of this time. The musical and visual tasks were piloted to ensure that they were equally engaging for this duration, and that both afforded a high degree of interpersonal interaction in a group of four.

After this task, participants were asked to return to their first location. Their seats were now arranged in new groups of four: the out-group to be compared with in-group members in the forced-choice trust scenarios.[107] After they completed the questionnaire, participants submitted their responses and were led to a separate area to be debriefed about the design and purpose of the experiment.

Data analysis

There were three predictor variables for the initial analysis: mode of interaction (music or visual), goal condition (no goal, shared or non-shared), and the individuals' specific task type (no-goal, depicting elephants, or birds).

The questionnaire data is mostly ordinal (Likert rating scales etc.) therefore quantitative statistical tests could be run, using a mixture of ordinal regression, and non-parametric independent-samples comparisons (Kruskal-Wallis), with t-tests run on the memory scores (treating these as interval data).[108] All post-hoc comparisons were Bonferroni-corrected. The specific trust preference data were calculated as the total number of in-group members chosen in the forced-choice in-group/out-group comparisons. Tests of memory accuracy (correct/incorrect for each) for attributes of other participants were combined, forming a mark correct out of three. All other rating scales were analysed in their raw form. Non-parametric correlations were run on the dependent measures to test how they were related.

---

[107] Participants were sat far enough apart to ensure that they completed the questionnaire individually, and without worrying that others might see their responses.
[108] It is unclear whether the memory scores are interval data, as the scores from all three questions were totalled for each participant. Therefore, the parametric tests run on this outcome should be interpreted with caution.

Results

Baseline mood and prior knowledge of others

Ordinal logistic regression, using mode, goal condition, and task type as predictors, was used to test whether there was significant difference in baseline mood between groups. There was no difference for negative or positive PANAS pre-task baseline scores. Similarly, there was no significant difference between the groups for prior knowledge of the interaction group (or for the trust measure out-group).

Ratings of connectedness

The ordinal regression model (with mode, goal condition and task type as predictors) was significant for ratings of connectedness (Chi-square = 10.245, df 4, p = .037), with effect sizes of r^2 = .226 or .235 (according to Cox & Snell/Nagelkerke).[109]

Within this model, the only component which had a significant contribution was mode (Wald = 5.611, df 1, p = .018).[110] A 2-tailed Kruskal-Wallis test also showed a significant difference in connectedness ratings according to mode: participants in the music conditions rated connectedness higher compared to participants in the visual conditions: H(1) = 6.917, p = .009, Cohen's *d* = .913, *r* = .415 (see Figure 1).[111]

The same ordinal regression model (inputting mode, goal condition and task type) did not predict specific in-group trust scores (nor general ratings of trustworthiness), IOS scores (although this was nearing significance: Chi-square = 8.627, df 4, p = .071), memory scores, or the other ratings of successfulness, performance, same team perception, entitativity, difficulty, reliance on or confidence in others.

---

[109] Observed power = .91

[110] Inputting the same predictor variables into a linear regression model (for connectedness) in stages (1 - mode, 2 - goal sharing & mode, 3 - goal sharing & mode & task type) revealed a significant change when mode was inputted, but not when goal sharing, or goal task type were added to the model (although all three models were significant).

[111] All effect sizes were computed from means and standard deviations for pairwise comparisons using an online calculator found here: https://www.uccs.edu/lbecker/
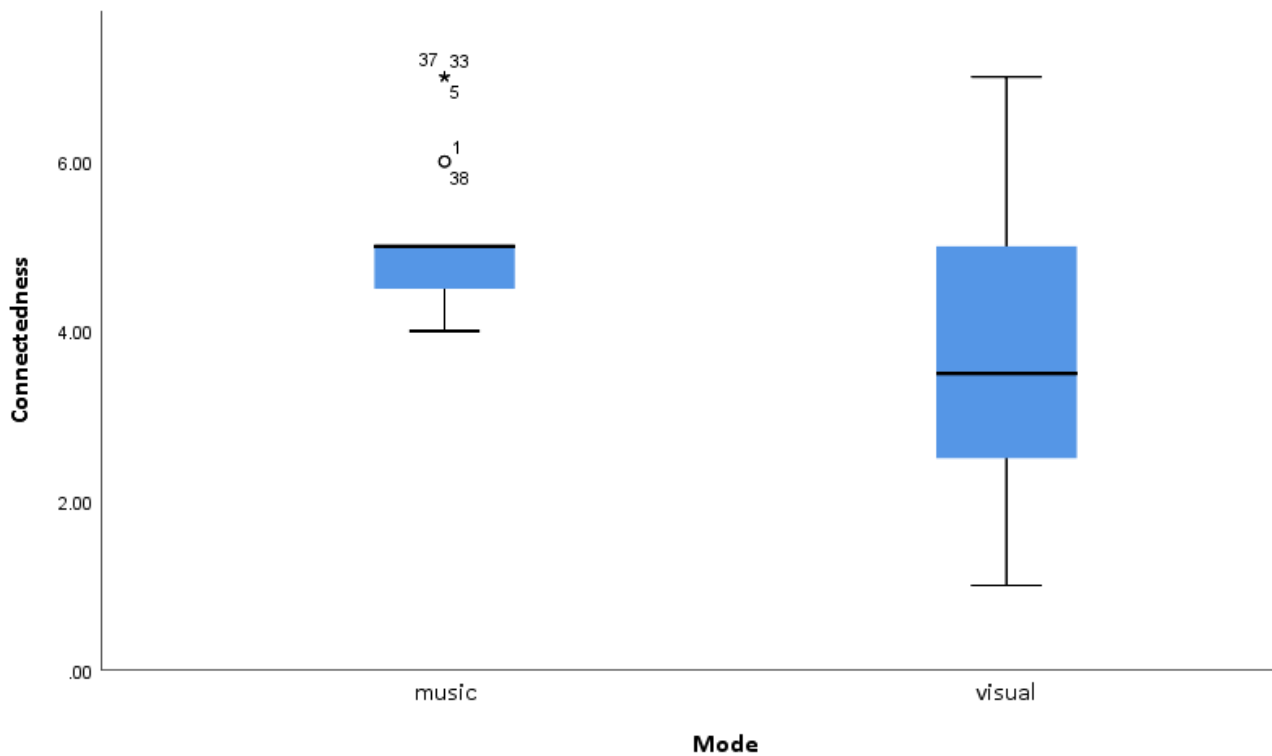
Figure 1 Boxplot showing median, interquartile range (blue shading), and top and bottom 25% of ratings (with outliers at the upper end of the music condition) of connectedness according to interaction mode. Participants in the music condition rated connectedness to their interaction group significantly higher than those in the non-music condition.

Memory for group member attributes

Memory scores were treated as interval data in an independent-samples t-test. Memory scores were significantly higher following the visual interaction task than the music task: $t(37.87) = -2.147$, $p = .038$.[112]

A different ordinal regression model was created, excluding the music no-goal condition, inputting only mode, goal sharing (shared/non-shared) and their interaction. This model was significant for the memory scores (Chi-Square = 9.805, df 3, $p = .020$) with effect sizes $r^2 = .264$ or .297 (Cox & Snell/Nagelkerke). The interaction between mode and sharing was the only significant component: Wald = 2.315, df 1, $p = .005$.[113]

---

[112] These scores are normally distributed; however, Levene's test was significant (p = .003), therefore results given are corrected as equal variance is not assumed.
[113] This same model was not significant for any other dependent measure, including connectedness.

Similarly, an ANOVA with mode, goal sharing, and their interaction as predictors was significant: $F_{(3)}$ = 3.843, p = .020, power = .762, with mode and mode*sharing as significant components. Memory scores were significantly higher following the visual than after the music task – $F_{(1)}$ = 4.392, p = .045; and there was a significant interaction – $F_{(1)}$ = 6.863, p = .014; observed power .525 and .715 respectively (see Figure 2 below).
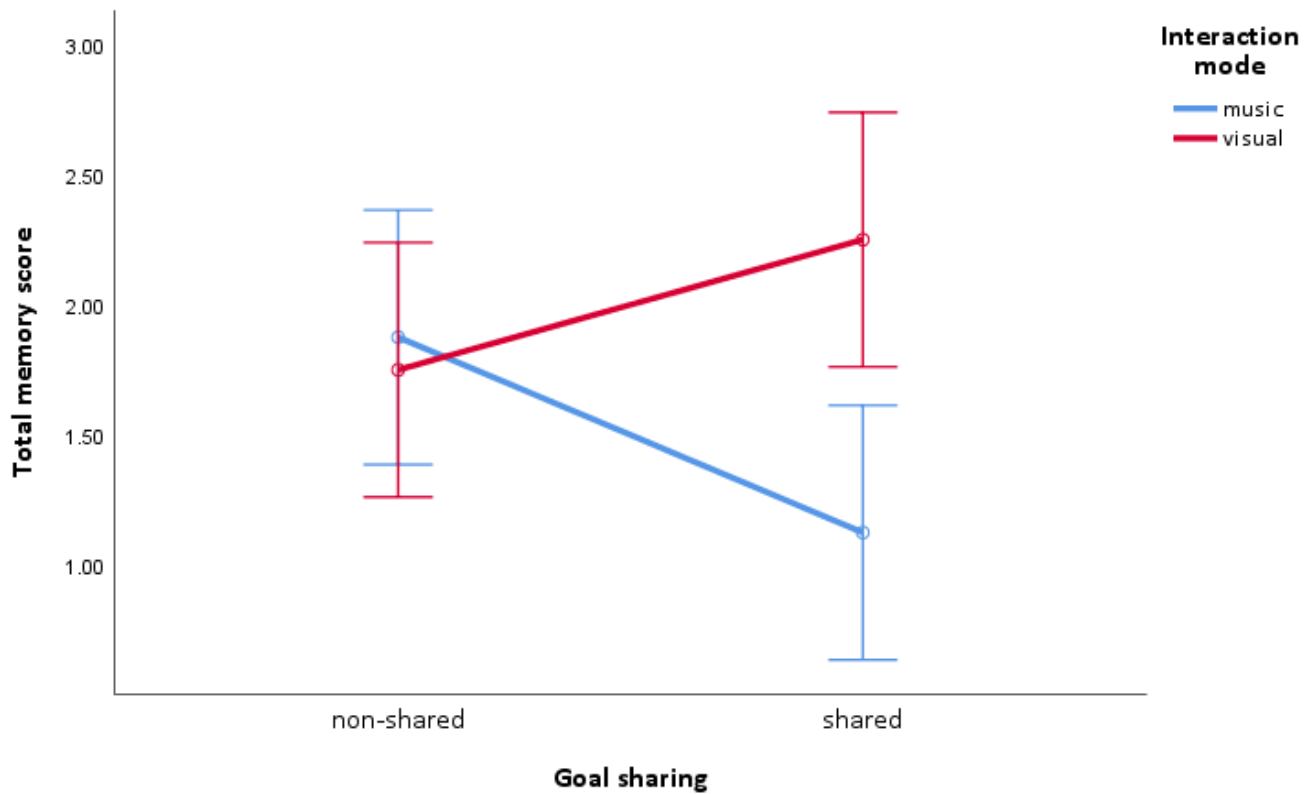

Figure 2 Interaction graph showing estimated marginal means and 95% confidence intervals (error bars). There is a significant main effect of interaction mode, and significant interaction between mode and goal sharing, on the total memory score.

Perceived closeness: IOS ratings

A 2-tailed Kruskal-Wallis test showed a significant difference in IOS scores depending on goal condition: $H(2) = 7.352$, $p = .025$. Pairwise comparisons (Bonferroni-corrected) revealed that IOS scores were higher when no-goal, compared with the non-shared goal condition - $H = 12.594$, $p = .028$, Cohen's $d = 1.251$, $r = .530$ (and marginally significant between the no-goal and the shared goal conditions - $H = 11.156$, $p = .064$) - but with no significant difference as a result of sharing or not sharing the goal (see Figure 3 below).
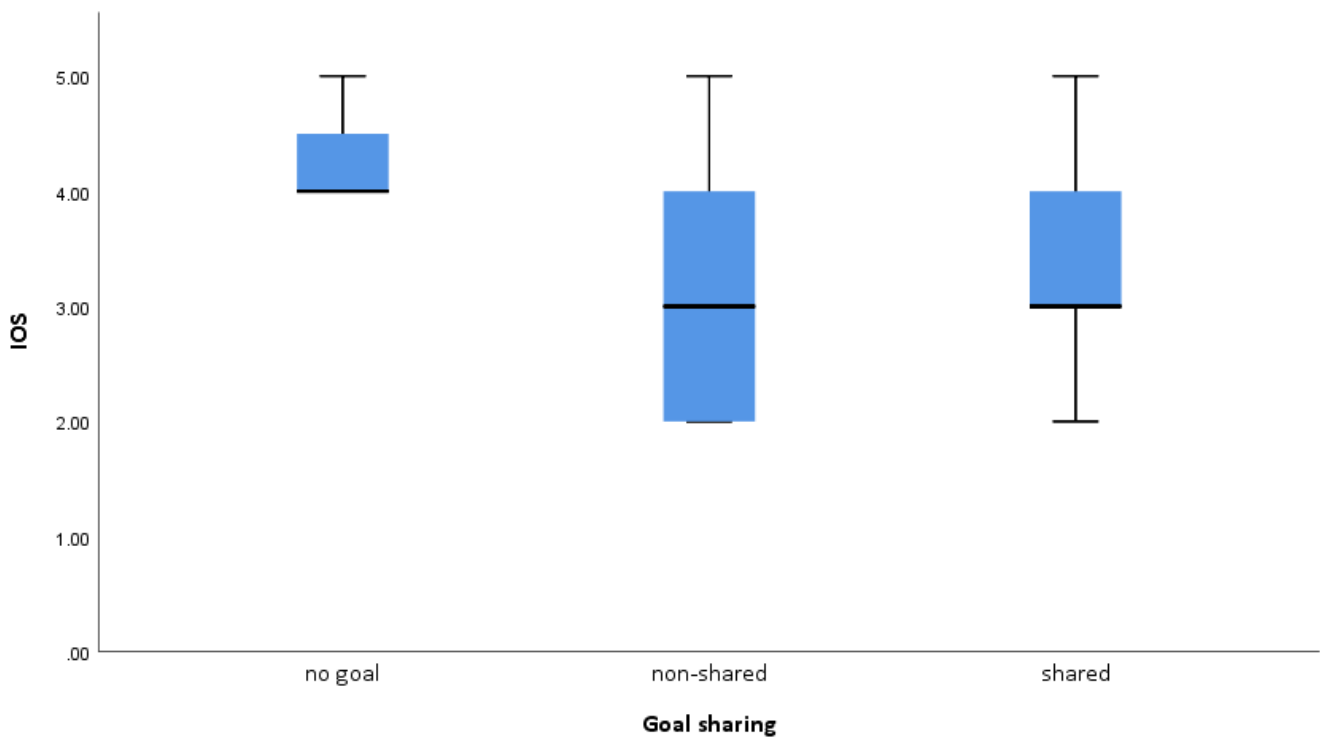
Figure 3 Boxplot showing median, interquartile range (blue shading), and top and bottom 25% of ratings of IOS ratings following different goal conditions. Ratings were significantly higher after the no-goal condition than the non-shared goal condition.

Perception of being 'on the same team'

A 2-tailed Kruskal-Wallis test showed a significant difference in ratings of same team perception depending on goal condition: H(2) = 6.770, p = .034. In this case, there were no significant pairwise comparisons after Bonferroni corrections. There was a trend towards higher ratings of team membership when participants shared than when they did not share a goal[114] and when they had no goal, compared to non-shared goal (see Figure 4 below).
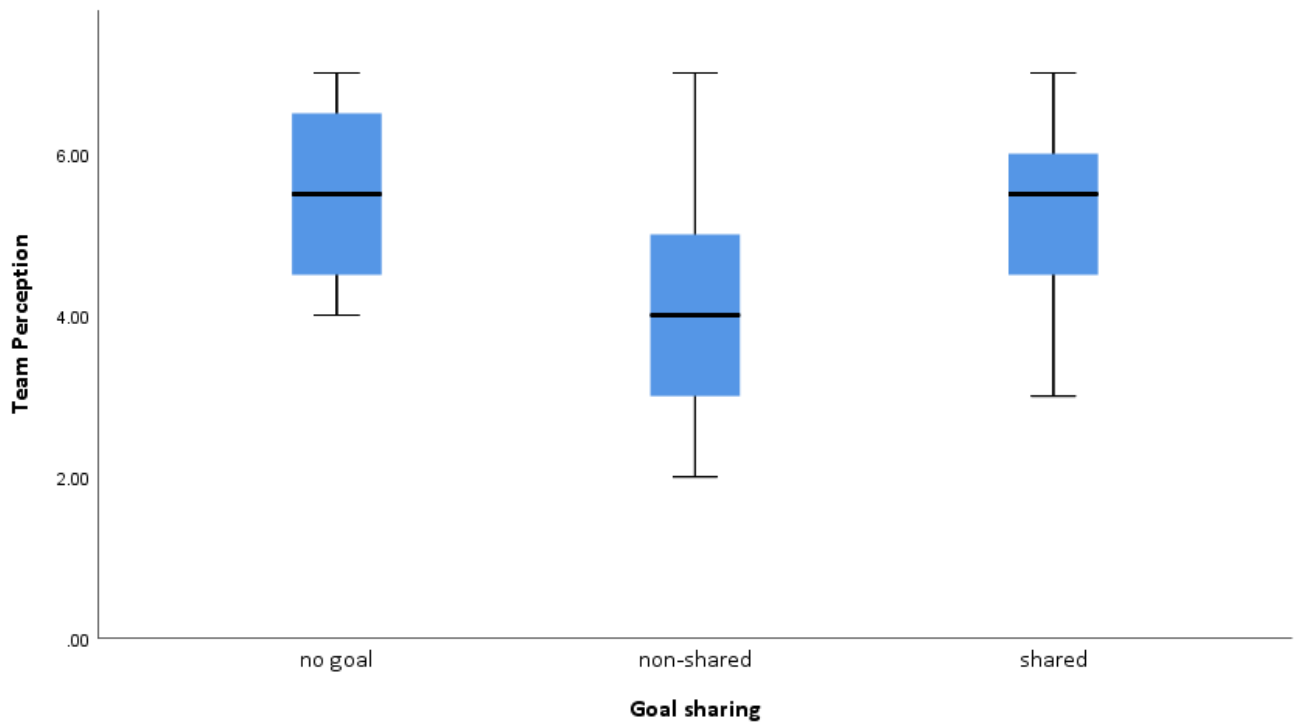

Figure 4 Boxplot with median, interquartile range (blue shading), and top and bottom 25% of ratings, showing effect of goal condition on the perception of being on the same team

Correlations between outcome measures

The following tables show Spearman's *rho* correlations for the dependent measures highlighted in the above analyses. Connectedness was related to other bonding measures, and also general indices of trust (especially IOS and confidence - Table 12).

---

[114] This was the pairwise comparison nearest to significance - H = 10.469, p = .074.

Table 12 *Significant non-parametric correlations with connectedness*

	IOS	Entitativity (interaction)	Team perception	Reliance	Confidence	General trust	No. social words
Connected-ness	.581 (p<.001)	.334 (p=.040)	.451 (p=.003)	.468 (p=.002)	.546 (p<.001)	.389 (p=.013)	.374 (p=.017)

Note: each box shows *rho* (p-value).

Connectedness did not correlate significantly with specific in-group trust preference, memory scores, interaction difficulty, performance, successfulness, or the other entitativity measures (similarity/entitativity).

IOS scores are related to several of the same measures as connectedness (including confidence and same team perception - Table 13). Perception of being on the same team also correlates with many of the same rating scales (Table 14).

Table 13 *Significant non-parametric correlations with IOS scores*

	Entitativity	Entitativity (interaction)	Team perception	Reliance	Confidence
IOS	.520 (p=.001)	.436 (p=.006)	.510 (p=.001)	.316 (p=.047)	.546 (p<.001)

Note: each box shows *rho* (p-value).

Table 14 *Significant non-parametric correlations with same team perception*

	IOS	Success-fulness	Entitat-ivity	Entitativity (interaction)	Reliance	Perfor-mance	Confid-ence	General trust
Same team perception	.510 (p=.001)	.436 (p=.005)	.406 (p=.009)	.493 (p=.002)	.468 (p=.002)	.457 (p=.003)	.492 (p=.001)	.385 (p=.014)

Note: each box shows *rho* (p-value).

Memory scores did not correlate with any other measure. Moreover, specific in-group trust preference scores correlated weakly with two components of the entitativity scale - similarity (rho = -.330, p = .04) and interaction (rho = .350, p = .034) - but with no other ratings (not even the general rating of trustworthiness).

There were no significant correlations between any of these scores and prior knowledge of participants in either the interaction group or the trust comparison out-group.

Changes in mood

The change in positive and negative PANAS scores pre- and post-interaction task was calculated. Ordinal logistic regression using mode, goal condition and task type as predictors was not significant for any change in negative mood. However, this model did significantly predict change in positive mood (Chi square = 9.594, df 4, p = .048, r^2 = .213 or .215) with goal condition (between no-goal and non-shared) as the significant component (Wald = 7.558, df 1, p = .006). This is also shown in Kruskal-Wallis independent-samples test: H(2) = 6.337, p = .042. No individual pairwise comparisons were significant after Bonferroni corrections, but there were trends towards a larger decrease in positive mood for no-goal compared to non-shared goal condition (H = -11.719, p = .060), and for no-goal compared to shared goal (H = -11.406, p = .070); there seemed to be no net change in positive mood for the shared/non-shared goal conditions.

Sex differences

Kruskal-Wallis independent-samples comparisons were run for all dependent measures according to sex. No differences were found except for ratings of difficulty of the task (higher ratings for male than female: H(1) = 4.495, p = .034), confidence in others (higher for female than male: H(1) = 7.212, p = .007), and for general trust (higher for female than male: H(1) = 6.537, p = .011). When included as a covariate in ordinal regression models (with mode, goal condition and task as predictors), sex did not have any effect.

Experience of the group interaction task - free responses

Comments by participants in the musical, shared-goal groups were largely to report a 'good experience', they 'worked well together', it being 'enjoyable' and 'fun', feeling 'comfortable' with lots of 'laughter'. Similarly, in the musical, non-shared groups, they reported lots of 'laughter', while 'sharing similar feelings' and 'communicating', with it being a bit more 'confusing'.

The non-musical, shared-goal groups emphasised the role of 'non-verbal communication' and 'comedy', 'working well together', completing the task easily and 'quickly', 'from instinct', but 'didn't find it enjoyable' in the same way as the music participants seemed to. Not sharing a goal had a bigger impact on the visual task: they noted that they didn't work well, it was 'confusing' as they 'worked against each other', with 'saboteurs' in the group; they commented that the experience was 'very awkward' and confusing until they realised they had different goals. These responses reveal that the effect of sharing or not sharing a goal was more important in the visual than in the music task.

179

The musical, no-goal groups reported an experience similar to the other music groups: 'quite amusing', with laughter, while not working very well together as they had no goal to focus on. However, they still seemed to bond as a group: becoming 'less shy' as the task progressed.

Summary of results

Table 15 gives a summary of the main quantitative findings from my statistical analysis. The free responses corroborated the effect of mode on enjoyment and sense of bonding, with more reported fun/laughter etc. in the music condition, regardless of goal condition.

Table 15 *Summary of data analysis*

Dependent measure	Mode condition comparison	Goal condition comparison
Connectedness	Music > visual **	
Memory scores	Visual > music **	Mode*goal sharing ** visual – shared > non-shared music – non-shared > shared
IOS ratings		No goal > non-shared ** No goal > shared goal
Same team perception		Shared > non-shared No goal > non-shared
Decrease in positive mood		No goal > shared No goal > non-shared

Note: asterisks (**) indicate statistical significance.

Discussion

In this experiment, I manipulated interaction mode (music vs. visual) and goal conditions (shared vs. non-shared vs. no goal) and measured how far each of these affected group formation, using a variety of indices of bonding. In line with hypothesis (a), there was a clear effect of mode, with higher ratings of connectedness to other group members following the music interaction task, in comparison with the non-music, visual condition. Connectedness was correlated with other items pertaining to general bonding with the group, including closeness scores (IOS), reliance on and confidence in the others, general trustworthiness and the interaction pictorial scale (the predictor for dynamic entitativity - see Rutchick et al., 2008). Connectedness ratings have previously been found to increase following synchronous walking (Wiltermuth and Heath, 2009 - study 1); here, this measure is applied to a context of group music-making, providing an indication of group formation or bonding as a result.

This effect of mode on connectedness did not seem to be as a result of increased attentiveness towards others; the memory scores show that the opposite is the case i.e. higher memory recall for others under the visual than the musical condition. These findings contrast with the results of attentiveness and memory conducted in dyads (e.g. Woolhouse et al., 2016). It might be the case that, in the context of group musical interaction, attention is not directed towards specific others, but is a more general awareness of the collective (compare the idea of a central 'ear' of a quartet, operationalised in Glowinski et al., 2013). The effect of mode on connectedness also does not seem to be related to any increase in positive mood as a result of the musical interaction task, measured using self-report PANAS pre-/post-interaction ratings. This contrasts with previous work linking collective effervescence or arousal to group bonding (see work on emotion contagion work by Spoor & Kelly, 2002, and Barsade, 2002; however, they use video coding of emotional behaviour as well as the PANAS).

Neither hypotheses (b) nor (c) were supported by this experiment. There were no effects of the goal conditions on connectedness ratings, nor on any of the main bonding measures. In fact, the IOS scores indicated that the highest ratings of self-group overlap (or closeness) occurred following the musical, no-goal condition, with no apparent difference between shared and non-shared goal conditions for either modality. The ratings of 'same team' perception showed a marginally significant effect of goal conditions, with a trend towards higher ratings following shared than non-shared goal, but no pairwise comparisons were significant. There were also no effects of these conditions on mood, apart from a trend towards a decrease in positive mood following the no-goal condition.

Limitations

The main limitations of this experiment concerned the sensitivity of the measures we used. There was no effect of the presence of shared (or any) goals with the self-report, quantitative measures; however, the free responses (see above) indicate something that might not have been detected in these ratings. A more subtle effect might have been revealed through closer analysis of the nature of the interaction itself - for example, measuring moment-by-moment alignment in video analysis.[115] This has been studied in conversational contexts, with different degrees of multimodal alignment depending on whether partners agree or conflict on certain issues (Paxton & Dale, 2013).

Similarly, music-making might have had an effect on mood if we used more sensitive (but time-consuming) measures. Other studies showing the emotional or affective significance of music listening or performance use indirect methods such as pain threshold (as a proxy for endorphin levels

---

[115] See an example of this measurement in my video analysis (reported later), where I analyse the effect of presence or absence of an interaction goal. I could use a similar methodology in future research to compare interaction involving shared or non-shared goals.

- see Dunbar et al., 2012), or activation of reward centres in the brain (e.g. Blood & Zatorre, 2001). These might tap into more general reward or arousal processes, rather than sensations which can be self-reported more explicitly (as in the PANAS measure we used here).

Our interpretations about participants' attention might not be reliable, given a couple of problems with the accuracy of the memory scores.[116] First, higher memory scores under the non-music condition could have been due to more focus on group members' hands during the visual activity (which involved moving shapes on a table), leading to more accurate recall of wristbands. Second, the other interpersonal memory items were problematic: it is difficult to assess what counts as blonde hair, and pictures of group members had been displayed already in the trust measure (so participants might not have been completing these questions from memory of the interaction, but rather from earlier in the questionnaire; see footnote 102, p. 161).

There are a number of possible reasons for the lack of any effect (of either mode or goal condition) on the specific in-group trust measure. The forced-choice task involved choosing between specific individuals; however, any sense of group formation might have been related to more of a collective focus, a more general bonding to the group as a whole (see suggestion above). Moreover, trust might be an inappropriate outcome for this reason: deciding whether or not to trust someone relies on knowledge and perception of another individual, rather than any affiliation to a group.[117] Alternative explanations include reputation effects - choosing between specific others is a potentially costly task; subsequent investigation of this might emphasise more the anonymity of responses.

Finally, the free responses highlighted a greater incidence of general group laughter in the musical compared to the visual condition; this behaviour might have contributed to social bonding (see Gervais & Wilson, 2005). A follow-up experiment (see Appendix 2) was conducted to ensure that the effect of mode on connectedness was not just as a result of the humorous nature of the musical task.

Experiment 2 - music in team formation

This experiment is an investigation of how *team* affiliation might be affected by properties of the interaction. In particular, I tested the efficacy of musical goals in contributing to team identity (or affiliation on the basis of team membership). I compared two goal types, corresponding to two team names, with a no-goal condition. I hypothesised higher levels of in-group preference (on the basis of team membership) following musical interaction with a goal than a no-goal task.

---

[116] Moreover, the main effect of mode emerged only in the more powerful parametric test, treating the memory scores as interval data.

[117] Group membership can contribute to person perception in this way, but probably not in the short time-span involved in this experiment.

Method

Participants

A total of 20 undergraduate students (16 male and 4 female) took part in this experiment. All participants had a high level of musical training (they were recruited through the music faculty). They were aged 17-26 years (mean 18.7).[118] They all gave informed consent to take part in the experiment, which was reviewed by the Faculty of Music Research Ethics committee.

Design

This comprised a between-groups one-way design, varying the interaction goal (goal or no-goal). Ten participants were assigned to the no-goal condition; they were split into two teams (elephant and bird), but then interacted in a no-goal music task. Another ten participants were assigned to the goal condition; they were also split into the elephant and the bird team, but then interacted with the goal to depict a herd of elephants or a flock of birds (corresponding to their team identity).   The distribution of these goal conditions is shown in Table 16 below.

Table 16 *Number of participants assigned to each team, and to each interaction goal condition*

Team identity (badge)	Interaction goal	No. participants
Elephant	No goal	5
Bird		5
Elephant	Elephant	5
Bird	Bird	5

All participants in the no-goal condition were male; interaction conditions with a goal (bird or elephant) comprised mixed-sex groups.[119] All interaction groups comprised members of both teams (see Table 17).

---

[118] Although all participants were undergraduates, one was slightly younger (aged 17 years). They gave their informed consent in the same way as all the other participants.
[119] This difference in sex composition was a potential confound - any differences caused as a result of goal presence might have been attributable to sex differences. However, as there were no significant effects of sex in experiment 1, this was not likely to be problematic.

Table 17 *Tasks assigned to each interaction group*

Interaction group	Interaction goal	Team affiliations	No. participants
1	No goal	Elephant	2
		Bird	3
2	No goal	Elephant	3
		Bird	2
3	Elephant	Elephant	2
	Bird	Bird	3
4	Elephant	Elephant	3
	Bird	Bird	2

Note: there were two groups in the no-goal session (comprising members of both teams). There were two groups with a goal condition; participants in each group had goals corresponding to their team identity (thus each group involved non-shared goals).

This experiment had two dependent variables - group bonding (within the interaction group i.e. mixed teams) and team identification (preference for team members, both from within and outside the interaction group). These were administered via a post-interaction questionnaire. Group bonding measures included ratings of connectedness for members of the interaction group (especially other-team members), ratings of interaction success and entitativity. Team identification measures included specific trust preference, the difference in connectedness (between ratings for own-team versus other-team members), and an IOS scale of overlap with the team as a whole (see Appendix 1.2 for full details of measures used). The short form of the PANAS mood questionnaire was administered pre- and post- interaction task (Kercher, 1992; reviewed by Mackinnon et al., 1999). Participants were also asked to rate (and detail, if necessary) how much they knew other members of their own and the other team prior to the experiment (7-point scales).

Procedure and apparatus

Photographs for the trust questionnaire had been taken, with consent, the previous day. Participants were tested in groups of ten - one session for the goal condition, and the other the no-goal condition.

1) Minimal team formation

Participants were first divided into two separate teams (five participants in the bird team, and five in the elephant team). These teams were sat separately (in two circles, both teams in the same room).

Participants were each given a badge to signify their bird or elephant team identity with the team symbol (see Figure 5 below).

Figure 5 Elephant and bird team symbols

Here, they signed the consent form, completed the pre-task PANAS and then read their individual instructions. Participants were told that they were part of 'an improvisation competition at the music faculty'; each of them was a member of one of two teams in this competition.

2) Interaction conditions

Participants were instructed to move to one of two locations in the room, where they were each given a kazoo and told to:

interact with each other, and improvise together, using only these kazoos.

They were asked to complete the task without talking to one another, interacting and improvising only within (and not between) their assigned group.

Participants in the no-goal condition were given no additional instructions. Those in the goal condition given a depiction task, corresponding to their team identity:

interact with each other...to perform music inspired by a herd of elephants/a flock of birds

Each interaction group involved members of both teams (see above) therefore the groups with a goal condition had non-shared interaction goals. All groups interacted for 10 minutes.

Participants were told that they could earn points for their own team, being 'assessed on how well you listen and co-ordinate with the others.' Video cameras were set up for each group to reinforce the sense of competition.

After this task, participants were seated individually to complete the questionnaire. After this, they were led to a separate area to be debriefed about the design and purpose of the experiment.

Data analysis

Goal was the main predictor variable. Comparisons were run between goal and no-goal conditions; also the goal conditions were separated to form three levels - the bird goal, elephant goal, and the no-goal condition.

Items on the questionnaire measured either team formation (bonding to individuals on the basis of shared team membership, from within or outside the interaction group) or group formation (bonding to individuals within the interaction group regardless of team membership). Trust scores were formed by totalling the number of own-team choices for all scenarios (for familiar team members from the interaction task, for unfamiliar team members, and a combined score). Connectedness ratings comprised three different scores: towards own-team members, other-team members, and the difference (how much *more* connected participants felt towards own-team members within their interaction group).

Kruskal-Wallis independent-samples non-parametric tests were run on all dependent measures. All post-hoc tests include Bonferroni corrections for multiple comparisons. Non-parametric correlations were run to test the relationship between team-focussed (e.g. IOS) and group-focussed (e.g. group success) items.

Results

Baseline differences

    1)   Prior knowledge of others

Kruskal-Wallis comparisons for prior knowledge of other participants (own-team and other-team members) were run. These revealed significant differences between goal conditions. The difference in prior knowledge of own-team members was significant: $H(2) = 6.984$, $p = .030$; participants in the bird goal condition knew their own team better than those in the no-goal condition ($H = -7.650$, $p = .036$). There were also significant differences in prior knowledge of other-team members: $H(2) = 8.933$, $p = .011$; again, those in the bird goal condition gave higher ratings than those in the no-goal condition ($H = -8.350$, $p = .010$).

One participant (in the bird goal condition) reported much higher ratings of prior knowledge for members of their own and other-team members. Kruskal-Wallis comparisons were re-run, excluding this participant. There was still a significant difference in prior knowledge of members of the other team: $H(2) = 7.286$, $p = .026$ (again, higher ratings in bird goal than no-goal condition: $H = -7.275$, $p = .032$). However, there was no longer a significant difference between the goal conditions for prior

knowledge of members of participants' *own* team.[120] The following statistics reported were run on this smaller dataset (n = 19).

2) PANAS mood scores

Kruskal-Wallis comparisons showed no difference between the groups according to goal type for baseline positive and negative mood, measured by the PANAS. Moreover, there was no effect of goal type on changes in negative or in positive mood pre- and post-interaction task.

3) Sex differences

As noted earlier, there were 10 male participants in the no-goal condition, and a mix in the goal conditions (2 female, 2 male in bird condition, and 2 female, 3 male in elephant condition). Therefore, there were baseline differences between these conditions.

Ratings of connectedness (towards others within interaction group)

1) Towards own-team members

Kruskal-Wallis independent-samples comparisons showed no significant effect of goal presence, or goal type, on ratings of connectedness with participants' own team members.

2) Towards other-team members

However, there was a significant effect of goal presence on connectedness towards members of the other team: there were higher ratings following the no-goal than the goal conditions: H(1) = 9.444, p = .002, Cohen's *d* = 1.961, *r* = .700.[121]

---

[120] This rating, and not the prior knowledge of members of the other team, might confound the team identification results if significant. On the contrary, if there is a higher rating of prior knowledge of members of the other team, this is actually likely to act in the opposite direction i.e. to reduce any potential own-team preference (unless, of course, their previous encounter with the out-group was a very negative experience).
[121] Observed power = .975

Separating goals into bird and elephant depiction tasks, there was again a significant effect of goal type: $H(2) = 9.506$, $p = .009$. In particular, this was higher following the no-goal than the elephant goal ($H = 8.000$, $p = .020$) and nearing significance when compared to the bird goal ($H = 7.100$, $p = .077$) (see Figure 6 below).
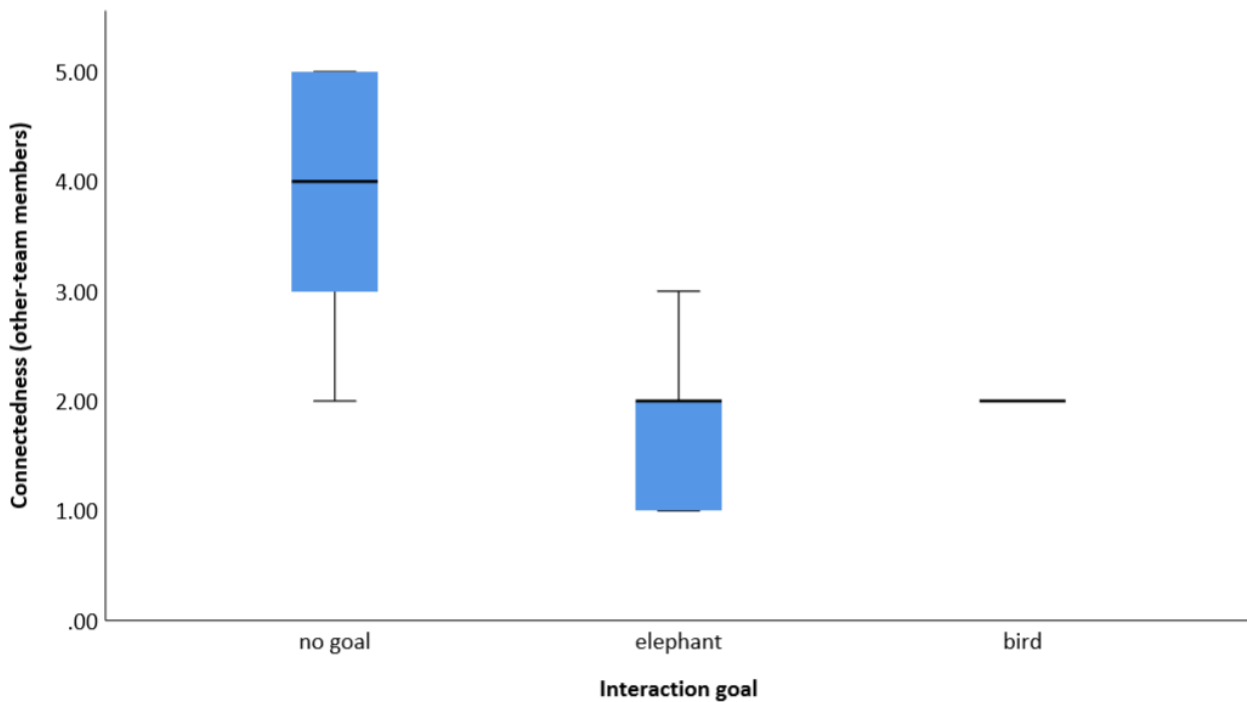

Figure 6 Boxplot (with median, interquartile range - blue shading - and top and bottom 25% of ratings) showing effect of interaction goal on connectedness towards members of the other team. Ratings of connectedness were significantly higher following the no-goal than the goal conditions (specifically, the elephant depiction task)

3) Towards own-team members *more* (difference between own and other-team ratings)

A Kruskal-Wallis test revealed that participants felt significantly more connected to their own-team members (than the other-team members, all within the interaction group) following the goal conditions than the no-goal condition: $H(1) = 5.507$, $p = .019$, Cohen's $d = .737$, $r = .346$.[122] Again, splitting up the goal condition revealed a significant main effect of goal type: $H(2) = 7.683$, $p = .021$. There was a larger difference in the bird condition than in the no-goal condition ($H = -8.925$, $p = .018$) (see Figure 7).
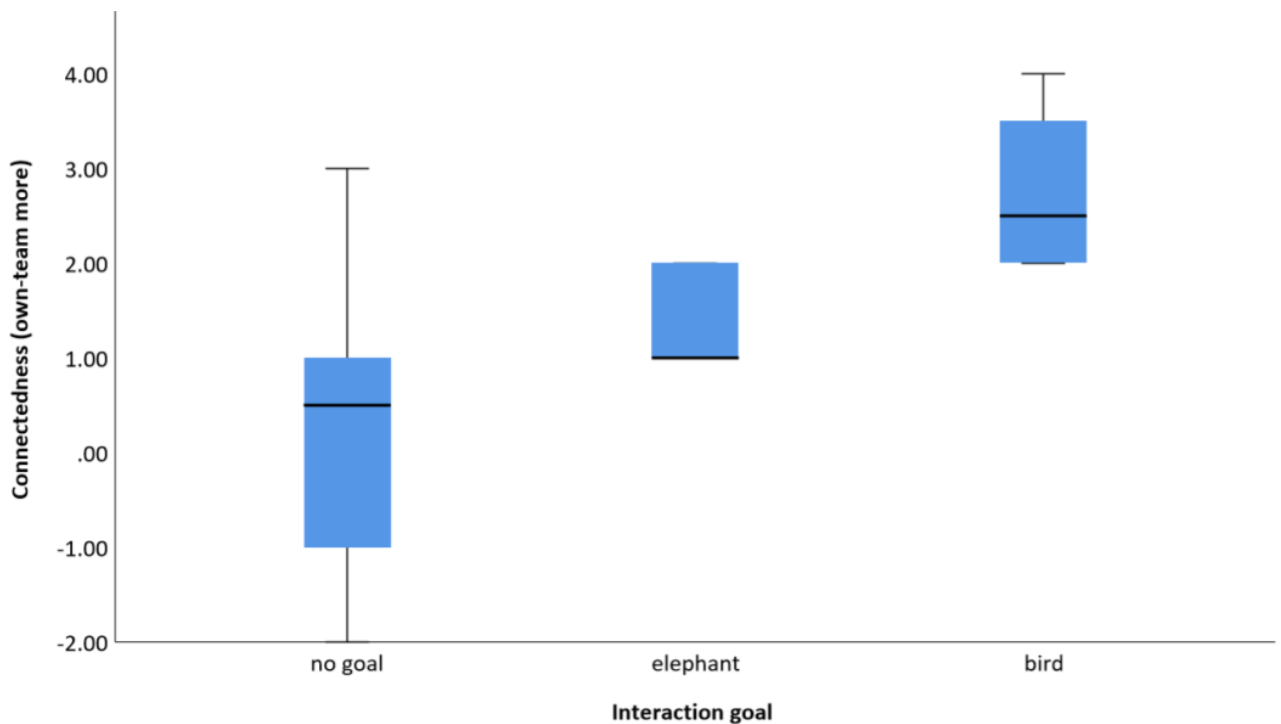
Figure 7 Boxplot showing effect of interaction goal on how much more connected participants felt towards members of their own team than members of the other team (within their interaction group). This is significantly higher following the bird task than the no-goal interaction (plot shows median, interquartile range, and top and bottom 25% of ratings).

In line with this, there were marginally significant effects of sex. Specifically, there were higher ratings of connectedness towards other-team members amongst male than female participants ($H = 3.946$, $p = .047$) and towards own-team members amongst female than male members (nearing significance: $H = 3.804$, $p = .051$). This corresponds with the sex imbalance between the goal and no-goal conditions (see above).

---

[122] Observed power = .314

IOS ratings (towards the team as a whole)

Kruskal-Wallis comparisons revealed a significant effect of goal type on IOS ratings towards their team identity: H(2) = 6.788, p = .034. As above, post-hoc comparisons showed a significantly higher team IOS following the bird goal compared to no-goal condition (H = -8.300, p=.031, Cohen's *d* = 1.933, *r* = .695) (see Figure 8). There were no effects of sex on IOS scores.
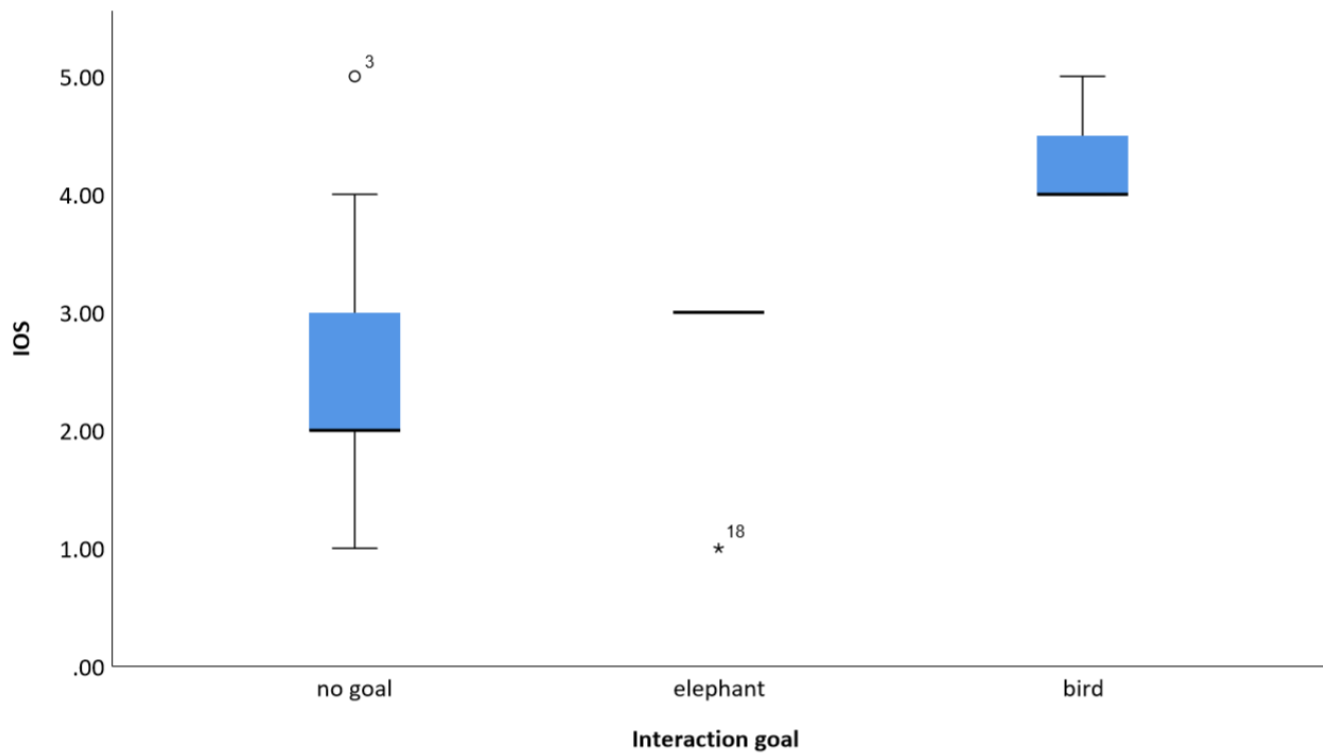

Figure 8 Boxplot showing effect of interaction goal on rating of IOS towards team as a whole. Participants rated closeness to team (IOS) significantly higher following the bird goal condition than the no-goal task (plot shows median, interquartile range, top and bottom 25% of ratings, and outliers).

Group success

There was a marginal effect of goal type on success shown in a Kruskal-Wallis test: $H(2) = 5.270$, $p = .072$. This indicated a trend for higher success ratings in the bird than the elephant condition (but those in the no-goal condition spanned the whole range of scores) (see Figure 9).
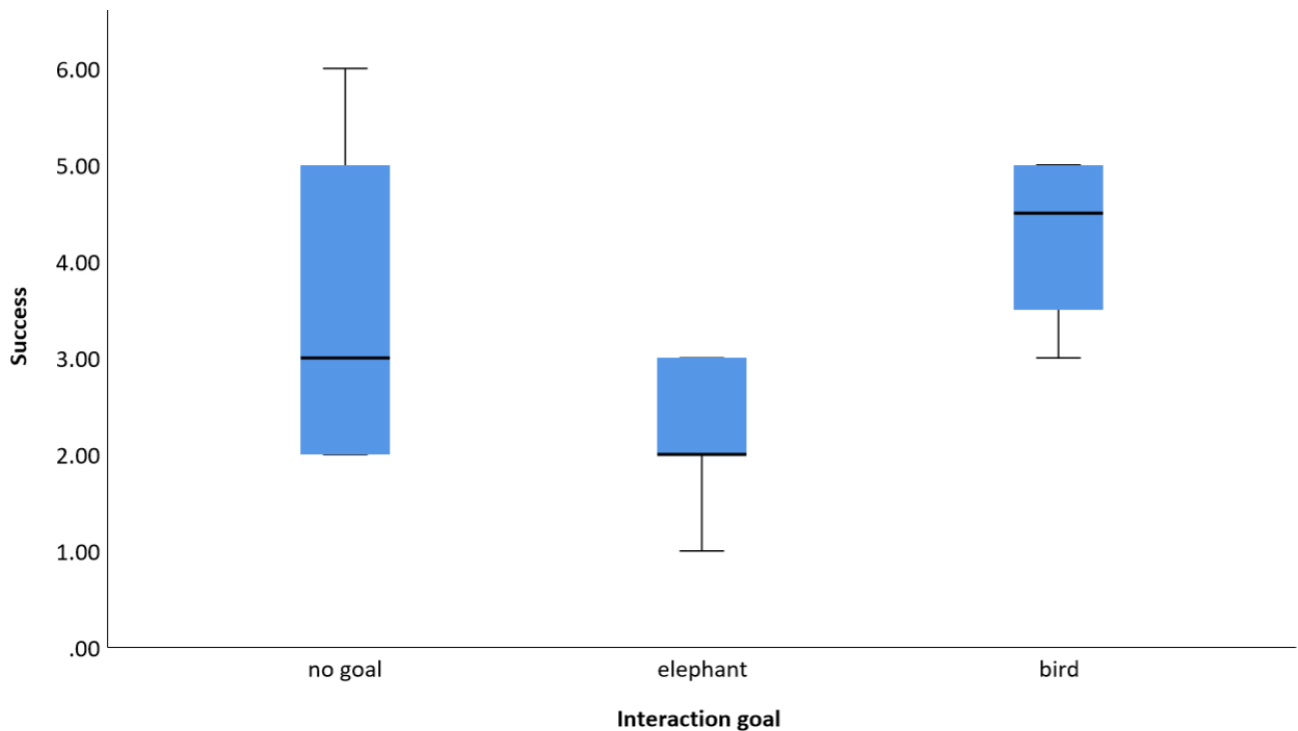

Figure 9 Boxplot showing effect of interaction goal on ratings of group interaction success (plot shows median, interquartile range, and top and bottom 25% of ratings).

Ratings of group interaction success were correlated with ratings of connectedness towards own-team members: Spearman's *rho* = .648, $p = .003$ (see Figure 10 below).[123] Group success was also correlated with the pictorial scale of entitativity (*rho* = .502, $p = .029$), and with the baseline measure of positive mood (*rho* = .509, $p = .031$), but with no other measure.

---

[123] This correlation was found only for connectedness to own-team members, not other-team members, or the difference.
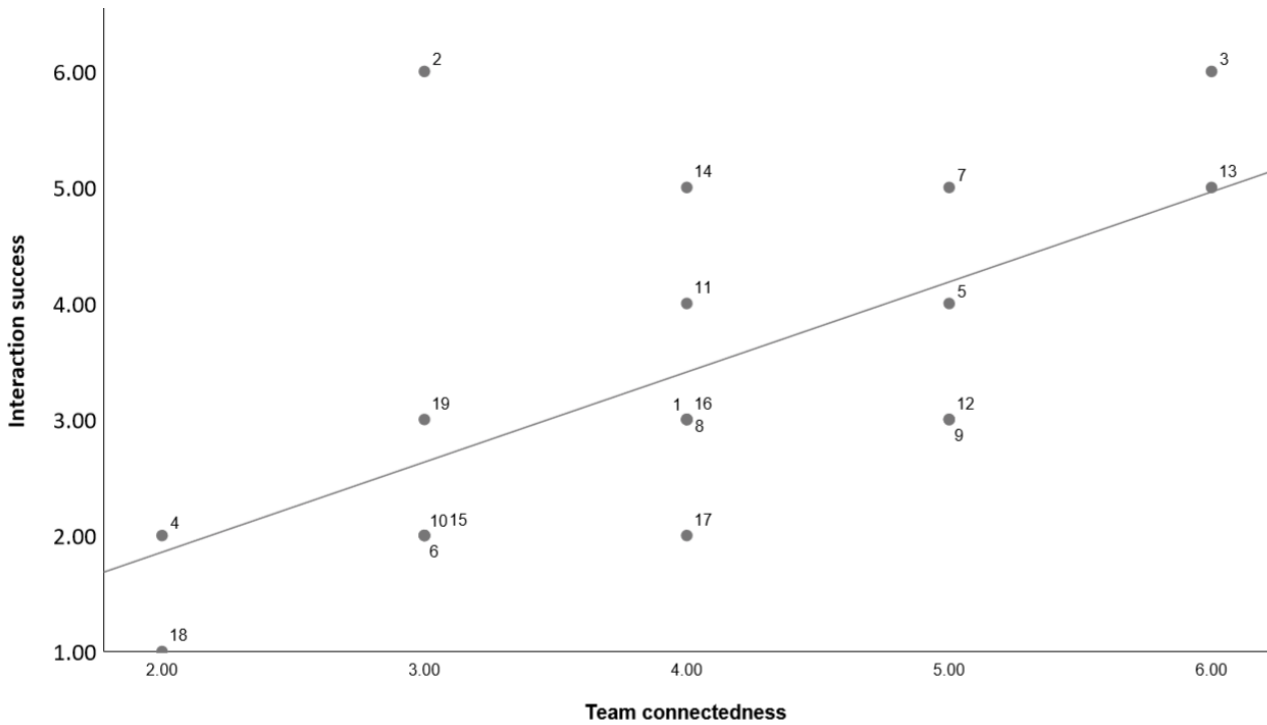
Figure 10 Significant non-parametric correlation between connectedness towards own-team members (who were also members of the interaction group) and ratings of group interaction success.

Other ratings

Kruskal-Wallis comparisons revealed no effect of goal type on the specific trust questionnaire (whether just familiar or unfamiliar individuals, or for the combined score). There were also no effects of goal type on any of the pictorial entitativity measures.

Baseline positive mood correlated with trust towards unfamiliar others (*rho* = .631, p = .005), the total trust score (*rho* = .595, p = .009), as well as with group success ratings (see above).

As in experiment 1, ratings of connectedness[124] were correlated with general trust (towards unfamiliar others) and closeness to the team (see Table 18).

Table 18 *Non-parametric correlations between own-team connectedness, trust and IOS ratings*

Measure	Specific trust (unfamiliar others on same team)	IOS (towards team as a whole)
Connectedness (to own-team members)	.503 (p = .028)	.484 (p = .036)

 Note: each box shows *rho* (p-value).

---

[124] Again, this refers to the ratings for own-team members (not connectedness to other-team members, or the difference).

192

Experience of the group interaction task - free responses

Participants in the no-goal condition largely seemed to focus on the group interaction and communication between group members; a few commented that they didn't notice the team differentiation, but interacted equally.

Participants in the goal conditions focussed less on group communication, but instead highlighted the difficulty of the task (especially participants given the elephant goal condition).

Responses from some participants (across goal conditions) did not make a clear distinction between group and team identities, using 'teamwork' to describe the group interaction.

Summary of results

Table 19 provides a summary of the results of experiment 2.

Table 19 *Summary of data analysis*

Dependent measure	Goal condition comparison	Goal type comparison
Connectedness (other-team)	No goal > goal **	No goal > elephant ** No goal > bird
Connectedness (own-team more)	Goal > no goal **	Bird > no goal **
IOS		Bird > no goal **
Success		Bird > elephant

Note: asterisks (**) indicate statistical significance. The relative difficulty of the elephant condition was corroborated by the free responses.

Discussion

In this experiment, I distinguished between group bonding and team identification. The latter was established by perceptual differences in bird and elephant identities (using badges). I then measured the effect of having an interaction goal - comparing the bird and elephant depiction conditions with a no-goal music-making task. The group interaction tasks (in all conditions) deliberately crossed teams, forming a group membership that was somewhat distinct from team identity.

In line with the hypothesis, participants with an interaction goal that corresponded to their team membership felt more connected to other members of their own team than those on the other team (within their interaction group), compared to participants without such a goal. This effect was carried by those who had the bird goal task (ratings of connectedness following the elephant goal lay in between those from the bird and no-goal conditions). Similarly, those in the bird goal condition felt more self-other overlap with the wider team membership (higher IOS scores) than participants in the

no-goal task (again, the elephant scores lay somewhere in between). Therefore, an interaction goal - an external form of common ground, an element of a more extrinsically-oriented encounter - supports team identification i.e. affiliation with on the basis of team membership, which can go beyond the face-to-face group context.

However, this effect was only shown by participants in the bird goal condition, not those with the elephant goal. This might be related to the trend towards higher perceived success for the bird goal, and the difficulty of the elephant goal task as reported via the free responses: only the goal which the participants felt was more successful contributed to team identity. In line with this, ratings of the group interaction success were correlated with one index of team affiliation. These findings provide evidence for the direct role of interpersonal interaction on the formation of team-based affiliation. In particular, the correlation between team identity and perceived group success aligns with the idea (part of Social Identity Theory) that our self-esteem is tied to our social (as well as personal) identities.

On the other hand, participants who interacted in a no-goal, music-making task felt significantly more connected to members of the opposing team (in their interaction group) than participants who had an interaction goal which corresponded to their team identity. This highlights the potential for music-making (especially of a more intrinsic orientation - emphasising its phatic, relational attributes) to provide a context for social bonding which overrides team (or other category) divides (replicating e.g. Gaertner et al., 1993: see Shifting Goals section, Chapter 3, on processes of de- and re-categorisation).

Again, a sense of connectedness was correlated with more general social affiliation - in this case, trust towards unfamiliar individuals, and sense of overlap with team identity as a whole; this is similar to the correlations found in experiment 1 between connectedness and general trustworthiness. Thus, connectedness seems to be an important index of relationships, but one that is not directed towards specific individuals on the basis of knowledge or prior experience of that person. Rather, it seems to be related to general affiliation towards the group or team membership as a whole.

Limitations

One limitation of the present experiment is the gender imbalance between the goal and the no-goal conditions, which might confound the effect on connectedness. However, this is not likely to be problematic, given previous findings. In general, male participants are likely to have a stronger in-group preference (see e.g. Van Vugt et al., 2007). However, here, male participants rated higher connectedness towards other-team members than did female participants, whereas female

participants rated higher bonding for the own-team members. This is more likely to be an effect of the goal conditions. Moreover, differences between the effects of the two depiction goals would not be confounded by sex (as both of these goal conditions had a gender mix).

Prior knowledge of other participants might also limit the validity of these findings. This should not affect the results too dramatically: the participant with the highest ratings was removed from the dataset, leaving only a significant difference in prior knowledge of *other*-team members. However, the experiment should be replicated, with a larger sample, to ensure that this did not confound findings relating to team affiliation.

Finally, the participants might have experienced some confusion regarding the group/team distinction within the experimental design - some of the free responses conflate the terms 'group' and 'team'. However, this is not likely to be important for the present findings. Often, the connectedness ratings were calculated as an explicit comparison between those who shared and did not share team membership (within the interaction groups); the IOS scores were also clearly associated with the team symbol (and the use of the terms 'group' and 'team' might have been used loosely in the free responses).

Experiment 3 - music in team identity

This experiment built upon previous findings to investigate further whether different attributes of the musical interaction - presence of a depiction goal, and perceived interaction success - contribute selectively to interpersonal bonding or a sense of team identity. Based on earlier results, I hypothesised the following:

(a) Higher interpersonal bonding following the no-goal (when the interaction is more intrinsically-focussed, privileging the phatic) than the goal conditions (where participants are focussed on an external depiction goal)

(b) Higher team identification (higher recall of team-related words) following interaction with a team-related goal that was found to be more successful, in comparison with a less successful goal (or a no-goal task) i.e. higher following the bird goal than after the elephant or no-goal conditions.

Method

Participants

A total of 36 participants took part in the experiment.[125] They were aged between 17 and 25 years (mean = 20.89).[126] There was a mixture of male and female participants in each condition (7 female in the bird goal, 8 in elephant, and 9 in no-goal condition). Participants interacted in same- or mixed-sex pairs (bird goal: 2 female, 1 male, 3 mixed; elephant goal: 2 female, 4 mixed; no-goal: 3 female, 3 mixed). They were recruited from across the University. They all gave informed consent to take part in the experiment, which was reviewed by the Faculty of Music Research Ethics committee.

Design

This experiment comprised a between-groups design, varying interaction goal: no goal, bird goal (easy goal) and elephant goal (more difficult goal). Where participants were assigned to a goal condition, these corresponded to their team identity (see Table 20).

Table 20 *Team identity and interaction goal assigned to participants*

Team identity	Interaction goal	No. participants
Elephant	No goal	6
Bird		6
Elephant	Elephant	12
Bird	Bird	12

 Note: n = 12 in each goal condition.

There were two dependent variables: team identification (measured using both implicit and explicit indices) and bonding with interaction partner (self-reported affiliation; partners shared team membership).

Procedure

1) Minimal team formation

Participants were tested in groups of 4. They were first split into two teams (2 in the "bird" team, and 2 in the "elephant" team). The experiment was introduced as a competition between these teams; they were told that they were going to do a couple of activities to earn points for their team. They were asked to put on a badge, showing their team symbol. This was similar to the 'minimal' team

---

[125] This was based on a pilot study. A priori calculations predicted power = .95 for n = 36.
[126] As in experiment 2, all were students at the University.

formation phase conducted by e.g. Good et al. (2017) and Tunçgenç & Cohen (2016). Participants were never in a group where there was a gender divide between teams.

2) Interaction conditions

Participants first completed an interaction task, supposedly to earn points for their team (video cameras were set up - not recording - to reinforce this). This task was completed with their partner (who shared the same team identity). Both pairs of participants completed this task simultaneously, in neighbouring rooms (the rooms were counterbalanced across conditions). They could not see the other pair, but they could hear what they were doing.

All participants were each given a simple musical instrument (a kazoo) and told that they could interact with their partner using only this kazoo (without talking or otherwise communicating). It was expected, from pilot experiments, that the bird condition would be easier than the elephant condition. Participants engaged in the interaction task for five minutes.

Easy goal condition – participants allocated to the bird condition were given the following task:

> make music inspired by a flock of birds, using only the kazoos. Improvise together, listening and co-ordinating with each other. In your joint improvisation, you could allude to e.g. a variety of bird calls, their manner and/or speed of movement, interaction with potential mates, wings flapping and their protection of territory, amongst other things.

Hard goal condition – participants allocated to the elephant condition were given the following task:

> make music inspired by a herd of elephants, using only the kazoos. Improvise together, listening and co-ordinating with each other. In your joint improvisation, you could allude to e.g. trumpeting sounds, their manner and/or speed of walking, interactions with their calves, ears flapping, and the sound of drinking/spraying water, amongst other things.

No-goal condition – participants allocated to the no-goal condition (half in the bird team, and the other half in the elephant team) were given the following task:

> make music, using only the kazoos. Improvise together, listening and co-ordinating with each other.

3) Implicit measure of team identification

Immediately after completing the interaction task, each participant was seated individually and given a word recall task (again allegedly to earn points for their team). Participants were presented with a written list of 22 words. Half were related to elephants (bull, herd, tusk, calf, trunk, ivory, mammal, leather, trumpet, mammoth and stampede) and half to birds (beak, eggs, wing, song, flock, chick,

aviary, feather, plumage, nesting and migratory). These were presented alternately, and in counterbalanced order; the bird and elephant lists were matched as far as possible in word length and familiarity. Participants were given 20 seconds to study the words, and then two minutes for free, written recall.

4) Explicit self-report measures

Participants were then asked to complete a questionnaire. This contained measures of team identification (IOS towards team as a whole, reward allocation to unfamiliar individuals associated with the team) and interpersonal bonding (trust, connectedness, similarity and likeability) towards their partner. Participants were asked to rate their need to belong (using the single-item test – see Nichols & Webster, 2013). They were also asked about their perceived performance on the task (difficulty ratings, and how well they think they performed individually), how successful they thought they and their partner were at the interaction task, and how embarrassed they felt. (See Appendix 1.3 for full details.) The short form of the PANAS mood questionnaire was administered pre- and post-interaction task (Kercher, 1992; reviewed by Mackinnon et al., 1999). Participants were also asked to rate (and detail, if necessary) how much they knew their team partner prior to the study (7-point scale), and whether they recognised any individuals pictured in the reward allocation task. They were asked to specify how many years of musical training they have had in the past, whether they are practising musicians of any sort now, and - if so - how regularly (on a 5-point scale: monthly, fortnightly, weekly, more than once a week, every day). After they completed this questionnaire, participants were thanked, debriefed and paid for their participation.

Data analysis

To test the effect of interaction goal on partner affiliation, independent-samples Kruskal-Wallis comparisons were run on the self-report affiliation ratings (ordinal data including connectedness, IOS, trust, likeability) according to interaction condition (no-goal, elephant goal, bird goal).

Before investigating the effect of perceived interaction success, Kruskal-Wallis comparisons were first run as a manipulation check, to test whether the bird and elephant goal conditions varied systematically in difficulty (using ratings of interaction success, difficulty, performance, and embarrassment). However, there were no significant differences according to interaction goal. Therefore, the effect of perceived interaction success on implicit team affiliation was tested across all goal conditions using linear regression (or ordinal regression for the IOS scores). Word recall was measured as a percentage of own-team related words out of total recall (including some words that were remembered correctly but with minor mistakes e.g. wrongly pluralised, 'sing' instead of 'song'). Own-team reward allocations were totalled and calculated as a percentage of total reward (excluding

any comparisons where participants claimed to recognise individuals in the photographs). Correlational analyses were run on the dependent measures to test the distinction between affiliation-related and team-related ratings.

Results

Baseline measures

Kruskal-Wallis comparisons were run to check whether there was any difference in baseline measures across interaction conditions. There was no difference in Need to Belong, the number of years of musical training, the frequency of musical activities, or in prior negative mood. However, there was a significant difference in positive mood (from the PANAS self-report test prior to interaction): $H(2) = 8.123$, $p = .017$; pairwise comparisons no goal > bird goal condition ($H = 11.167$, $p = .027$).

Kruskal-Wallis comparisons were run after excluding four participants (two with the lowest mood in the bird goal condition, and two with the highest mood in the no-goal condition). There was no longer a significant difference in baseline positive mood according to interaction condition. The rest of this analysis was run using this smaller dataset ($n = 32$).

Effect of interaction condition

Partner affiliation and trust

Kruskal-Wallis comparisons according to interaction condition (bird goal, elephant goal and no goal) showed significant effects on ratings of partner trustworthiness, connectedness, and similarity to partner. The statistics are reported in Table 21 (all pairwise comparisons are Bonferroni corrected). There were no effects on reliance, likeability or confidence in partner. (See Figures 11-13 below.)

Table 21 *Significant main effects, pairwise comparisons, and effect sizes of goal type on measures of partner affiliation*

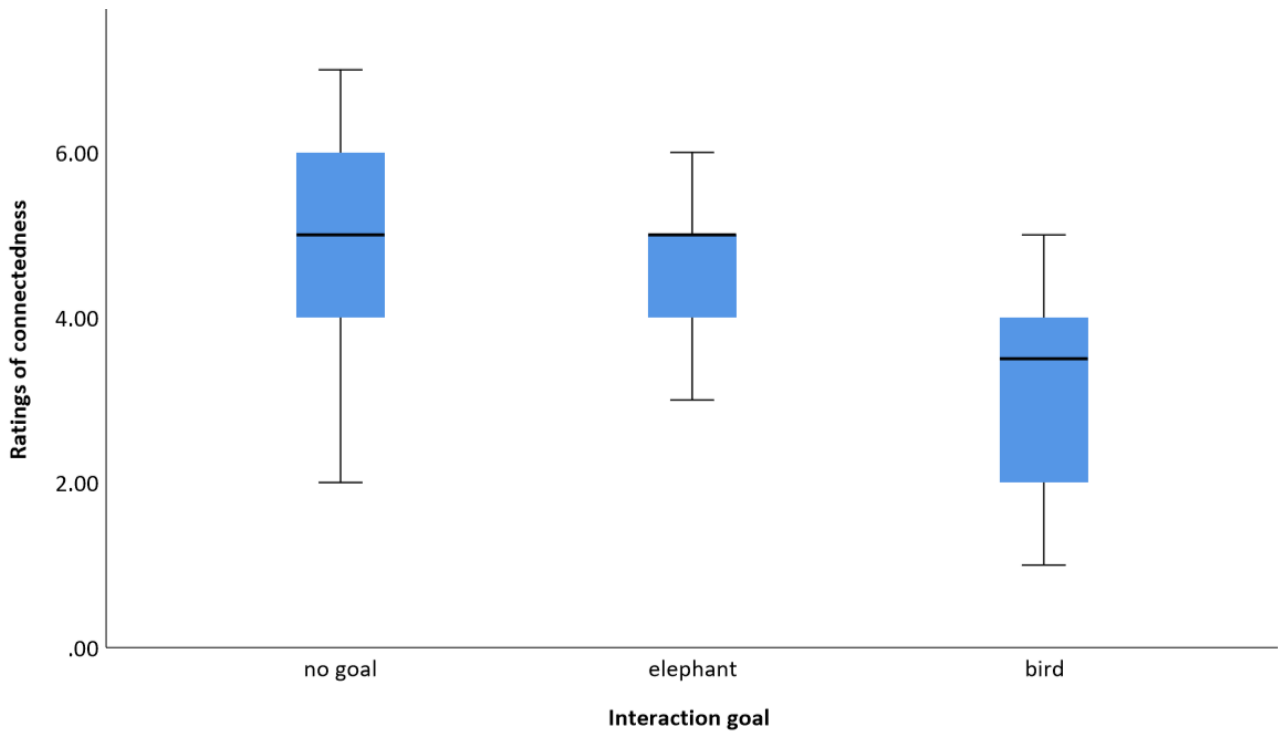| Measure	Main effect	Pairwise comparison		Effect size
Trustworthiness	H(2) = 7.055 p = .029	No goal > Bird goal	H = 10.350 p = .024	Cohen's *d* = 1.157 *r* = .501
Connectedness	H(2) = 7.970 p = .019	No goal > Bird goal	H = 10.750 p = .023	Cohen's *d* = 1.200 *r* = .515
Similarity	H(2) = 6.273 p = .043	Elephant > Bird goal	H = 9.717 p = .038	Cohen's *d* = 1.325 *r* = .552

Figure 11 Boxplot showing effect of interaction goal on ratings of connectedness towards interaction partner. Ratings were significantly higher following the no-goal than the bird goal condition (plot shows median, interquartile range, and top and bottom 25% of ratings).
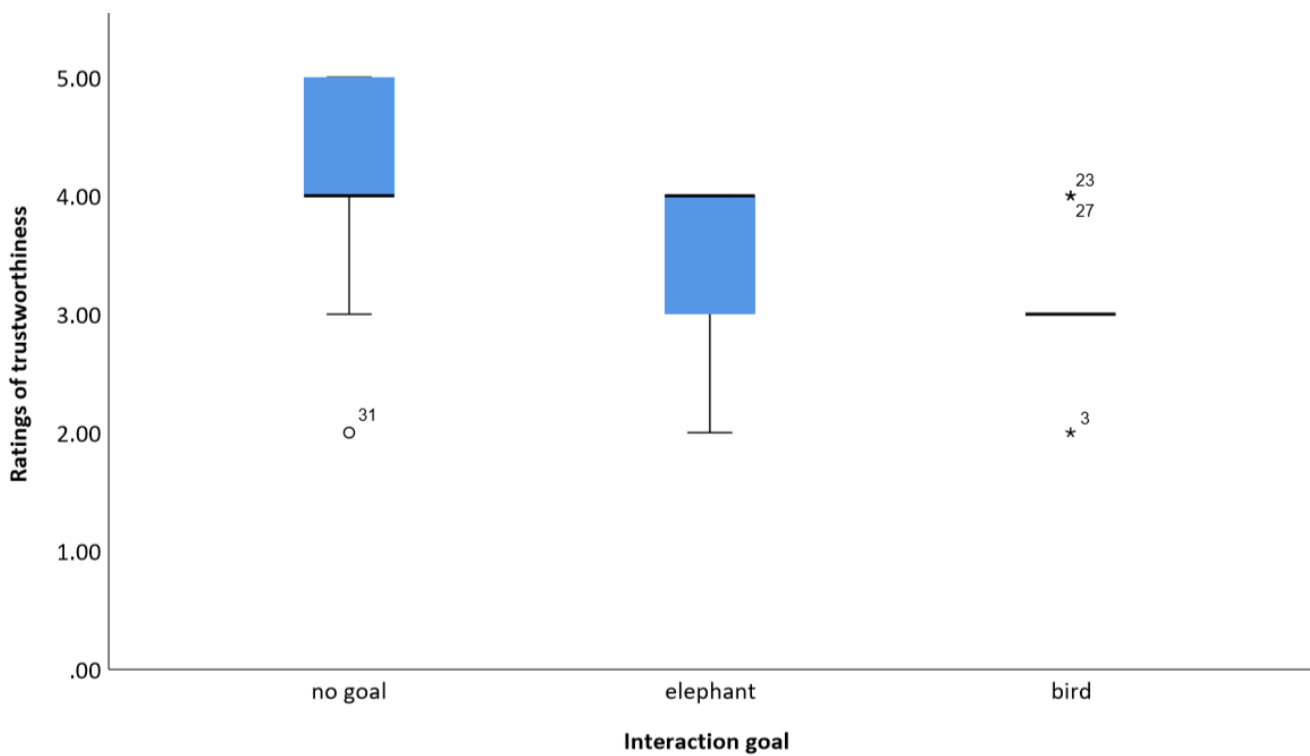

Figure 12 Boxplot showing effect of interaction goal on ratings of partner trustworthiness. Ratings were significantly higher following the no-goal than the bird goal condition (plot shows median, interquartile range, top and bottom 25% of ratings, and outliers).
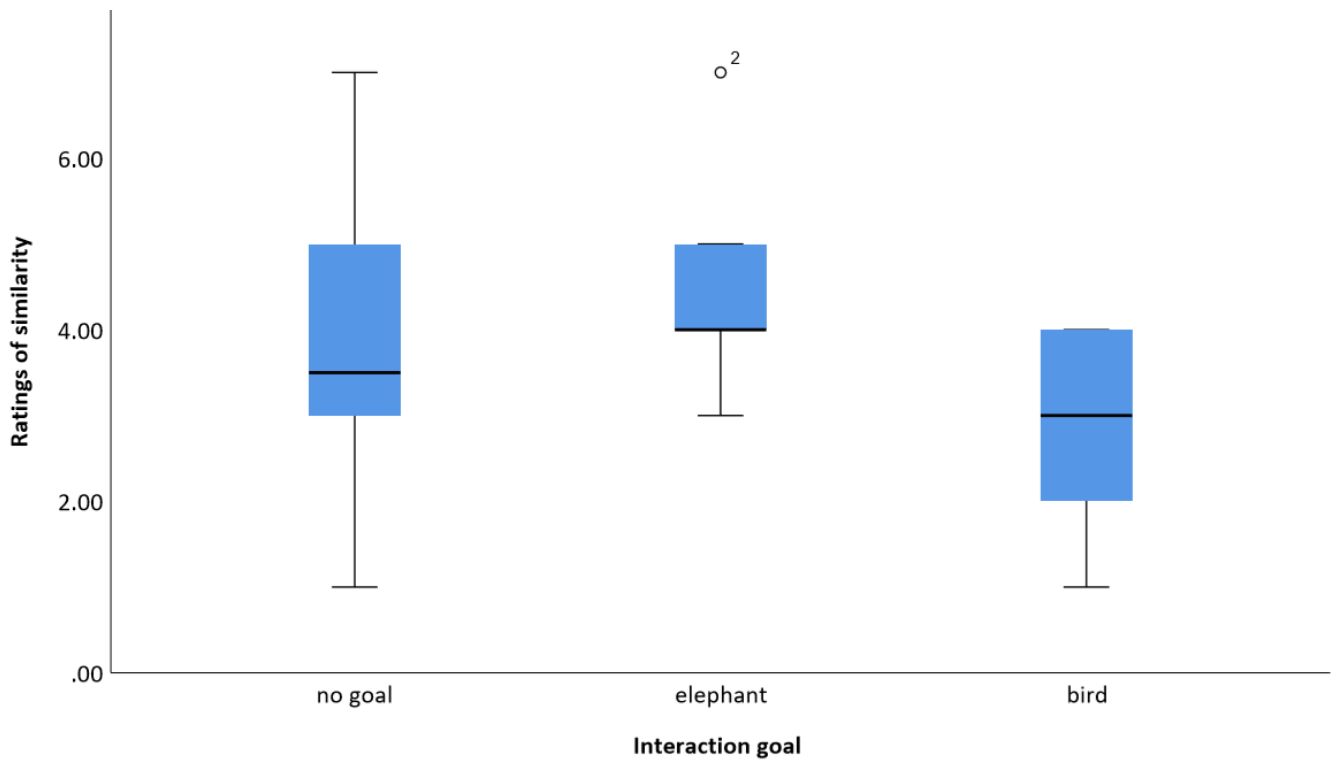
Figure 13 Boxplot showing effect of interaction goal on ratings of similarity to partner. Ratings were significantly higher following the elephant goal than the bird goal condition (plot shows median, interquartile range, top and bottom 25% of ratings, and one outlier).

Task success

There were no effects of interaction condition on ratings of difficulty, performance or task success. However, there was a significant effect on ratings of embarrassment: $H(2) = 6.224$, $p = .045$. There were no significant pairwise effects, except a marginal contrast between the bird and no-goal condition - bird > no goal, $H = -9.700$, $p = .052$.

Team identification

There were no significant effects of interaction condition on explicit team identification (IOS ratings). There were also no main effects of condition on reward allocation, or on team-related word recall (% own-team words, including and not including small errors).

Predicting social identification: ratings of success

A multiple linear regression model for own-team word recall (an implicit measure of team identification), using measures of interaction success as predictors (difficulty, individual performance, success, and embarrassment) was significant: $F(4) = 3.231$, $p = .027$, adjusted $r^2 = .224$. The effects of individual predictors within this model are shown in Table 22. Ratings of *individual* performance are negatively correlated with social identification, but *joint* success is positively correlated (nearing significance as an individual predictor). Embarrassment is negatively correlated with social identification.

Table 22 *Performance, success and embarrassment as predictors of team identification*

Predictor	Beta coefficient	T	P
Individual performance	-.627	-2.944	.007
Embarrassment	-.406	-2.451	.021
Joint success	.385	1.943	.063

Note: team identification is as shown implicitly in own-team word recall.

Using non-parametric correlations, own-team word recall is correlated only with ratings of embarrassment (Spearman's *rho* = -.357, $p = .045$). In turn, embarrassment does not correlate with any other ratings (apart from another negative correlation with ratings of interpersonal connectedness: *rho* = -.443, $p = .011$).

This linear regression model did not predict reward allocation. Moreover, linear regression using the above measures of partner affiliation as predictors (similarity, trustworthiness, and connectedness) was not significant in predicting own-team word recall (or point allocation).

Predicting IOS ratings: both partner affiliation and success

An IOS scale was used as an explicit measure of team identification. However, the results of this seem to be (weakly) predicted by both interaction success and ratings of partner affiliation and closeness. The results of non-parametric Spearman's ranked correlations are shown in Table 23.

Table 23 *Significant non-parametric correlations with rating of IOS with team*

Measures	Connectedness	Similarity	Likeability	Confidence	Success
IOS rating	.593 (p < .001)	.568 (p = .001)	.503 (p = .003)	.378 (p = .033)	.359 (p = .044)

Note: each box shows *rho* (p-value).

An ordinal regression model of IOS ratings using those significant measures as predictors (connectedness, similarity, likeability, confidence and success) was significant: Chi-square = 68.591 df 25, p > .001, r^2 = .883 (Cox & Snell).

Affiliation and performance: are they separate factors?

A factor analysis was run on the following variables: IOS, own-team word recall, confidence, trustworthiness, connectedness, success, similarity, likeability, difficulty, performance and embarrassment (reliance was excluded as it correlated strongly with confidence).[127]

Principal Component Analysis extracted three component factors (above Eigenvalue of 1), and then orthogonal Varimax rotation was run.[128] The resulting factor loadings of each variable are shown in Table 24.

Table 24 *Factor loadings on affiliation, confidence and identification*

1 - affiliation	2 - confidence	3 - identification
IOS	Confidence	Word recall (negative)
Connectedness	Trustworthiness	Embarrassment
Similarity	Joint success	
Likeability	Difficulty (negative)	
	Individual performance	

 Note: only factor loadings above .5 are shown.

Most of the partner affiliation measures load onto component 1, alongside IOS team identification. However, own-team word recall loads onto a different component (with ratings of embarrassment), as do ratings of success, confidence, trustworthiness, difficulty and performance.

Mood changes, musical training and sex differences

Spearman's *rho* correlations indicated statistically significant relationships between baseline mood, mood changes and a number of rating scales (see Table 25).

---

[127] This dataset met the usual assumptions for factor analysis: KMO > .5, and Bartlett's Test of Sphericity was significant at p < .001. However, these results should be taken as a guide only due to a smaller than usual sample size for this type of test.

[128] Orthogonal rotation was used as these variables are not strongly correlated. However, there is reason to believe that e.g. connectedness and trustworthiness are related (and IOS and confidence - see above), thus a different sort of rotation might be used if this experiment is repeated with a larger sample size.

Table 25 *Non-parametric correlations with baseline and change in mood*

	Difficulty	Perfor-mance	Success	Trustwor-thiness	Reliance	Confidence	Connected-ness	Likeable
Baseline +	-.482*	.436	.439	.352				
Baseline -		.360	.512*	.551*	.617*	.485*	.470*	.397
-ve change		-.403*	-.505*	-.489*	-.546*	-.450*	-.500*	-.533*

Note: all are significant at p < .05 level; those marked * are significant at p < .01.

Baseline positive mood correlated with ratings of interaction success; baseline negative mood correlated with ratings of confidence in or reliance on the partner, and also partner affiliation measures. An increase in negative mood from pre- to post-interaction had a negative correlation with those indicators of partner trustworthiness and affiliation, as well as success of the interaction.

Kruskal-Wallis comparisons indicated a significant effect of musical training on ratings of connectedness. Those who indicated that they were *not* practising musicians rated partner connectedness as higher than those who were practising musicians: H(1) = 5.036, p = .025. Similarly, degree of musical training and connectedness were negatively correlated (*rho* = -.431, p = .014), as were connectedness and frequency of musical activities (*rho* = .-380, p = .032). This seems to indicate that those who were musically trained were less likely to feel connected to their partner.

Kruskal-Wallis comparisons showed no effect of sex, or of the sex composition of the pairs, on any indicators of partner affiliation or team identification.

Discussion

This experiment tested the relationship between different modes or goals of interaction, and measured (perhaps) dissociable effects on interpersonal bonding and team identification. In line with hypothesis (a), I found higher ratings of partner connectedness, similarity and trustworthiness following the no-goal than the bird goal condition. Contrary to expectations, the bird and the elephant goal conditions did not vary systematically in difficulty or perceived success (see below for potential reasons for this). Therefore, there was no relationship between goal type and team identification. However, there was a relationship between team-related word recall (implicit measure of category formation or identification) and interaction success (specifically, how embarrassing it was, across all interaction conditions). In line with hypothesis (b), word recall was predicted by perceptions of interaction success (marginally significant), and by lower levels of embarrassment. Interestingly, it had a negative relationship with perceptions of individual performance. These findings are discussed in more detail below.

The no-goal musical interaction task led to significantly higher partner affiliation than the musical interaction with an added bird depiction task. This seems to provide evidence that the mode of the interaction - the amount of freedom in the improvisatory task, and whether or not it involves an explicit external goal (or extrinsic orientation) - affects its efficacy in forming social bonds within a dyad (or, by extension, a group). In contrast, team identification, as measured using word recall, was related to interaction success - specifically, degree of embarrassment. This is evidence for the role of self-esteem and 'face' in team identification: if the interaction is more embarrassing, the participant is less likely to identify with that team (and is also less likely to feel connected to their partner). However, *lower* ratings of individual performance predicted higher word recall. This seems to contradict these findings; however, individual performance is different to any form of perceived joint success. If an individual feels that they personally "did well", then they are less likely to associate with the others on their team than if they feel it was a more joint effort, or that their association was a more successful one (see e.g. Kokal et al., 2011).

Mood might have been important in shaping perceived interaction success: baseline positive mood correlated with interaction success, while baseline negative mood correlated with higher ratings of confidence in and affiliation with the partner. These did not correlate with ratings of social identification. Feeling more negative over the course of the interaction, however, had a negative relationship with these ratings of partner affiliation and interaction success.

The respective roles of interaction mode on partner affiliation, and the joint success of the interaction on team identification, seem to point to a conceptual distinction between interpersonal, group-like bonding, and social identification with a category or team. This is supported to some extent by the factor analysis, with connectedness, similarity and likeability loading onto a different component to confidence, success, difficulty and performance.

Limitations

The previous relationship between perceived success and interaction goal (bird and elephant tasks) was not found here, perhaps due to musical training (in the previous experiment, it is likely that a higher proportion of participants were musically trained than in the present study), or due to differing interaction contexts (here, they were in dyads who shared a goal, whereas previously participants were in groups of five, where the goal was not shared amongst all participants). Furthermore, there was no apparent effect of goal presence on team identification, in contrast with findings from Experiment 2; this might well also relate to the fact that, previously, teams were mixed in the interaction task (so goal *difference* between teams was more apparent), whereas, in this experiment, participants interacted in same-team pairs.

Another possible limitation is the potential confound with baseline positive mood, which was higher amongst participants in the no-goal than those in the bird goal condition. This should not affect the results reported here, as participants were selectively removed on this basis; however, this experiment should be replicated with a larger sample size. With a higher N, we could also run the factor analysis more reliably, to explore further a potential dissociation between indicators of group bonding and team identification.

The explicit rating of team identification - the IOS scale - was predicted by a larger range of factors, spanning both interaction success ratings as well as various indicators of partner affiliation. This is perhaps due to the nature of the scale, which asks participants to depict their felt association with the team identity visually, in terms of degree of overlap between two circles. This more explicit representation might relate more to their interaction experience with their team partner (rather than with the more abstract concept of the team itself).

Finally, the relationship between interaction difficulty, success, embarrassment and individual performance is still unclear from the present analysis. Although embarrassment and individual performance are related to team identification, joint success shows only a marginal relationship. Moreover, embarrassment correlates more strongly with interpersonal connectedness than with joint success. While it is apparent that team identification emerges from the social success of the encounter in some sense, embarrassment has more far-reaching implications (as one might expect) for the personal affiliations within the interaction, as well as the identities one develops as a result.

Video analysis - experimental interaction

In this section, I present the findings from a case study video analysis, in which I compared the behaviour involved in a no-goal condition (a more relational interaction) with the goal-focussed (extrinsically-oriented) musical tasks. I expected that these interaction conditions would have different goal orientations, and thus might necessitate different types of common ground - one more internal, or relational, and other more external, transactional, its focus outside of the interaction.

I observed the gaze behaviour of each participant, anticipating that their looking patterns might offer insights about the interaction focus (see Chapter 2 on the role of gaze in grounding). I also measured the duration of interaction behaviour in each condition, and the amount of this time spent in joint synchrony (of note onsets). Synchrony has previously been linked to social bonding and fluency in dialogue (see Chapters 4 and 2, respectively). I expected that higher levels of synchrony (and an intrinsically-focussed interaction) might lie behind higher levels of interpersonal affiliation (no-goal/elephant goal > bird goal - see experiment 3).

This was an exploratory case study. Prior work investigating gaze behaviour and musical interaction has investigated the role of looking patterns within a performance informed and regulated by specific conventions (see work by Nikki Moran, 2010, 2013, on gaze and gesture in North Indian classical music). The present analysis complements this by comparing two more basic interaction tasks, activities which model an overall distinction between more musical and speech-like modes of relating (compare also e.g. Vandermoortele et al., 2018); no comparison of this sort has been made before. I shall now report my method, before presenting my main findings; then, I shall discuss my results in the context of the literature and my other experimental results.

Method

In what follows, I shall detail how I set up my interaction conditions, and my procedure for the video analysis.

Participants

A total of 12 participants (aged 19-22, mean = 19.75 years) took part in the study, forming six interaction dyads. Three dyads (age 19-22, mean = 19.5) were selected for video analysis. These were chosen as none of the six participants had prior knowledge of their partner (all rated prior knowledge as 1). All participants had a high level of musical training (they were recruited through the music faculty or local choirs). Two dyads were mixed sex, and one was female only. All participants provided informed consent to take part in the study, and for their interaction to be filmed.

Design

This comprised a between-groups design, varying interaction goal - no goal, bird goal (an easy goal), or elephant goal (a difficult goal).[129] The dyads chosen for video analysis each represented one of these different interaction tasks (see Table 26).

---

[129] The tasks had been piloted to ensure that the bird condition was considerably easier than the elephant condition. However, I found in experiment 3 that these goal conditions did not vary in self-report ratings of success or difficulty.

Table 26 *Participants assigned to each interaction condition*

Interaction goal	Participant	Sex
Bird	LC	Female
	BK	Female
Elephant	PC	Male
	SPC	Female
No goal	CL	Male
	AB	Female

Procedure

Participants were given a simple vocalising instrument (a kazoo) and told that they could interact with their partner using only this kazoo (without talking or otherwise communicating). The interaction task lasted for five minutes. The instructions given for each condition were similar to those used in previous experiments (and detailed in Appendix 3.1).

Afterwards, participants were asked to rate different aspects of their experience of the task in a self-report questionnaire, including their affiliation to their partner, how difficult they found the task, and how well they think they performed.[130] Participants were also asked to rate (and detail, if necessary) how much they knew their partner (7-point scale). After they completed this questionnaire, participants were thanked and debriefed.

Video and audio recording

Participants were filmed during their interaction, and their individual contributions recorded using personal microphones. This took place in a sound-proof recording studio. Participants were sat down for the duration of the task, and thus their proximity to one another was the same in each interaction condition. Video cameras were positioned so that the facial expressions (including eye movements) of each participant could be recorded head-on. The video footage allowed discrimination between whether participants were looking towards the face region of the other participant, or whether they were looking away; when both participants looked towards the other, this was defined as a period of mutual gaze (compare similar approach, for looking patterns rather than gaze, in Moran, 2013). The audio tracks were analysed to locate note onsets; this data was integrated with the gaze data for the behavioural analysis. (See Appendix 3.2 for details of the recording set-up and of how the audio tracks and video footage were processed.)

---

[130] This data was not analysed for the whole sample (n = 12), as some dyads knew each other well before the experiment. The data provided for the case study sample (n = 6) was not large enough to run statistical tests.

Behavioural analysis

Note onsets were used to show bouts of joint synchrony and any turn-taking behaviour. These were analysed alongside gaze onsets (looking towards the other), to investigate any patterns in the relationship between these variables.

Bouts of joint synchrony

Joint synchrony was defined in terms of the inter-onset intervals (IOI) of note onsets (using audio data). Bouts of joint synchrony were to have the following characteristics:

1) A clear beat, maintained at approximately the same rate throughout
2) This beat rate or tempo is shared between both participants
3) Both participants contribute at least once during that period

I calculated the beat rate or IOI for each participant using the inter-onset measurements between notes. Where the notes fit into a beat structure, but comprised a different rhythm, then the beat was calculated using the relative proportions of those inter-onset times e.g. adding together shorter notes which comprised a beat, or halving a longer note which lasted for two beats. In all cases, beat IOIs were around a second in length.[131]

To give an overall IOI for each bout of joint synchrony, I combined the data from both participants. This was calculated as the mean of local beat onsets (this allowed for some flexibility of beat rate within each bout). A new bout of joint synchrony was marked when there was a significant change in perceived beat IOI. A bout of joint synchrony ended when the beat IOI changed or ceased to be shared, e.g. when one participant introduced a new tempo that was not continued by their partner.

Two independent coders also indicated the start and end of bouts of joint synchrony, and whether they considered those bouts to be successful or unsuccessful.[132]

Gaze onsets in bouts of synchrony

I investigated the relationship between gaze onsets and beat IOI for each participant. This was calculated by measuring the time between gaze onset and the nearest beat onset (either before or after). Gaze onsets were considered "in time" when they occurred within the standard deviation for the beat IOI for the whole bout of joint synchrony.

---

[131] This is slightly slower than the spontaneous tapping tempo, but is well within the limits of 'subjective rhythmization' (see London, 2002).

[132] The experimenter used the note onsets (extracted using Praat). However, the independent coders used only the raw video material (with audio). They could start, stop and rewind as much as they needed to. They each viewed the goal conditions in a different order to counterbalance the effects e.g. of boredom. This was a quicker measurement, which complemented the more precise analysis of IOIs.

I also explored the relationship between gaze onsets and other features of the interaction (outside these bouts of joint synchrony). In the no-goal condition, I detected points of convergence; these had the following characteristics:

1) Both participants' note onsets occurred in synchrony i.e. within 250 ms[133]
2) At a structurally important moments in the phrase i.e. at its beginning or ending
3) With a harmonic relationship between the pitch height of each participants' note

The relationship between gaze onsets and endings (looking to their partner and away, respectively), and these points of convergence was measured.

In the elephant goal condition, there were both points of convergence and more general *periods* of convergence. Periods of convergence were times of interaction in which there was more than usual sustained synchrony between the interaction partners. Approximate *points* of convergence occurred both within these periods and as isolated points, and were moments when there was particular synchrony (250ms deviance or less) and/or harmonic consonance between the individuals' pitches.

Turn-taking and gaze behaviour

A pattern of turn-taking between the participants was detected in the two goal conditions. A new turn was defined according to the following rules:

1) A turn switch between the two participants OR
2) If the same individual started a new phrase, after a pause, and with a substantially different idea to their previous phrase.

I investigated the relationship between participants' turns, including, (a) who took each turn, (b) whether it was matched in pitch to the previous turn (i.e. whether any of the notes matched any of the notes of the previous turn), (c) whether it was in time with the previous turn (i.e. did the first note of the turn occur within the beat pattern established, within one standard deviation of the previous beat structure, if this was present), and (d) the relationship of the participant's gaze onsets and endings to the start of their turn.

The patterns observed through this analysis are reported in the next section.

Results

As in experiment 3, participants reported similar levels of difficulty and performance success across the different interaction conditions (see Table 27).

---

[133] 250ms is the 20% deviation in metronome beat detected by all participants in Stephan et al. (2002).

Table 27 *Ratings of difficulty, performance and perceived success for participants in each goal condition*

Interaction goal	Difficulty (/5)	Performance (/5)	Success (/7)
Bird	4	2	4
	2	3	5
Elephant	4	2	5
	2	4	5
No goal	4	3	6
	2	4	4

Bouts of joint synchrony

For each bout of joint synchrony, a mean IOI and standard deviation were calculated. For all bouts of joint synchrony (identified by the experimenter), the overall mean IOI was approximately one second, and the mean standard deviation for all bouts was about 10% of that total.

Graphs in Figure 14 show (a) the overall mean IOI for all bouts of joint synchrony, and the standard deviation about that mean, and (b) the mean variability, or an average of all the standard deviations for each period of joint synchrony. These are all roughly the same, across all conditions (although the beat IOI was perhaps slightly quicker for the no-goal than the bird condition). (See Appendix 3.3.1 for the raw data - the mean and standard deviation IOI calculated for each bout of joint synchrony, in each condition.)
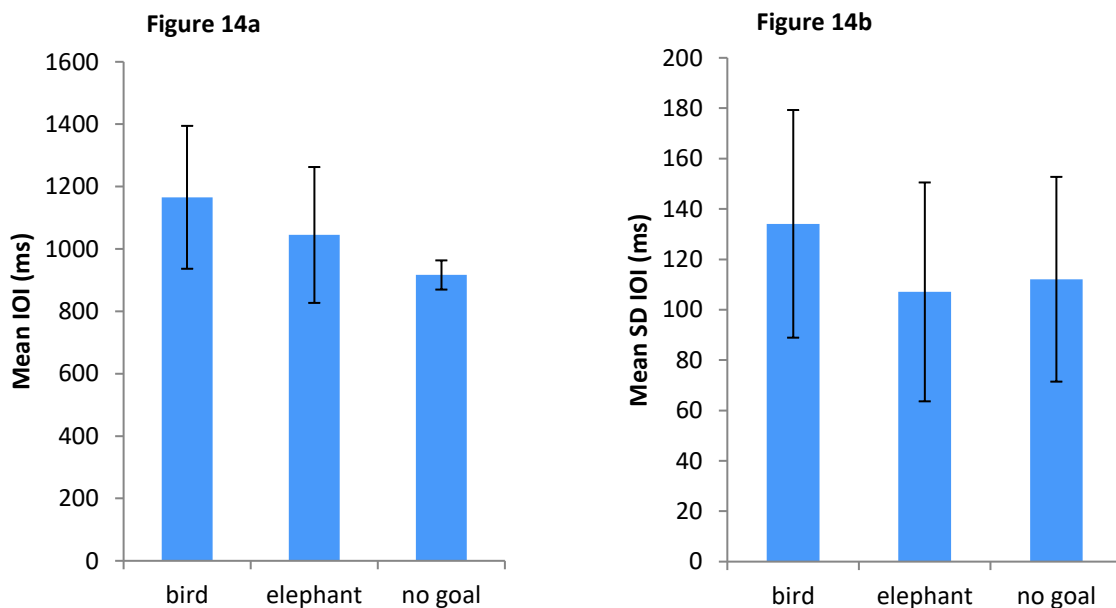

Figure 14 Overall IOI statistics for all bouts of joint synchrony, for each interaction condition. Figure 14(a) shows the mean IOI across all bouts (with standard deviation). Figure 14(b) shows the mean standard deviation in IOI for all bouts (with standard deviation).

However, there were clear differences between interaction conditions regarding the *duration* of those bouts of joint synchrony. Table 28 reports the total interaction duration, the total duration of bouts of joint synchrony, and the percentage of total interaction time spent in joint synchrony, for each interaction condition.

Table 28 *Length of overall interaction and, of that, the time spent in joint synchrony, for each interaction goal condition*

Interaction goal	Interaction duration (sec)	Joint synchrony (sec)	% joint synchrony
Bird	284.4	100	35.2
Elephant	309.8	164	52.9
No goal	293.6	258	87.9

Although the amount of interaction time is roughly the same across conditions, the duration of this spent in joint synchrony is very different. Taking into account the ratings by the two independent coders gave a more conservative estimate of the number of bouts. (See Appendix 3.3.2 for the durations of joint synchrony identified by the experimenter, and by the two coders.)

Table 29 shows the bouts of synchrony for which there was complete agreement (between the experimenter and both coders) and whether they were considered successful or unsuccessful - (1) a successful bout has SD IOI < 10% of the mean IOI, and (2) and (3) whether they were classed Y or N (successful or unsuccessful) by the two independent coders.

Table 29 *Bouts of joint synchrony as identified by the experimenter and two independent coders, and whether they were considered successful*

Bird interaction condition					
Start of bout	**End of bout**	**Length of bout**	**Successful?**		
02:19	02:25	00:06	N	Y	N
05:09	05:14	00:05	Y	Y	Y
Elephant interaction condition					
00:30	00:33	00:03	Y	Y	Y
04:42	04:45	00:03	Y	Y	Y
No-goal interaction condition					
00:34	01:42	01:08	N	Y	Y
01:54	02:07	00:13	N	N	Y
02:16	04:24	02:08	N	Y	Y
05:12	05:32	00:20	N	Y	Y

The length of time spent in joint synchrony (identified by experimenter and validated by both coders) is shown in Table 30 (and presented in Figure 15).

Table 30 *Time spent in joint synchrony for each interaction goal condition.*

Interaction goal	Joint synchrony (sec)	% joint synchrony
Bird	11	3.9
Elephant	6	1.9
No goal	229	78.0

Note: this shows joint synchrony as validated by both coders, also expressed as percentage of total interaction time.
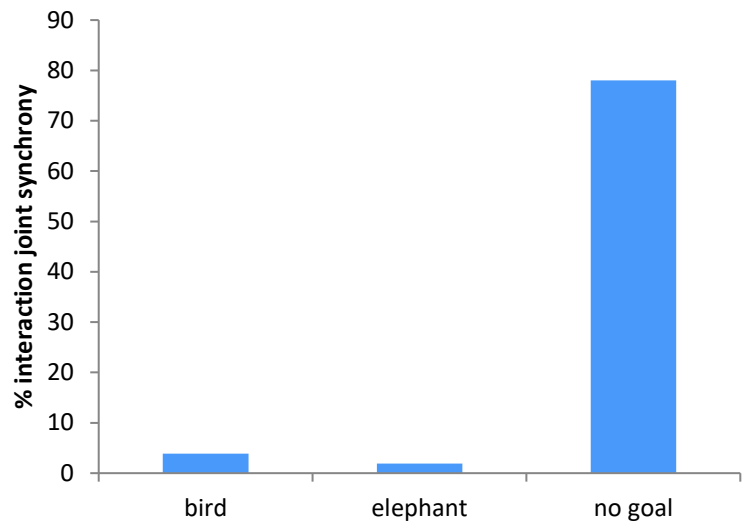
Figure 15 Percentage of interaction time spent in joint synchrony for each goal condition

Gaze behaviour

Table 31 shows the total incidence of individual gaze by each participant in each condition, and how much of this was mutual gaze.

Table 31 *Gaze behaviour of each participant*

Interaction condition	Participant	Individual gaze		Mutual gaze	
		Number	Length (sec)	Number	Length (sec)
Bird	BK	71	99.872	52	37.744
	LC	57	121.124		
Elephant	PC	29	81.688	31	39.192
	SPC	48	99.368		
No-goal	AB	55	124.829	13	9.674
	CL	14	14.046		

Note: this shows total incidence and length of individual gaze, and incidence and length of mutual gaze (when both participants look to each other i.e. individual gaze overlaps).

The no-goal condition had the least mutual gaze, but also the largest imbalance of gaze (as AB looks to CL more than vice versa).

There was no striking difference in the number of gaze onsets which corresponded to the beat, within bouts of joint synchrony (defined according to the experimenter's ratings). Table 32 gives a

percentage of gaze onsets which occurred within one standard deviation of the mean beat IOI, out of the total number of gaze onsets which occur within those bouts of joint synchrony (across both participants).

Table 32 *Relationship between gaze and interaction beat*

Interaction goal	No. gaze	No. gaze within 1 s.d.	% gaze within 1 s.d.
Bird	43	12	27.9
Elephant	39	12	30.8
No goal	56	15	26.8

Note: this shows the total incidence of gaze in bouts of joint synchrony, and the number of gaze onsets occurring "in time" with the beat (i.e. within one SD of the bout IOI), also expressed as a percentage.

Analysing the interaction as a whole, not just the bouts of joint synchrony, revealed some interesting patterns and some differences - especially between two of the interaction conditions: the no-goal task, and the bird goal condition. The elephant goal condition seemed to fall between these two extremes. These observations will be reported next.

No-goal condition - observations

The no-goal condition comprised more simultaneous playing, involving both participants throughout. There was no detected turn-taking behaviour. Their interaction was divided into clear phrases, and they would breathe between the phrases together. This phrase structure displayed general adherence to Western tonal harmony, in which they were both highly trained.

As a result of this structure, there were clear points of convergence, where both participants played in synchrony, and with pitches that had a harmonic relationship to each other. These occurred at structurally-important moments - at the start or at the end of phrases. Gaze towards the other occurred before and/or during this point of convergence. (See Appendix 3.4 for all of these points of convergence, and the gaze behaviour which occurred within one second before and after this point.) Most (18 out of 26) points of convergence are preceded by gaze by one or both of the participants, within a second of the point of convergence; gaze then does not need to occur after this point. These points account for around 40% of all gaze onsets during the interaction.[134]

Mutual gaze also occurs around these moments, generally before or at a phrase transition (including some of the points of convergence documented above). Table 33 shows all the points of mutual gaze

---

[134] This figure includes an aggregate of both participants, and includes some gaze onsets in the phrase immediately leading up to the point of convergence (sometimes more than one second before). The percentage is out of the total incidence of gaze towards the other.

throughout the interaction. All occur at phrase boundaries, with a possible point of convergence (of varying synchrony and harmonic consonance) occurring afterwards. All of these points are in this case signalled and preceded by a breath together.

Table 33 *Mutual gaze during interaction, and co-occurrence with other interaction behaviour*

Mutual gaze		(Possible)	Synchrony	Pitch relationship	Other
Start	Finish	P.O.C.	(ms)		behaviour
00:50.6	00:52.8	00:56.7*	181	Octave	
01:11.2	01:11.6	01:12.0*	14	Fourth	
01:46.2	01:46.6	01:48.4	96	Fourth	
02:06.3	02:07.7	02:09.2	523	Non-consonant	Laughter
02:20.2	02:21.3	02:24.7*	134	Third	
02:47.0	02:48.4	02:48.3	305	Third	(Laughter)
02:50.6	02:50.9	02:52.1	84	Non-consonant	Laughter
03:00.0	03:00.3	03:01.6	42	Sixth	
03:46.2	03:47.1	03:45.9	16	Non-consonant	
04:06.3	04:06.8	04:05.6	88	Non-consonant	Laughter
04:23.7	04:24.2	04:24.0	107	Non-consonant	Laughter

Note: the points denoted with an * are points of convergence also documented above (see Table 48). The others do not fall into that category because they (a) have synchrony of > 250, (b) are not consonant, due to breaking into laughter shortly after, or (c) are at less structurally-important phrase boundaries. (Laughter) denotes behaviour which signals the start of a frame break, which then occurs at 02:52.

The rows highlighted in grey do not follow the standard pattern: the points of convergence occur before rather than after the period of mutual gaze. In general, however, the same pattern emerges - gaze behaviour occurs before moments when the participants play together, co-ordinating those points where they coincide - in some sort of harmonic or temporal synchrony.

The following picture sequence (Table 34) shows the first of these moments of mutual gaze unfolding.

Table 34 *Sequence of events during and after a moment of mutual gaze in no-goal condition*

Time point	AB	CL	Observations
00:47.1			Continuing phrase, both are looking away
00:48.3			AB looks towards CL
00:50.9			CL reciprocates, establishing a period of mutual gaze

00:52.8			CL ceases to gaze towards AB
00:56.6			AB looks away just before the point of convergence (which forms the start of the next phrase)

Note: left: AB; right: CL.

There are several points when the interaction is disrupted by laughter. These sometimes lead to a complete breakdown in the interaction. The points of laughter occur immediately (or a short time) after mutual gaze (an estimation of eye contact). See Table 35 for details of what happens when these points occur.

Table 35 Co-occurrence of mutual gaze and mutual laughter in no-goal condition

Mutual gaze		Mutual laughter onset (est.)	Frame break?
Start	End		
00:50.6	00:52.8	00:51.98	No
02:06.3	02:07.7	02:07.17	Yes
02:50.6	02:50.9	02:52.20	Yes
03:00.0	03:00.3	(cont.)	No
03:46.2	03:47.1	03:55.11	No
04:06.3	04:06.8	04:07.14	No
04:23.7	04:24.2	04.24.13	Yes

Note: a frame break is when the interaction completely breaks down, and they resume together a short time afterwards.
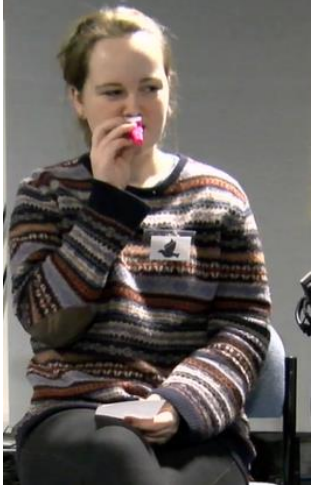
Bird goal condition - observations

The bird goal condition displayed another consistent pattern of gaze behaviour - but one which contrasted to that shown in the no-goal task. In the bird goal condition, the dyad kept to a turn-taking structure, with less simultaneous playing. Participants BK and LC took it in turns to play. Each looked away before their onset (or were already looking away), and then checked back after their onset. They looked away between 5 seconds and 41 ms before the onset of their turn (mean: 1770ms before; or, in six cases, were not looking at all), and then checked back between 68 ms and 5 seconds after the turn onset (mean: 1649ms after; all gaze after is included when it is within a second after turn onset, or occurs at some point during the turn). This pattern held for 66 out of a total of 80 turns (including both participants) - i.e. 82.5% of turns. Of the turns that conformed to this pattern:

- 51 (77.3%) were matched in pitch to the previous turn
- 28 (42.4%) were in time with the previous turn
- 21 (31.8%) were both matched in pitch and timing

There were no significant differences between participants within the dyad.

The following picture sequence (Table 36) shows one example of this pattern unfolding:

Table 36 *Sequence of events during a turn transition in bird goal condition*

Time point	BK	LC	Observations
01:08.8			BK starts her turn, with averted gaze
01:09.9			BK checks back at LC shortly after her turn starts
01:12.3			LC starts her turn, with averted gaze

01:14.3			LC checks back, again during her turn
01:14.9			BK starts her next turn, with averted gaze

Note: left: BK: right: LC.

In this case, instances of mutual gaze often occurred in the cross-over between turns, and were not related either to particular points of convergence, or to moments of breakdown or laughter. Of the remaining 14 turns that did not conform to the look away/check back pattern described above:

- 9 had either (a) no gaze before or (b) looked away beforehand (as expected) but then did not check back after the turn onset
- 1 where the participant looked away at the onset of the turn (rather than before) and checked back shortly after
- 4 had gaze before and during the turn onset (they did not look away)

On other occasions, the participants played simultaneously, and so a different pattern of gaze behaviour emerged. This occurred on several occasions, but especially from 03:14.8 to 03:30.5 and

03:45.1 to around 03:54.8. During these periods, the participants gazed at the other both before and during their turn onset, rather than averting their gaze. The first incidence of this (from 03:14.8) displays an interesting asymmetry between the participants: each time, LC joins simultaneously, looking towards BK during her turn onset. However, each time BK enters, she starts a new turn as normal, gazing away from LC beforehand. During the second period of simultaneity (from 03:45.1), both participants gaze towards each other at their turn onsets, either imitating the other, or continuing their own previous turn material, while the other continues to play. This same gaze pattern then later occurs only for LC, when she joins simultaneously, sometimes imitating BK (this occurs at 03:57.4, 04:02.6, 04:14.4, and at 04:26.1).

Thus, with a bird goal, the interaction assumes a turn-taking structure for the majority of the interaction time. However, there are short periods of the interaction which deviate from this pattern - participants can play simultaneously, and so gaze towards each during (rather than after) their turn onsets, rather like the points of convergence in the no-goal condition. This mixture of patterns is more obvious in the elephant goal condition, as will be reported next.

Elephant goal condition - observations

The elephant depiction goal was chosen as it was expected to be a harder goal to achieve in this context (especially with the musical instrument selected for the interaction). Although this was not reflected in the self-report ratings of success and performance (see above), the difficulty of the goal might be reflected in other patterns of behaviour. In this case, the turn-taking structure of the bird goal occurs for some of the time, but also there is more recourse to simultaneous, synchronous playing. This might be because the goal itself was harder to maintain, so it was a less important part of the interaction (in comparison with the bird goal condition).

First, the pattern of gaze aversion for turn onset and checking back afterwards occurs for both participants, for a total of 23 times. (See Appendix 3.5.1 for details of each of these.) On each of these occasions, the participants are often playing simultaneously, but at each turn they introduce their own independent material. Second, a similar pattern of convergence to what is displayed in the no-goal condition occurs, and more regularly than the occasions in the bird goal condition. In this case, instead of particular points of convergence, there are periods of more sustained synchrony. These periods always begin with either sustained gaze from one or both participants before and/or during onset, or the establishment of synchrony very early on (meaning that gaze is not essential to co-ordinate their interaction). Often, there is a point of convergence in the middle of this period,

facilitated by a period of mutual gaze. On other occasions, there are isolated points of convergence. (See Appendix 3.5.2 for more precise details.)[135]

In this way, the elephant goal involves both patterns of gaze behaviour - with gaze aversion before and checking back after each turn onset; and gaze behaviour which occurs before or during particular points (or periods) of convergence. Most (all but two) of the instances of mutual gaze occur during these points of convergence. Similar to the no-goal condition, the interaction breaks down with laughter on a number of occasions - all after periods of mutual gaze. See Table 37 for a summary of these points.

Table 37 *Co-occurrence of mutual gaze, mutual laughter, and interaction breakdown in the elephant goal condition*

Mutual gaze		Mutual laughter onset (est.)	Frame break?
Start	End		
00:55.27	00:57.62	00:56.99	Yes
01:31.94	01:33.29	01:32.46	No
04:01.00	04:03.32	04:02.69	No
05:00.95	05:02.50	05:02.17	Yes

The mixture between two distinct patterns might demonstrate that the goal in this case is less salient or important in the interaction, perhaps as a result of it being harder or less effective in this context. Therefore, the form of interaction is more towards the 'musical' end of the spectrum - it is more intrinsically-oriented (this is exhibited even more so, of course, in the no-goal condition) - maybe as a result of the difficulty of the goal task.

Summary of findings

Table 38 summarises the main features for each interaction condition.

---

[135] Sometimes, it is difficult to categorise their behaviour as either turn-taking (where an individual takes their turn, to exchange with the other) or periods of convergence (where there is more simultaneous, synchronous playing). There is inevitably a degree of overlap between the two.

Table 38 *Duration of joint synchrony, turn-taking, points or periods of convergence, and incidence of interaction breakdown in each interaction goal condition*

Condition	Duration of joint synchrony	Turn-taking behaviour	Simultaneous behaviour	Frame breaks?
No goal	High	None	Gaze (incl. mutual gaze) occurs before points of convergence	After mutual gaze
Bird goal	Medium	Gaze aversion before turn onset; check back after onset	Short periods: gaze before onset	None
Elephant goal	Low	Gaze aversion before turn onset; check back after onset	Periods of synchrony: gaze at/before onset; mutual gaze at points of convergence	After mutual gaze

These results reveal interesting patterns in the gaze behaviour. There are two distinct patterns which emerge depending on the prominence of the goal task. This might indicate different types of common ground, depending on whether the interaction is extrinsically-or intrinsically-focussed (with or without a feasible depiction goal). The role of mutual gaze is of particular interest in this respect. In addition, the incidence of frame breaks (i.e. breaking into mutual laughter) provides insights into common ground (as signalled by mutual gaze) from a different perspective. The findings relating to the duration of joint synchrony are also of interest. These interpretations will be discussed (alongside findings from my other experimental work) in my discussion.

Discussion

This video analysis reveals differences between forms of interaction which vary only in the presence (and type) of goals. As discussed in Chapter 2, interaction which has no external goal beyond that just to maintain the affiliative, interactive context, is more intrinsically-oriented. Whereas, adding an external, here depiction, goal creates more of an extrinsic focus. Both of these interaction types necessitate common ground - but the nature of this might differ. We expected that any differences would be revealed through patterns of both individual and mutual gaze (see role of gaze in conversation, detailed in Chapter 2).

On one hand, in the no-goal interaction condition, participants did not exhibit any turn-taking behaviour, but instead played simultaneously throughout, with long periods of joint synchrony. By looking towards each other just before moments at which they converged, in synchrony, with

harmonic relationships, perhaps also breathing together as they start a new phrase, it appears that our participants focussed primarily on co-ordinating these moments, aligning their sounds and intentions in time.[136] Thus, the process of *grounding* in this context seems to concern synchrony, co-ordination, and a convergence of intentions within the interaction itself - rather than sharing understanding about anything outside.

On the other hand, where the interaction involves a clear goal (in this case, to depict birds), it resembles more closely the turn-taking structure characteristic of speech (Sacks et al., 1974). Even more precisely, the pattern of gaze aversion-gaze checking, before and after turn onset, which we see here is also common in spoken exchanges (Kendon, 1967). Mutual gaze occurs in the cross-over between turns (as also shown in e.g. Bavelas et al., 2002). This might indicate that the grounding which occurs here is analogous to what happens in speech - even though the only form of reference we have in this experiment is to depict a very general scene, sharing just the broad focus on *birds*. Even here, just with this vague external focus, we have primed a more extrinsically-oriented form of interaction - something more speech-like rather than musical. Previous work has found that gaze aversion might occur in this turn-taking context in order to reduce cognitive load - in order to re-direct attentional resources away from observing the other to forming the next utterance (Doherty-Sneddon & Phelps, 2005). Checking back is then important to ensure that the other is attending, and to register their back-channelling responses - which indicate that they have understood (Bavelas et al., 2002). In this musical context, it is unclear what specific meaning could be part of this shared understanding. Perhaps, we see this gaze pattern here because it is the expected or automatic form which structures our interaction - a structure we would usually need when external goals and references are introduced.

The patterns exhibited in the elephant-goal condition were a mixture of these two types - perhaps because the goal was less successful in its own right, leaving the participants to pursue more relational goals for a larger part of the interaction (which was, as a result, more intrinsically-focussed). Some of the same turn-taking pattern of gaze aversion and checking was displayed; this differed to the bird task, as participants generally played simultaneously (although I still considered these turn-taking bouts, given that each participant "came in" with a contribution that was distinctive, introducing new melodic material each time). At other points, there were more periods of synchrony: participants did not contribute separate, or even overlapping, turns, but rather kept time together, often in iconic "marching" behaviour. Even though this was presumably designed to resemble our external elephant goal, the gaze patterning was similar to our no-goal case - gaze

---

[136] Compare an observation made by Moran (2013) in her video analysis of duo improvisation in a North Indian classical tradition: moments of eye contact occurred at the *sam* - the point of coherence at the first beat of the metric cycle.

(including mutual gaze) occurred before or at broad points of convergence, again perhaps to co-ordinate their synchronisation. (In the bird condition, there were a couple of shorter periods of simultaneous playing where a similar pattern was displayed.)

If the nature of the common ground is indicated by gaze patterns, then it seems that the no-goal condition (or the elephant-goal condition, where the goal was harder to achieve) involved shared focus on co-ordination, i.e., a relational focus. In the bird-goal condition, the gaze pattern is analogous to what occurs in speech - behaviour usually oriented towards mutual understanding of something external. This provides initial evidence that there is a real, behavioural difference between interaction modes according to their goal - namely, whether the exchange is more intrinsically- or extrinsically-focussed, and thus involves common ground which is oriented internally or external to the interaction.

Mutual gaze also signals a moment of grounding even where the interaction frame breaks down - as in instances of mutual laughter, disrupting the flow of the interaction. On these occasions, participants look towards each other (as expected), achieve eye contact, but then the interaction rapidly deteriorates. This laughter again seems to be a point of grounding - a moment of mutual recognition reached at that moment of eye contact. Laughter might index positive affect, perhaps as participants appreciate the humour in the incongruence of playing a kazoo in the University Faculty of Music. However, laughter is also generally a response to embarrassment, as a way of saving face in those contexts. The task to make music, playing a slightly silly musical instrument, with strangers - usually a situation in which one tries to make a good first impression - might well have contributed to a feeling of self-consciousness and embarrassment. Laughing together might have been a joint effort to protect the face of the other, jointly acknowledging the situation while working together to diffuse or mask any negative feelings (see discussion in Chapter 2). Interestingly, this joint laughter occurred only in the elephant- and no-goal conditions. These were the more intrinsically-oriented of the three tasks; also, however, they were the cases in which there was a mix of genders (whereas both participants were female in the bird condition). The presence of someone of the opposite sex might well have increased the social pressure, perhaps making such frame-breaks, and moments of mutual embarrassment, more likely. Together with my experimental findings, this suggests an important role of embarrassment (or of motivations to protect, perhaps enhance, self-esteem) in managing or evaluating these social relationships.

I also found a strong disparity in the amount of joint synchrony displayed. In particular, the no-goal condition exhibited much more synchronous playing - perhaps a factor in the higher social connectedness emerging following this condition (compared with the goal conditions) in experiment 3. This is unlikely to explain completely the differences in interpersonal affiliation shown in our

previous experiment: while the video analysis revealed very few bouts of joint synchrony in the elephant goal condition, the degree of affiliation following this task (in experiment 3) fell somewhere in the middle (between the no-goal and bird-goal tasks). It seems more likely that the effects of interaction on social bonding relate *both* to the degree of synchrony and to the differences in mode or goal focus. Of course, a more relational goal itself involves a greater emphasis on joint synchrony; this is clearest in the no-goal condition. These findings also correspond to the effect of musical interaction on group bonding (compared with a non-musical task, in experiment 1), and the effect of the no-goal condition on connectedness to members of the *other* team (in experiment 2).

The work of this chapter has made some substantial, if exploratory, contributions to our understanding of interaction mode, of the nature of musical interaction, and the effect of interaction goal focus on the way we relate to others. This also sheds light on the types of social relationships we establish - whether a sense of interpersonal or group affiliation, or a new team identity (which might even exclude others who do not belong). My conclusion will bring all of this work together, providing a summary of what I have shown so far, and our next steps in this new, wider project.

Conclusions

Making music is an activity that we do together. For the duration of that performance, we focus our efforts at co-ordinating with each other, losing ourselves in the task to become more aligned with those around us, sometimes striving to perform something well, or correctly. It is a powerful way in which we assert a group identity - whether as believers in a church congregation or supporters of a football team. We do not just reflect that group identity, but we enhance, perhaps even enact it, in that performance.

Belonging to a social group and forming meaningful relationships with those around us is crucial for our wellbeing. We interact with people time and time again, building social networks and relationships of trust. Our group identity comes from our affiliation with others, through interdependent relationships - working on something together, making music, or holding a conversation. This process, however, is not without its limits: our new group identity does not include everyone. For an in-group to exist, it must have a boundary, dividing those who belong from those who do not.

This thesis has been about the intersection of these spheres - how we interact through music, and how this structures our social relationships. I have pulled together many different fields of study, constructing a theory and proposing some specific hypotheses about the role of interaction in group formation.

Instrumental music, speech, song, and even dance all stem from our capacity to relate to and understand others, and meet our need to belong and secure lasting connections. I have explored the nature of these everyday interactions: we might interact to get something done, to communicate some information about something outside our relationship; at other times, our primary goal is to relate to each other, to share a sense of intentionality and affiliation that does not require external knowledge of a common ground. I have also explored the nature of our group identities: we bond with others through face-to-face engagement, or interdependent relationships; however, these occasions are often shaped by judgements we have made about others before we are even introduced. Their appearance - looking or sounding different to us, perhaps with a different skin colour, strange clothes, or a strong accent - can lead us to categorise them as "other", making us less inclined to engage with them; however, if we see them regularly, or find that we need to work together, we can break down these barriers.

The way that we interact is crucial in managing, reinforcing, and enacting these group bonds and boundaries - building relationships with some, but excluding others who do not belong. I have

proposed that it is the goal of our interactions, and the prerequisite common ground, which predict whether our encounter brings people together and encourages them to participate, or whether it constructs boundaries between who can and who cannot be involved. More musical forms of engagement, which have an intrinsic orientation, being more about our shared participation than anything beyond this, are powerful in bonding us together through relationships of interdependence. However, where we speak different languages (whether in speech or in other, more prescribed forms of music-making), we find it much harder to engage: interaction with an extrinsic focus, where the goal is to talk about something, relies on a common code, shared access or understanding, and is thus limited to a few, select others. This reinforces the boundaries we construct through categorisation of ourselves and others.

I have described many examples of this in action, from playing games, singing football chants, to everyday camp-fire songs and discussion groups. However, I also adopted the task of attempting to provide empirical data for this framework. Substantiating my theory calls for many experiments - many more than what is possible in a thesis of this length; I devised a methodological paradigm and used this to conduct an initial exploration.

In my first experiment, I explored how musical interaction contributes to group formation, as measured by ratings of interpersonal connectedness. Improvised, simple music-making was more effective than an interactive, visual task. There were no effects of group members sharing, not sharing, or even not having an external goal: group formation does not rely on the presence of an explicit shared goal - in fact, having a goal sometimes seemed to distract from the self-other bonding process. In my second experiment, I corroborated and extended these findings. In this case, I compared musical interaction with and without external goals. Where my participants had no additional goal beyond the task to interact, they subsequently felt more connected to those assigned to an opposing team (where their different team identities were reinforced by corresponding musical goals, they instead felt a stronger affinity to their team-mates). Again, in experiment three, the no-goal interaction task resulted in higher feelings of affiliation towards one's partner than when the interaction involved some external, shared goal. Thus, musical interaction, and especially music-making in which the performers are focussed upon each other rather than on something else, can bring and bond people together more easily than other tasks. Music foregrounds the phatic, the relational components of interaction, whereby individuals can co-ordinate their actions within a framework of floating intentionality without relying on shared meaning or external common ground.

Forms of interaction with more salient, external goals might instead contribute to processes of categorisation. I introduced this question in my second experiment, assigning participants to different teams and then measuring how far the interaction task enhanced their identification with

that team, specifically. Musical interaction with a corresponding depiction goal enhanced connectedness towards others who shared that team affiliation, promoting intergroup bias between the "birds" and the "elephants". This effect emerged only as a function of the success of that interaction task. In experiment three, I clarified this relationship. Using an implicit measure, I showed that participants felt stronger team identification when they thought the interaction was a joint success - or rather, was less embarrassing. When they were more embarrassed, or when they felt they were more personally responsible for the success of the task, they were less likely to incorporate this team identity into their self-concept. This has parallels with Social Identity Theory, specifically the role of self-esteem. In defining oneself in relation to a particular category, we seek enhanced self-evaluation through the positive distinctiveness of our group or category (see Turner et al., 1987). Moreover, Goffman's work on maintaining 'face' in interaction is relevant here. He emphasises the importance of maintaining 'positive social value', through regulating the flow of interaction, i.e. 'face-work' (Goffman, 1967: 5). A social situation which is unsuccessful, or embarrassing, causes the individuals to distance themselves from the associated category or team as a face-saving strategy. And perhaps this is the case vice versa: positive interaction that is associated with a particular team identity is likely to result in adoption of that social identity, driven by motivation to enhance self-esteem.

The results of these three preliminary experiments are presented in Tables 39 and 40. Overall, they make a start at identifying the components of musical (and other forms of) interaction which structure our social relationships - bonding us together (via affiliative interdependence), and dividing us from others who are different (through our sense of team identity).

Table 39 *Experimental results showing factors which might contribute to group or interpersonal bonding*

Experiment	Variable	Measure
1	Music > Visual	Ratings of connectedness
	No goal > (Non-shared) goal	IOS ratings
2	No goal > Goal	Connectedness to other-team members
3	Elephant/no goal > Bird goal	Partner trust, connectedness, similarity

Table 40 *Experimental results showing factors which might enhance team-related affiliation, or a sense of team identification*

Experiment	Variable	Measure
2	(Bird) goal > No goal	Connectedness to own-team members *more*
	(Bird) goal > No goal	IOS ratings for team
	Goal success	Connectedness to own-team members
3	Embarrassment (-ve)	Team-related word recall

My video analysis explored in more depth the behavioural correlates of interaction with different goal orientations - again, with or without an external depiction goal. This case study offers some insight into the possible roots of what I measured elsewhere using scales of connectedness and other indicators of team identity. The no-goal interaction task, i.e. our more intrinsically-oriented condition which had resulted previously in higher levels of affiliation, involved much more synchronous playing than either of the two goal conditions. Furthermore, different patterns of individual and mutual gaze suggested distinct grounding processes. On one hand, participants given the no-goal task focussed on mutual co-ordination, achieving alignment at important moments in the musical phrase. On the other hand, the pair with the bird goal (perhaps the more achievable of the two goals) engaged in a speech-like manner, looking away to process the "information" conveyed, and to form a turn, and then checking to see if the other had "understood" something. Furthermore, mutual gaze enacted alignment in a different sense - this was often the moment of breakdown, when the exchange disintegrated into laughter. Goffman described this very situation in his discussion of embarrassment; I argue here that this emotion, and our motivation to save face in the light of this, is yet another form of grounding. It concerns our immediate relationship with another, but also affects how we feel about the social identity we might represent together.

Future directions

This work has implications not only for our understanding of social group formation, and the complex role of music-making in establishing our communities, but also for our understanding of interaction more generally. The crucial difference between interaction which is more "speech-like" and forms which are more "musical" might be one of goal orientation, a different type of shared focus - or perhaps a different type of common ground. I have demonstrated here a novel, but simple, way of modelling those interaction types, so we can use the controlled setting of a psychology laboratory to understand more those processes which we see every day in "real life".

Further research should not only attempt to replicate the results reported here, perhaps with larger and more diverse samples (to tackle the problem identified by e.g. Henrich et al., 2010), but could

also use different, perhaps more sensitive or reliable measurement techniques. Eye tracking and motion sensor technology could be used to analyse a larger corpus of behavioural data; physiological measures could detect smaller changes in affect or mood over the course of the interaction task; a more reliable interpersonal memory test (similar perhaps to Woolhouse et al., 2016) could assess more accurately whether attention towards others varies as a result of shared or conflicting goals (a condition which seemed to have no effect in experiment one, but which appears to be significant from my other results). Further research could ascertain whether the results of my implicit team identification measure (memory for team-related words) are reliable - and whether they do indeed tap into a sense of identity, rather than mere concept salience. Moreover, more data collection could clarify the relationship between joint success and embarrassment, and demonstrate how these are related to a sense of team identity (and also to interpersonal connectedness, more generally).

Research could employ the same experimental paradigm to investigate other factors which might shape group formation, or the nature of the interaction: for example, the number of people in a group, its gender composition, and the relationship between gender and other social or team divisions. Does gender identity overshadow any alternative social categorisation? Furthermore, how does goal orientation shape the nature of the interaction when it includes three rather than two participants? What happens if they have conflicting depiction goals? Further work is needed to clarify the nature of embarrassment in this context - what other behavioural correlates are there, beyond mutual gaze and laughter? Is it more likely where there is a mix of genders? Do friends show a different response, compared to the unacquainted individuals recruited here? Furthermore, more clarity is needed regarding how embarrassment and interaction goals each relate to the formation of social identity, and a cohesive social group. Can the interaction itself build a team identity without it being established at the start of the experiment? And are there other types or components of interaction, beyond a depiction goal, which might create common ground in an experimental context? Finally, beyond the role of embarrassment, my work has not seriously probed the role of emotion, or affect. We know that music is a rewarding activity - but how exactly does its effect on our mood shape the personal relationships it enacts? We know that interaction varies in goal orientation - but our emotions (including embarrassment) supersede this divide: they affect our relationships, but in ways tied to our joint experience of the world, our successes and failures as an individual and as a group.

This thesis has answered some important questions - and it has raised many more. It has presented a theoretical framework, and a research methodology, which makes it possible to explore this area with new impetus. This is a time of new and rising social divisions. We need to understand better how we relate to others, and how we can overcome those barriers which divide us.

Appendices

Appendix 1 - Experiment outcome measures

1.1 Experiment 1

The outcome measures of group bonding are listed in Table 41 below, in order of presentation:

Table 41 *Outcome measures of group formation used in experiment 1*

Outcome	Measure	Reference
In-group trust	Forced-choice responses to various scenarios (half reverse-scored); choice between specific individuals who belong to the in-group, or the out-group (photographs shown; randomised order of presentation).	Scenarios adapted from the Specific Interpersonal Trust scale (Johnson-George & Swap, 1982; taken from Wrightsman, 1991)
Experience of the group interaction task	Free written response; participants asked to describe their experience, level of enjoyment, and how successful they felt they were at working together.	N/A
Memory for group member attributes	Asked how many members (a) were wearing glasses, (b) have blonde hair, and (c) were wearing a pink wristband	Similar to the test administered in Woolhouse et al. (2016)
Connectedness	7-point Likert scale, 'How connected did you feel with the other participants in your group during the activity?'	Wiltermuth & Heath (2009)
IOS scale of bonding of the individual to the group	5-point pictorial likert scale	Swann et al. (2009)
Successfulness	7-point Likert scale: 'How successful do you think your group were at the task?' 5-point Likert scale: 'I found this task difficult' 5-point Likert scale: 'I think I performed well on this task'	N/A
Entitativity	Three 5-point pictorial scales of entitativity, similarity, and interaction	Rutchick et al. (2008)

Team perception	7-point likert scale, 'How far did you feel you were on the same team with other participants in the task?'	Wiltermuth & Heath (2009)
General trust in group	5-point Likert scale: 'I could rely on the people in my group.' 5-point Likert scale: 'I had confidence in the others in the group' 5-point Likert scale: 'Overall, I think the people in my group are generally very trustworthy'	N/A

Other filler items were included: rankings of interaction with other individual group members (rank 1-3 if interacted with them most to least); a social word completion game was also included (taken from Knowles & Gardner, 2008).

Pictures used in depiction goal instructions

Figure 16 shows the pictures used for the depiction tasks in the musical and non-musical goal conditions.

(a)   (b)   (c)   (d)
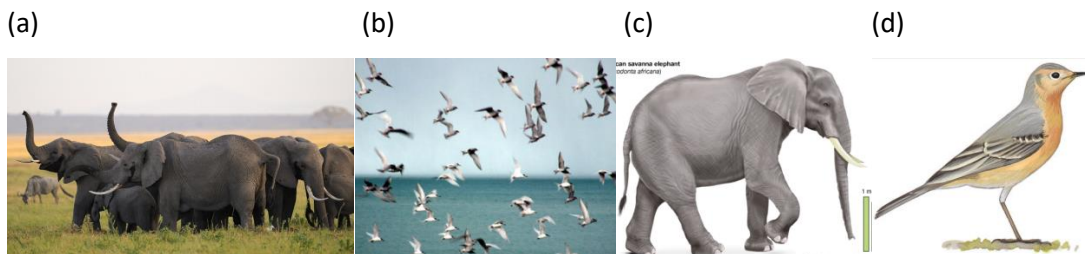
Figure 16 Pictures used for depiction goal tasks (a) the elephant goal in the musical condition, (b) the bird goal in the musical condition, (c) the elephant goal in the non-musical condition, and (d) the bird goal in the non-musical condition. Pictures (a) and (b) were also used in the goal conditions of Experiment 2.

1.2 Experiment 2

Two different types of dependent variable were used in this experiment – some measuring group formation (towards the interaction group, i.e. mixed teams - see blue text), and others measuring category or team formation (towards specific team members both from within and outside the interaction group - see red text). These were administered in a post-interaction questionnaire. The main items are shown in Table 42 below, in order of presentation:

Table 42 *Outcome measures of group formation and team-related affiliation or identification in experiment 2*

Dependent variable	Measure	Reference
Specific trust preference (own-team members vs. other-team members)	Forced-choice questions for a number of different scenarios. Participants chose between photographs of own-team and other-team members (some familiar from the interaction, and others unfamiliar i.e. not involved in the experiment). Each photograph was associated with own or other team symbols (randomised order of presentation; half items reverse-scored).	Scenarios adapted from the Specific Interpersonal Trust scale (Johnson-George & Swap, 1982; taken from Wrightsman, 1991)
Experience of the group interaction task	Free written response: describe experience, level of enjoyment, success at working together across the two teams.	N/A
Connectedness for own and other team members (in interaction group)	Two 7-point Likert scales, 'How connected did you feel with members of your team/the other team during the activity?'	Wiltermuth & Heath (2009)
Rating of successfulness (of the group interaction)	7-point Likert scale, 'How successful do you think your group were (as a whole) at the task?'	N/A
Entitativity (ratings of the group as a whole)	Three 5-point pictorial scales of entitativity, similarity, and interaction	Rutchick et al. (2008)
IOS scale of bonding of the individual to their team as a whole	5-point pictorial Likert scale, associated with their team symbol	Swann et al. (2009)

Note: measures of group formation are in blue text and team-related affiliation or identification in red.

1.3 Experiment 3

Measures in the questionnaire, in order of presentation (measures relating to team identification are shown in red and to partner affiliation in blue), are detailed in Table 43.

Table 43 *Outcome measures of partner affiliation and team identification in experiment 3*

Dependent variable	Measure	Reference
Participants were shown a picture of their team symbol, and asked to 'think about your team – the bird/elephant team – as a whole'		
IOS scale of bonding of the individual to their team	'Please select the picture which best represents your relationship with your team.' 5-point pictorial likert scale	Swann et al. (2009)
Reward allocation to unfamiliar team members	'Below are pictures of other members of the bird and the elephant teams who have taken part in the experiment already. Their team is shown by the symbol. Imagine that you have 4 pots of £5 in £1 coins. ... Decide how many £1 coins to allocate to each participant and write a number in each box. This must total £5 on each line.' Forced-choice decision to allocate unequal rewards to unfamiliar individuals associated with their own and other team symbols. Photographs and team membership were counterbalanced.	Similar to Hetherington et al. (2014) and Locksley et al. (1980)
Participants were asked to 'think back to the musical task you did with your partner.'		
Experience of the interaction task	Free written response: 'describe your experience of the task, how much you enjoyed it, and how successful you think you were as a pair.'	N/A
Trust in/reliance on partner	5-point Likert scales: 'I think I could rely on my partner' 'I had confidence in my partner' I think my partner is generally very trustworthy'	N/A
Connectedness to partner	7-point Likert scale: 'How connected did you feel with your partner during the activity?'	Wiltermuth & Heath (2009)
Similarity to partner	7-point Likert scale: 'How similar do you feel to your partner?'	Rabinowitch & Knafo-Noam (2015)
Partner likeability	9-point Likert scale: 'How likeable is your partner'	Hove & Risen (2009)

Note: measures of partner affiliation are in blue text and team identification in red.

Amongst the measures above, they were also asked about their perceived performance on the task (5-point Likert scales: 'I found this task difficult' and 'I think I performed well on this task'), how successful they thought their interaction was (7-point Likert scale: 'How successful do you think your pair was at the task?') and how embarrassed they were (7-point Likert scale: 'How embarrassed were you during the task?'; see e.g. Sabini et al., 2000).

Appendix 2 - Experiment 1 follow-up

A short follow-up experiment after experiment 1 tested further whether there were any experimental artefacts arising as a result of the particular nature of the music task. The musical interaction might have caused greater connectedness (in comparison to the non-musical, visual interaction task) for several different reasons.

First, musical interaction might foreground phatic, interpersonal dimensions more than when there is a more explicit focus on a visual goal. Both forms of interaction relied upon non-verbal interaction of some sort, but the music condition, with its affordance for sounded gestures and melodic contour, made the interaction itself (rather than something external) the central focus.

Second, the musical task was much more flexible: there were no real constraints on the notes, or on how the interaction was to proceed. The visual task enabled some freedom, but moving shapes afforded much less improvisation and had a possible "finish" point. This freedom might have helped or hindered the relationship between participants - in the former by enhancing moment-by-moment attention, or by making the interaction more risky and potentially embarrassing in the latter.

Third is the possible role of positive affect. While music generally results in positive affect (Dunbar et al., 2012), some elements of the music condition particular to this experiment (and not more generalizable as effects of music-making) might have prompted positive behavioural responses (e.g. laughing together) resulting in enhanced connectedness (see Gervais & Wilson, 2005). For instance, the musical instruments produced a buzzy and humorous timbre; also, the depiction tasks were unusual and perhaps more light-hearted than other forms of participatory, improvisatory music.

In order to investigate which attributes of music-making might have contributed to group connectedness, I ran a smaller-scale case study experiment, manipulating certain aspects of the music interaction task. By restricting the melodic content (making the interaction more scripted), or by using an instrument limited to the pentatonic scale, I investigated the potential effect of interaction freedom. These are compared to a free improvisation task, using the same simple vocalising musical instrument (kazoo) as was used in the original experiment. This set-up also tests

whether bonding only occurs as a result of the peculiar characteristics and sounds of the kazoo (via group laughter), or whether it might be a result of other aspects of the music-making.

Method

Participants

A total of 9 participants took part in this case study (8 female, 1 n-b), all aged between 20 and 24 years (mean = 20.89). All were musically trained (recruited from the music faculty). They all provided informed consent to take part in the study, which was reviewed by the Faculty of Music Research Ethics committee.

Design

This was an exploratory, between-groups design, varying the limitations and modality of the interaction condition on three levels: (1) free improvisation using simple vocalising instrument (kazoo), (2) improvisation around a melodic 'anchor', the well-known tune of 'Happy Birthday', using the same instrument and (3) free improvisation using a pentatonic percussion instrument (balafon). This simultaneously measured any specific effect of the kazoo (being an instrument which might enable social bonding merely through laughing together), while testing the effect of limiting the freedom of interaction, either on the basis of a melodic structure, or limiting the notes available in a consonant, pentatonic framework.

The number of participants was split equally, with three assigned to each condition. Participants interacted in groups of three (all group members assigned to the same condition)

The dependent variables included various measures of social bonding to the interaction group as a whole, and towards specific members of the group. These were administered in a questionnaire. Table 44 below summarises the items used, in order of presentation.

Table 44 *Outcome measures of individual and group bonding*

Dependent variable	Measure	Reference
Experience of the interaction task	Free written response: 'describe your experience of the task, how much you enjoyed it, how comfortable you were, and how well you think you all worked together'	N/A
Connectedness to group/group members	7-point likert scale: 'How connected did you feel with the other participants in your group during the activity?' They were also asked how connected they felt to the others individually (identified using letter A, B or C on badge)	Wiltermuth & Heath (2009)
IOS scale of bonding of the individual to their group	'Tick the diagram which best represents your relationship with the group.' 5-point pictorial likert scale	Swann et al. (2009)
Entitativity	Three 5-point pictorial scales of entitativity, similarity, and interaction	Rutchick et al. (2008)

Participants were also given 5-point Likert scales to report how far they could rely on the people in their group, how difficult they found the task, how well they think they performed, their confidence in the others, and how generally trustworthy their group is. They were also asked about their need to belong (using the single-item Need to Belong test – see Nichols & Webster, 2013). A social word completion game was also included (taken from Knowles & Gardner, 2008). There were 7-point Likert scales on how successful they thought the group as at the task, how embarrassed they felt (see e.g. Sabini et al., 2000) and whether they felt they were on the same team as the others (Wiltermuth & Heath, 2009), as well as 9-point scales on how likeable the other members are (as a group, and individually – see Hove & Risen, 2009). They were given a simple memory test (eye colour of each other group member, and whether they were wearing glasses during the task). The short form of the PANAS mood questionnaire was administered pre- and post- interaction task (Kercher, 1992; reviewed by Mackinnon et al., 1999). Participants were also asked to rate (and detail, if necessary) how much they knew the two other group members (two 7-point scales).

Procedure

Participants took part in groups of three. They were each given a badge (A, B, or C) to identify themselves (for the questionnaire administered later). They were given the following instructions, according to their assigned condition.

Free improvisation – 'Your task is to make music together, using only the kazoos. Improvise together, listening and co-ordinating with each other.'

Improvisation around melodic anchor – 'Your task is to make music together, using only the kazoos. Improvise on and around the tune of Happy Birthday. Work together, listening and co-ordinating with each other.'

Improvisation in pentatonic framework – 'Your task is to make music together, using only the balafons. Improvise together, listening and co-ordinating with each other.'

All groups were told not to talk or otherwise communicate, and were given five minutes for this interaction task. They were filmed during this interaction task, and their individual contributions recorded using microphones.

Immediately after the interaction task, participants were seated individually and asked to complete the questionnaire. They were then thanked and debriefed.

Results

Non-parametric independent-samples Kruskal-Wallis tests were run according to interaction condition. The only main effect was on ratings of connectedness towards the group as a whole: $H(2) = 6.095$, $p = .047$. There were significantly higher ratings of connectedness in the free improvisation condition compared to the melodic anchor condition: $H(2.160) = 5.333$, $p = .041$ (see Figure 17 below). Ratings of connectedness were not correlated with any other measure.
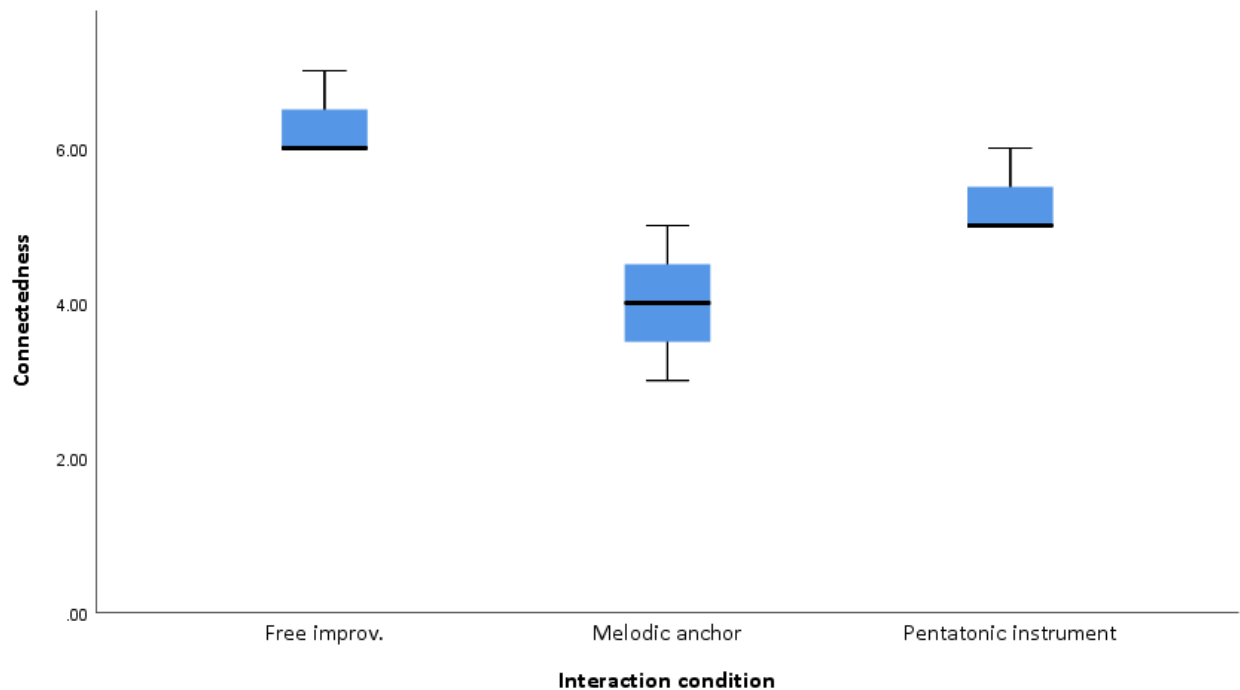
Figure 17 Boxplot showing the effect of interaction condition on ratings of connectedness with the group. Connectedness was significantly higher in the free improvisation compared the melodic anchor condition (plot shows median, interquartile range, and top and bottom 25% of ratings).

Baseline mood and prior knowledge

There were no significant differences between the three conditions for baseline mood (using PANAS self-report scale before interaction), nor were there significant differences in level of prior knowledge amongst the participants in each group.

Similarly, the ratings of connectedness were not correlated with ratings of prior knowledge of the individual group members. Ratings of connectedness towards individuals were also not correlated with prior knowledge; nor were they correlated with change in mood pre- and post-interaction.

Discussion

This outcome enables us to exclude the possibility that increased connectedness resulted only from specific task characteristics (the humorous nature of the kazoos): this instrument was used in both free improvisation and the melodic anchor condition. In contrast, higher connectedness resulted following the conditions with a greater degree of melodic flexibility, supporting the suggestion that it is the freedom of the improvisatory musical task which contributes to group bonding. (Instead of the possibility of having to overcome embarrassment in this condition, there was actually a trend towards higher embarrassment for the melodic anchor condition, although this did not reach significance.)

In summary, this empirical work demonstrates that five or ten minutes of very simple music-making results in a sense of group connectedness amongst adults, in comparison to a non-musical depiction task. This could be due to the multimodal, phatic nature of the musical task, or due to the greater freedom enabled by an improvisatory context.

Appendix 3 - Video analysis supplementary materials

Videos can be made available on request.

3.1 Task instructions

The instructions given for each task are detailed in Table 45.

Table 45 *Task instructions for different interaction conditions in video analysis study*

Interaction goal	Instructions
Easy goal - bird task	'make music inspired by a flock of birds, using only the kazoos. Improvise together, listening and co-ordinating with each other. In your joint improvisation, you could allude to e.g. a variety of bird calls, their manner and/or speed of movement, interaction with potential mates, wings flapping and their protection of territory, amongst other things.'
Hard goal - elephant task	'make music inspired by a herd of elephants, using only the kazoos. Improvise together, listening and co-ordinating with each other. In your joint improvisation, you could allude to e.g. trumpeting sounds, their manner and/or speed of walking, interactions with their calves, ears flapping, and the sound of drinking/spraying water, amongst other things.'
No goal	'make music, using only the kazoos. Improvise together, listening and co-ordinating with each other.'

3.2 Interaction set-up and data processing

The configuration for video recording the interaction condition is shown in Figure 18. Participants (and the direction they were facing) are shown by the grey squares. Cameras are denoted by the large black arrows. Participants were also filmed from above using a GoPro camera. The microphones used were AKG 414 ULS microphones set to cardioid (picking up sound from one direction, and attenuating it from other directions), held in position using stands, near to each participant, but not obstructing their interaction with each other.
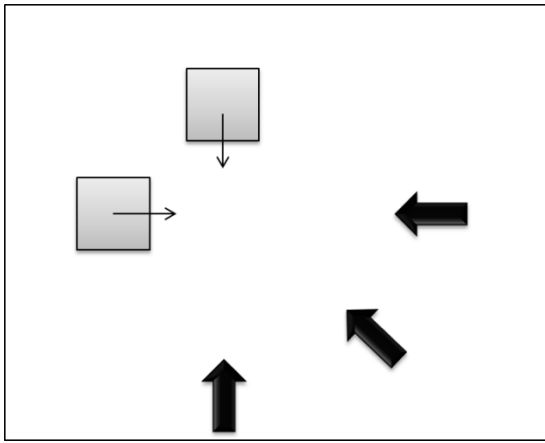
Figure 18 Diagram representing room layout for video recording (not to scale)

The audio tracks and video footage were synchronised using Final Cut Pro X. From this, a multi-cam clip, including the footage of direct gaze for each participant, was created for each dyad. Annotations of gaze behaviour were completed using ELAN.[137] Annotations were created frame-by-frame to show the start and end of gaze towards the face region of the other, within a precision corresponding to the frame-rate of 25 frames per second. The overlaps between the gaze annotations for each individual in a dyad were defined as periods of mutual gaze.

Analysis of the audio track was completed using Praat.[138] Audio tracks were recorded separately for each participant. Note onsets were defined as the beginning of an amplitude increase. Note changes within a continuous phrase were also marked as separate onsets, corresponding to detected changes in frequency. These markers were imported into ELAN to be compared to the video annotations. Inter-onset intervals for each note were calculated automatically as the distance between each marker.

3.3 Bouts of joint synchrony

3.3.1 IOIs - means and standard deviations

Table 46 documents the mean IOI for each period of joint synchrony, with the standard deviation about that mean, in each interaction condition. Bouts of joint synchrony are not necessarily the same length. The mean of each of these (plotted in the graphs in Figure 14) is found at the bottom.

---

[137] A linguistics software developed at the Max Planck Institute, retrieved February 2019 from https://archive.mpi.nl/tla/elan.
[138] This was developed by Boersma and Weenick, and retrieved from http://www.praat.org in March 2019.

Table 46 *Mean and standard deviation IOI for each period of joint synchrony, in each interaction condition, with overall mean (and mean SD) at bottom of each column*

Bird		Elephant		No goal	
Mean IOI (s)	SD IOI (s)	Mean IOI (s)	SD IOI (s)	Mean IOI (s)	SD IOI (s)
1240.7	153.3	808.2	80.1	854.6	68.0
1694.5	211.0	1271	116.3	921.6	124.0
1221.9	73.4	1021.4	96.3	921.2	182.8
882.4	114.7	1233.3	77.6	997.8	102.0
1092.7	169.3	1044.5	75.9	905.9	79.7
1201.6	124.8	1050.1	104.8	898.2	116.1
1331.2	201.1	1115.2	165.8		
984.5	98.7	603.5	61.4		
1203.4	140.0	970.5	162.6		
883.3	74.7	1418.8	124.8		
1294.7	98.6	843.5	172.3		
955.8	149.6	1431	185.7		
		821.8	61.3		
		1229.4	146.1		
		1000.5	76.5		
		967.7	108.5		
		1113.2	48.1		
		864.3	63.4		
1165.6	134.1	1044.9	107.1	916.5	112.1

3.3.2 Bouts of synchrony - coder agreement

Table 47 shows the bouts of joint synchrony extracted in each interaction condition - when they occurred, the average beat IOI and standard deviation of the beat IOIs (in ms), and how long the bout lasted. This also shows the bouts of joint synchrony coded (in one sitting) by two independent coders. Boxes highlighted in red show a bout which was identified by two coders, or by the experimenter and one coder; rows in yellow show bouts for which there was agreement by all three.

Table 47 *Bouts of joint synchrony in each interaction condition*

Bird interaction condition					Coder 1			Coder 2		
Start of bout	End of bout	Mean IOI	SD IOI	Bout length	Start	End	Success?	Start	End	Success?
00:48	00:53	1241	153	00:05				00:49	00:55	N
00:59	01:06	1695	211	00:07						
01:06	01:21	1222	73	00:15	01:08	01:32	Y			
01:22	01:28	882	115	00:06						
01:32	01:40	1093	169	00:08	01:32	01:48	N			
01:51	02:06	1202	125	00:15	01:48	02:08	Y			
					02:08	02:19	N			
02:19	02:27	1331	201	00:08	02:19	02:30	Y	02:19	02:25	N
02:31	02:38	985	99	00:07	02:31	02:41	Y			
					02:41	02:52	N			
					02:52	03:14	N			
					03:31	03:45	Y			
					04:01	04:04	Y			
04:04	04:14	1203	140	00:10	04:04	04:25	Y			
								04:17	04:24	Y
					04:36	04:41	N			
					04:51	04:57	Y	04:51	04:57	Y
					04:57	05:04	N			
05:08	05:16	883	75	00:08	05:09	05:14	Y	05:08	05:16	Y
05:21	05:25	1295	99	00:04						
05:27	05:34	956	150	00:07						
Elephant interaction condition										
00:05	00:15	808	80	00:10						
00:23	00:30	1271	116	00:07						
00:30	00:34	1021	96	00:04	00:29	00:34	Y	00:30	00:33	Y
00:37	00:42	1233	78	00:05	00:39	00:43	Y			
00:46	00:50	1045	76	00:04						
00:51	00:55	1050	105	00:04						
01:08	01:24	1115	166	00:16						
01:39	01:46	604	61	00:07						
01:50	02:10	971	163	00:20						
02:11	02:26	1419	125	00:15						
02:26	02:30	844	172	00:04						
02:46	02:57	1431	186	00:11						
03:12	03:16	822	61	00:04						
03:18	03:30	1229	146	00:12						
03:31	03:35	1001	76	00:04						
03:51	04:07	968	109	00:16						
04:16	04:25	1113	48	00:09						
04:40	04:52	864	63	00:12	04:42	04:58	Y	04:41	04:45	Y
No-goal interaction condition										
00:11	00:18	855	68	00:07	00:12	00:26	N			
00:34	01:42	922	124	01:08	00:26	01:47	Y	00:32	02:08	Y
01:54	02:07	921	183	00:13	01:47	02:07	N			
02:16	04:27	998	102	02:11	02:10	04:26	Y	02:10	04:24	Y
04:46	04:54	906	80	00:08	04:46	05:12	N			
05:01	05:32	898	116	00:31	05:12	05:34	Y	05:00	05:32	Y

Note: red shading: this period of joint synchrony was identified by the experimenter and one coder, or by both coders; yellow shading: this period of joint synchrony was identified by all three.

3.4 Gaze behaviour - no-goal condition

Table 48 shows all of the points of convergence, and the gaze behaviour which occurred within one second before and after this point. Here, synchrony is the deviation between the note onsets of each participant; gaze before shows the onset of a gaze which either continues or finishes within one second before the point of convergence, and gaze after includes the onset of gaze which occurs within one second after the point of convergence. Both participants are included in the gaze analysis (black = AB and blue = CL). Most points of convergence are preceded by gaze by one or both participants; the rows highlighted in grey are points of convergence which don't adhere to the pattern.

Table 48 *Points of convergence and associated gaze behaviour in no-goal condition*

Time	Synchrony (ms)	Pitch relationship	Phrase location	Gaze before?	Gaze during?	Gaze after?
00:16.8	14	Octave	End	00:15.7	Y	N
00:37.8	8	Octave	End	N	N	N
00:41.6	77	Octave	Start	N	N	N
00:52.8	157	Major 4-3	End	00:50.5	Y	N
				00:51.9	Y	N
00:56.7	181	Octave	Start		N	N
01:08.7	234	Fifth	End	00:59.1	Y	N
01:12.0	14	Fourth	Start	01:11.2	Y	N
				01:10.9	N	N
01:23.8	20	Third	End	01:22.1	N	N
01:25.5	61	Third	Start	01:24.9	N	N
01:39.5	120	Third	End	01:33.5	Y[139]	N
02:14.8	127	Fifth	End	02:11.6	N	N
02:16.7	4	Unison	Start	N	N	N
02:24.7	134	Third	Start	02:24.4	Y	N
02:36.3	23	Third	Start	02:35.5	Y	N
02:40.1	194	Octave	Start	02:40.0	Y	N
03:09.9	140	Octave	Start	03:03.5	N	N
03:25.6	170	Octave	Start	N	N	N
03:37.8	37	Third	End	N	N	03:37.9
03:41.8	160	Minor third	Start	N	N	N

---

[139] With slight deviation, but still towards CL

03:52.0	65	Third	End	03:51.8	Y	N
03:53.9	94	Octave	Start	03:53.7	Y	N
04:17.7	35	Octave	End	04:16.8	Y	N
04:21.8	203	Octave	Start	04:21.7	Y	N
05:04.8	99	Octave	Start	N	N	N
05:16.4	152	Octave	End	N	N	N
05:19.9	47	Third	Start	05:18.9	N	N

3.5 Gaze behaviour - elephant goal condition

3.5.1 Turn-taking

Table 49 displays the patterns of gaze aversion for turn onset and checking back afterwards. The time shown is the onset of the turn, gaze aversion documents the length of time *after* the end of the preceding gaze (if there is one), and checking back shows the length of time *before* the start of the checking gaze, if it occurs within 5 seconds of the turn onset. Again, the rows highlighted in grey display deviations from the usual pattern for turn-taking behaviour (gaze aversion beforehand, and checking afterwards) - always because they do not check back within five seconds of turn onset.

Table 49 *Turn-taking and associated gaze behaviour in elephant goal condition*

Time	Participant	Gaze aversion (ms)	Checking back (ms)
00:05.4	PC	1930	3541
00:11.3	PC	192	1796
00:16.0	PC	221	2663
00:20.1	SPC	674	1797
00:22.1	PC	1198	4151
00:24.9	SPC	372	18
00:27.7	SPC	20	189
00:28.9	PC	898	2588
00:36.2	PC	3489	4608
00:37.2	SPC	160	2163
00:43.6	PC	1413	(none)
00:44.1	SPC	(at onset) 0	242
01:03.3	PC	4404	4655
01:04.3	SPC	159	3750
01:28.4	PC	2494	3608

01:29.1	SPC	2186	2317
01:33.2	PC	61	(none)
01:34.8	SPC	1100	1446
02:32.0	PC	3309	3994
02:32.3	SPC	(none)	(none)
02:42.0	PC	583	2452
02:42.4	SPC	423	213
03:12.0	SPC	3779	2787
03:12.7	PC	3533	(none)
03:18.6	SPC	55	254
03:30.8	PC	3919	1965
03:31.6	SPC	(none)	(none)
04:14.1	PC	(none)	4517
04:26.8	PC	233	(none)
04:28.5	SPC	(none)	(none)
04:45.3	SPC	1831	(none)
04:45.4	PC	1029	(none)

3.5.2 Periods of convergence

Table 50 shows periods of convergence in the elephant-goal condition. Occasions marked * are characterised by synchrony very early on (removing the necessity for gaze to co-ordinate the moment) (PC gaze = black, SPC gaze = blue). Some of these overlap slightly with the turn-taking behaviour above (e.g. when one continues a turn, but the other joins in before they have finished).

Table 50 *Periods of convergence (with possible point of convergence) and associated gaze behaviour in the elephant goal condition*

Period of convergence		Gaze		Point of convergence?			
Start	End	Start	End	Time	Synchrony (ms)	Pitch relationship	Mutual gaze?
00:02.8	00:19.0	00:01.8	00:19.4	00:15.0	134	NC	Y
		00:13.1	00:15.8				
00:24.9	00:27.7	00:24.9	00:26.8	00:26.4	270	Octave	Y
		00:26.2	00:27.9				
00:30.0	00:33.2	00:28.5	00:32.3	00:32.0	63	Octave	Y
		00:31.4	00:32.7				
00:38.7*	00:42.3	00:39.3	00:47.4				
		00:40.7	00:42.2				
00:46.7*	00:57.4	00:55.2	00:57.6	00:56.5	183	NC	Y
		00:55.1	00:58.9				
01:03.2*	01:06.6	01:03.5	01:04.2	01:04.8	146	Third	N
01:39.1	01:54.4	01:46.3	01:48.1	01:46.5	87	Fifth	Y
		01:44.0	01:51.2				
01:56.3*	02:06.3						
02:07.2	02:30.0	02:11.5	02:14.2				
		02:06.6	02:07.0				
		02:38.5	02:42.0	02:40.7	182	NC	Y
		02:36.0	02:41.3				
03:00.0	03:11.9	02:59.0	03:01.6	03:03.8	150	NC	Y
		03:03.3	03:04.9				
		03:06.5	03:08.3	03:08.5	51	Octave	Y (before)
		03:00.5	03:09.2				
		03:23.9	03:25.4	03:25.3	39	NC	Y
		03:25.1	03:26.9	03:29.1	44	Fifth	N
		04:00.7	04:03.6	04:02.1	203	NC	Y
		04:01.0	04:03.3				
04:16.5*	04:25.6	04:18.6	04:23.5	04:23.1	70	Octave	N
04:37.1*	04:44.6	04:41.9	04:43.5				
		04:40.3	04:44.4				

Appendix 4 - A note on group selection

Human sociality depends on the capacities of individuals - we need to possess certain traits or predispositions to be able to engage with others. However, these social capacities or preferences only come to light in a group context (see Fehr & Fischbacher, 2003, who model the effects of individual co-operation on the whole population). This is especially pertinent for social preferences which are in fact potentially costly for individuals, but which benefit the group as a whole - as is often the case with prosocial or co-operative behaviour. It follows that, perhaps, in understanding the evolution of sociality or of co-operative behaviour, we should adapt our theory of selection to operate at a group level - considering the selective advantage or fitness of one group over another, rather than between competing individuals (and their traits, or genes). This theory is called group selection; in brief, it posits that social groups with norms of co-operation are likely to function more successfully than other groups, leading to the proliferation of more co-operative tendencies in subsequent generations (see original theory presented by Wynne-Edwards, 1962).

This has been the subject of much scrutiny and debate. Something like group selection emerges from models of kin selection and reciprocity; others have argued that within-group differences are far more significant, and outweigh any effect of between-group selection (see e.g. Leigh, 2010; and West et al., 2007a, and the subsequent exchange - West et al., 2007b, and Wilson, 2007). This controversy is ongoing; however, there are some new avenues which offer a productive way forward. For example, multilevel selection theory acknowledges that selective processes can operate both on a within-group and a between-group level, and that both of these offer useful insights (Wilson & Sober, 1994; see also Wade et al., 2010 - in answer to Wild et al., 2009). In conditions of high within-group conformity, for example, selection on a whole-group level might become more salient, perhaps then effecting overall shifts in altruistic behaviour on a population level (although the fitness benefits remain exclusively individual; see Wilson et al., 2008; see a computational model of this level in action in Gardner & Grafen, 2009). Another perspective is offered by niche construction theory: the group is not the unit of selection, but rather the selective environment, which itself selects or reinforces certain individual prosocial traits which are advantageous to the group as a whole. Finally, we could even consider group-level *traits* as the units of selection (rather than selecting groups, individuals, or genes).

> Group-level traits are not expressed by any single individual in the group, but emerge only from the structured organization of differentiated individuals. (Smaldino, 2014: 244)

Social behaviour - whether it is a preference for co-operation, or the display of some sort of 'collaborative interdependence' - is beneficial for groups and their individual members; this emerges only in a group context.

Traits at the level of individuals are the bedrock of human behavior and encompass the phenomenological experience of being human. But organization matters. Emergent group-level traits allow one group to outperform another, and they alter the physical and social environment, providing additional selection pressures and opportunities for new behaviors. (Smaldino, 2014: 254)

References

Aberson, C. L., Healy, M., & Romero, V. (2000). Ingroup Bias and Self-Esteem: A Meta-Analysis. *Personality and Social Psychology Review*, *4*(2), 157–173.

Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*, *60*(1), 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Amiot, C. E., & Aubin, R. M. (2013). Why and how are you attached to your social group? Investigating different forms of social identification. *British Journal of Social Psychology*, *52*(3), 563–586. https://doi.org/10.1111/bjso.12004

Anderson, B. (2006). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso (revised edition; ACLS Humanities E-Book).

Anshel, A., & Kipper, D. A. (1988). The Influence of Group Singing on Trust and Cooperation. *Journal of Music Therapy*, *25*(3), 145–155. https://doi.org/10.1093/jmt/25.3.145

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970. https://doi.org/10.1037/a0016923

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28*(2), 105–124. https://doi.org/10.1017/S0140525X05000038

Arbib, M. A. (2011). From Mirror Neurons to Complex Imitation in the Evolution of Language and Tool Use. *Annual Review of Anthropology*, *40*(1), 257–273. https://doi.org/10.1146/annurev-anthro-081309-145722

Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, *28*(3), 289–304. DOI: 10.2307/2786027

Argyle, M., & Kendon, A. (1967). The Experimental Analysis of Social Performance. *Advances in Experimental Social Psychology*, *3*, 55–98. https://doi.org/10.1016/S0065-2601(08)60342-1

Armstrong, G., & Young, M. (1999). Fanatical football chants: Creating and controlling the carnival. *Culture, Sport, Society*, *2*(3), 173–211. https://doi.org/10.1080/14610989908721852

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Aron, A., & McLaughlin-Volpe, T. (2001). 'Including others in the self: Extensions to own and partner's group memberships', in C. Sedikides & M. Brewer (Eds.) *Individual self, relational self, and collective self: Partners, opponents, or strangers* (pp. 89–108), Hove: Psychology Press.

Asch, S. E. (1956). Studies of Independence and Conformity: I. A Minority of One Against a Unanimous Majority. *Psychological Monographs: General and Applied*, *70*(9), 1-70, https://doi.org/10.1037/h0093718

Asch, S. E. (1987). *Social psychology*. Oxford: Oxford University Press.

Atkin, A. (2013). 'Peirce's Theory of Signs', in in E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. California: Stanford University Press (Summer 2013 edition, https://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/, accessed 6th August 19).

Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial Bias Reduces Empathic Sensorimotor Resonance with Other-Race Pain. *Current Biology*, *20*(11), 1018–1022. https://doi.org/10.1016/j.cub.2010.03.071

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, *211*(4489), 1390–1396.

Baily, J. (2005). So Near, So Far: Kabul's Music in Exile. *Ethnomusicology Forum*, *14*(2), 213–233. https://doi.org/10.1080/17411910500329658

Baily, J., & Collyer, M. (2006). Introduction: Music and Migration. *Journal of Ethnic and Migration Studies*, *32*(2), 167–182. https://doi.org/10.1080/13691830500487266

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556–1581. https://doi.org/10.1037/a0037737

Barraza, J. A., McCullough, M. E., Ahmadi, S., & Zak, P. J. (2011). Oxytocin infusion increases charitable donations regardless of monetary resources. *Hormones and Behavior*, *60*(2), 148–151. https://doi.org/10.1016/j.yhbeh.2011.04.008

Barsade, S. G. (2002). The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior. *Administrative Science Quarterly*, *47*(4), 644–675. https://doi.org/10.2307/3094912

Bashkow, I. (2004). A Neo-Boasian Conception of Cultural Boundaries. *American Anthropologist*, *106*(3), 443–458. https://doi.org/10.1525/aa.2004.106.3.443

Batson, C. D. (1987). Prosocial Motivation: Is it ever Truly Altruistic? *Advances in Experimental Social Psychology*, *20*, 65-122. https://doi.org/10.1016/S0065-2601(08)60412-8

Baumeister, R. F., & Leary, M. R. (1995). The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation, *Psychological Bulletin*, *117*(3), 497–529.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Report and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin Shapes the Neural Circuitry of Trust and Trust Adaptation in Humans. *Neuron*, *58*(4), 639–650. https://doi.org/10.1016/j.neuron.2008.04.009

Bavelas, J. B., & Chovil, N. (2000). Visible Acts of Meaning: An Integrated Message Model of Language in Face-to-Face Dialogue. *Journal of Language and Social Psychology*, *19*(2), 163–194. https://doi.org/10.1177/0261927X00019002001

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*(6), 941–952. https://doi.org/10.1037/0022-3514.79.6.941

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener Responses as a Collaborative Process: The Role of Gaze. *Journal of Communication*, *52*(3), 566–580. https://doi.org/10.1111/j.1460-2466.2002.tb02562.x

Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., & Castiello, U. (2012). Social grasping: From mirroring to mentalizing. *NeuroImage*, *61*(1), 240–248. https://doi.org/10.1016/j.neuroimage.2012.03.013

Beja-Pereira, A., Luikart, G., England, P. R., Bradley, D. G., Jann, O. C., Bertorelle, G.et al. (2003). Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nature Genetics*, *35*(4), 311–313. https://doi.org/10.1038/ng1263

Benzon, W. (2002). *Beethoven's Anvil: Music in Mind and Culture.* Oxford: Oxford University Press.

Bernhardt, B. C., & Singer, T. (2012). The Neural Basis of Empathy. *Annual Review of Neuroscience*, *35*(1), 1–23. https://doi.org/10.1146/annurev-neuro-062111-150536

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, *12*(2), 120–138. https://doi.org/10.1007/BF00986930

Berridge, K. C. (1996). Food Reward: Brain Substrates of Wanting and Liking. *Neuroscience and Biobehavioral Reviews*, *20*(1), 1–25. https://doi.org/10.1016/0149-7634(95)00033-B

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bishop, L. (2018). Collaborative musical creativity: How ensembles coordinate spontaneity. *Frontiers in psychology*, *9*, 1285. https://doi.org/10.3389/fpsyg.2018.01285

Bispham, J. C. (2009). Music's "design features": Musical motivation, musical pulse, and musical pitch. *Musicae Scientiae*, *13*(2_suppl), 41–61. https://doi.org/10.1177/1029864909013002041

Blacking, J. (1967). *Venda children's songs: a study in ethnomusicological analysis*. Johannesburg: Witwatersrand University Press.

Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.

Blocker, H. S., & McIntosh, D. N. (2017). Not All Outgroups Are Equal: Group Type May Influence Group Effect on Matching Behavior. *Journal of Nonverbal Behavior*, *4*, 395–413. https://doi.org/10.1007/s10919-017-0258-z

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*(1), 111–137. https://doi.org/10.1037/0033-2909.119.1.111

Botvinick, M. & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*, 756. doi:10.1038/35784

Bourhis, R. Y. & Giles, H. (1977). 'The language of intergroup distinctiveness', in H. Giles (Ed.) *Language, ethnicity, and intergroup relations* (pp. 119-34). London: Academic Press.

Bowlby, J. (1984). *Attachment and loss, Vol. 1. Attachment.* Harmondsworth: Penguin (2nd edition).

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. London: University of Chicago Press.

Boyd, R. T., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134

Boyd, R., & Richerson, P. J. (2010). Transmission coupling mechanisms: cultural group selection. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, *365*(1559), 3787–3795. https://doi.org/10.1098/rstb.2010.0046

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, *3*(327). https://doi.org/10.3389/fpsyg.2012.00327

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia*, *43*(1), 89–98. https://doi.org/10.1016/j.neuropsychologia.2004.06.018

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating Action Understanding: Inferential Processes versus Action Simulation. *Current Biology*, *17*(24), 2117–2121. https://doi.org/10.1016/j.cub.2007.11.057

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation : A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307–324.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482. https://doi.org/10.1177/0146167291175001

Brewer, M. B., & Gardner, W. (1996). Who Is this "We"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, *71*(1), 83–93. https://doi.org/10.1037/0022-3514.71.1.83

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brown, D. E. (1991). *Human universals*. London: McGraw Hill.

Brown, J. L. (1964). The Evolution of Diversity in Avian Territorial Systems. *The Wilson Bulletin*, *76*(2), 160-169.

Brown, P., Levinson, S., & Gumperz, J. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511813085

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind - evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience and Biobehavioral Reviews*, *30*(4), 437–455. https://doi.org/10.1016/j.neubiorev.2005.08.001

Brunner, H. G., Nelen, M., O'Breakefield, X., Ropers, H. H., & van Oost, B. A. (2009). Abnormal Behavior Associated with a Point Mutation in the Structural Gene for Monoamine Oxidase A. *Science*, *262*(5133), 578–580. https://doi.org/10.1126/science.8211186

Bryson, B. (1996). "Anything But Heavy Metal": Symbolic Exclusion and Musical Dislikes. *American Sociological Review*, *61*(5), 884–899. DOI:10.2307/2096459

Burgoon, J. K., Buller, D. B., Hale, J. L., & de Turck, M. A. (1984). Relational Messages Associated With Nonverbal Behaviors. *Human Communication Research*, *10*(3), 351–378. https://doi.org/10.1111/j.1468-2958.1984.tb00023.x

Burgoon, J. K., & Hale, J. L. (1984). The fundamental topoi of relational communication. *Communication Monographs*, *51*(3), 193–214. https://doi.org/10.1080/03637758409390195

Burrows, D. (1987). Instrumentalities. *The Journal of Musicology*, *5*(1), 117–125. DOI:10.2307/763827

Canonne, C., & Aucouturier, J. J. (2016). Play together, think alike: Shared mental models in expert music improvisers. *Psychology of Music*, *44*(3), 544-558. https://doi.org/10.1177/0305735615577406

Caporael, L. R., & Brewer, M. B. (1991). Reviving Evolutionary Psychology: Biology Meets Society. *Journal of Social Issues*, *47*(3), 187–195.

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role Reversal Imitation and Language in Typically Developing Infants and Children With Autism, *Infancy*, *8*(3), 253–278. https://doi.org/10.1207/s15327078in0803_4

Carter, C. S. (2005). The chemistry of child neglect: do oxytocin and vasopressin mediate the effects of early experience? *Proceedings of the National Academy of Sciences*, *102*(51), 18247–18248. https://doi.org/10.1073/pnas.0509376102

Cartwright, D. (1968). 'The Nature of Group Cohesiveness', in A. F. Zander & D. Cartwright (Eds.) *Group dynamics: research and theory* (pp. 91-109), London: Tavistock (3rd edition).

Castelfranchi, C. & Poggi, I. (1990). 'Blushing as a discourse: Was Darwin wrong?', in W. R. Crozier (Ed.) *Shyness and Embarrassment: Perspectives from Social Psychology* (pp. 230-252). Cambridge: Cambridge University Pres (Online publication, 2010) https://doi.org/10.1017/CBO9780511571183.009

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal Hierarchies in the Music of North India. *Journal of Experimental Psychology General*, *113*(3), 394-412.

Champagne, F. A. (2008). Epigenetic mechanisms and the transgenerational effects of maternal care. *Frontiers in Neuroendocrinology*, *29*(3), 386–397. https://doi.org/10.1016/j.yfrne.2008.03.003

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Cheng, Y., Lin, C. P., Liu, H. L., Hsu, Y. Y., Lim, K. E., Hung, D., & Decety, J. (2007). Expertise Modulates the Perception of Pain in Others. *Current Biology*, *17*(19), 1708–1713. https://doi.org/10.1016/j.cub.2007.09.020

Chomsky, N. (1959). Review: *Verbal behaviour* by B. F. Skinner. *Language, 35*(1), 26-58. doi:10.2307/411334

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-Based Helping: Is It Selflessly or Selfishly Motivated? *Journal of Personality and Social Psychology*, *52*(4), 749–758. https://doi.org/10.1037/0022-3514.52.4.749

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014a). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2014b). Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1658). https://doi.org/10.1098/rstb.2013.0400

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016). Social Effects of Movement Synchrony: Increased Infant Helpfulness only Transfers to Affiliates of Synchronously Moving Partners. *Infancy*, *21*(6), 807–821. https://doi.org/10.1111/infa.12140

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press (ACLS Humanities E-Book).

Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of Life Reviews*, *15*, 61–88. https://doi.org/10.1016/j.plrev.2015.09.001

Clayton, M. (2012). What is entrainment? Definition and applications in musical research. *Empirical Musicology Review*, *7*(1–2), 49–56. http://dro.dur.ac.uk/10102/

Clayton, M., Sager, R., & Will, U. (2005). In Time with the Music: The Concept of Entrainment and its Significance for Ethnomusicology. *European Meetings in Ethnomusicology*, *11*(1), 1-82. https://doi.org/10.1063/1.2756072

Clayton, M., Seeger, R., & Will, U. (2005). In Time with the Music: The Concept of Entrainment and its Significance for Ethnomusicology. *European Meetings in Ethnomusicology*, *11* (ESEM Counterpoint 1), 1-82. https://doi.org/10.1063/1.2756072

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377–395. https://doi.org/10.1016/0885-2014(94)90012-4

Coan, J. A. (2008). 'Toward a neuroscience of attachment', in J. Cassidy & P. R. Shaver (Eds.) *Handbook of Attachment: Theory, Research, and Clinical Implications* (2nd edition, pp. 241-265), New York: The Guildford Press.

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a Hand: Social Regulation of the Neural Response to Threat. *Psychological Science*, *17*(12), 1032–1039. https://doi.org/10.1111/j.1467-9280.2006.01832.x

Cohen, E. E. A., Ejsmond-Frey, R., Knight, N., & Dunbar, R. I. M. (2010). Rowers' high: behavioural synchrony is correlated with elevated pain thresholds. *Biology Letters*, *6*(1), 106–108. https://doi.org/10.1098/rsbl.2009.0670

Cohen, E., & Haun, D. (2013). The development of tag-based cooperation via a socially acquired trait. *Evolution and Human Behavior*, *34*(3), 230–235. https://doi.org/10.1016/j.evolhumbehav.2013.02.001

Collard, M., Vaesen, K., Cosgrove, R., & Roebroeks, W. (2016). The empirical case against the 'demographic turn' in Palaeolithic archaeology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1698). https://doi.org/10.1098/rstb.2015.0242

Collinson, I. (2009). 'Singing Songs, Making Places, Creating Selves:' Football Songs & Fan Identity at Sydney FC. *Transforming Cultures eJournal*, *4*(1). https://doi.org/10.5130/tfc.v4i1.1057

Conard, N., Malina, M. & Münzel, S. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature 460***,** 737–740. https://doi.org/10.1038/nature08169

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research*, *5*(3), 221–235. https://doi.org/10.1016/0022-3956(67)90004-0

Cook, N. (2017). 'Scripting Social Interaction: Improvisation, Performance and Western "Art" Music', in G. Born, E. Lewis, & W. Straw (Eds.) *Improvisation and Social Aesthetics* (pp. 59-77), Durham: Duke University Press.

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: From origin to function. *Behavioral and Brain Sciences*, *37*(2), 177–192. https://doi.org/10.1017/S0140525X13000903

Cosmides L. & Tooby J. (1992) 'Cognitive Adaptations for Social Exchange', in J. H. Barkow, L. Cosmides, & J. Tooby (Eds.) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 165-228), Oxford: Oxford University Press.

Coupland, J., Coupland, N., & Robinson, J. D. (1992). "How are you?": Negotiating phatic communion. *Language in Society*, *21*(2), 207–230. https://doi.org/10.1017/S0047404500015268

Crocker, J., & Luhtanen, R. (1990). Collective Self-Esteem and Ingroup Bias. *Journal of Personality and Social Psychology*, *58*(1), 60–67. https://doi.org/10.1037/0022-3514.58.1.60

Cross, I. (1999). 'Is music the most important thing we ever did? Music, development and evolution' in Suk Won Yi (Ed.), *Music, mind and science* (pp. 10-39), Seoul: Seoul National University Press.

Cross, I. (2003). Music and Evolution: Consequences and Causes. *Contemporary Music Review*, *22*(3), 79–89. https://doi.org/10.1080/0749446032000150906

Cross, I. (2005a). Music and Social Being. *Musicology Australia*, *28*(1), 114-126. DOI: 10.1080/08145857.2005.10415281

Cross, I. (2005b). 'Music and meaning, ambiguity and evolution', in D. Miell, R. Macdonald & D. Hargreaves (Eds.) *Musical Communication* (pp. 27-43), Oxford: Oxford University Press.

Cross, I. (2012a). Cognitive Science and the Cultural Nature of Music. *Topics in Cognitive Science*, *4*(4), 668–677. https://doi.org/10.1111/j.1756-8765.2012.01216.x

Cross, I. (2012b). 'Music and Biocultural Evolution'. In M. Clayton, T. Herbert & R. Middleton (Eds.) *The Cultural Study of Music: A Critical Introduction* (2nd edition, pp. 17-27). New York: Routledge.

Cross, I. (2013). "Does not compute"? Music as real-time communicative interaction. *AI and Society*, *28*(4), 415–430. https://doi.org/10.1007/s00146-013-0511-x

Cross, I. (forthcoming). Music in the digital age: commodity, community, communion.

Cross, I. & Tolbert, E. (forthcoming). 'Musical Epistemologies', in T. McAuley, J. Levinson., & N. Nielsen. (Eds.) *The Oxford Handbook of Western Music and Philosophy*.

Cross, I. & Woodruff, G. E. (2009). 'Music as a communicative medium', in R. Botha and C. Knight (Eds.) *The prehistory of language*, (pp. 113-144), Oxford, Oxford University Press.

Cross, L., Turgeon, M., & Atherton, G. (2019). Moving with the in-crowd: Cooperation and interpersonal entrainment in in- vs. out- groups. *Current Psychology*, *19*. https://doi.org/10.1007/s12144-019-00283-0

Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019). Reciprocal information flow and role distribution support joint action coordination. *Cognition*, *187*, 21–31. https://doi.org/10.1016/j.cognition.2019.02.006

d'Ausilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, *19*(3), 111–114. https://doi.org/10.1016/j.tics.2015.01.005

d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A., Soressi, M., et al. (2003). Archaeological Evidence for the Emergence of Language, Symbolism, and Music - An Alternative Multidisciplinary Perspective. *Journal of World Prehistory*, *17*(1), 1–70. https://doi.org/10.1023/A:1023980201043

Damasio, A. R. (2006). *Descartes' Error: Emotion, Reason and the Human Brain*. (London: Vintage Books).

Darwin, C. (1859, 1959 edition). *The origin of species; a variorum text*. Ed. M. Peckham. Philadelphia: University of Pennsylvania Press.

Darwin, C. (1871). *The descent of man: and selections in relation to sex*. Accessed online (eBooks@Adelaide).

Davis, A., Taylor, J., & Cohen, E. (2015). Social Bonds and Exercise: Evidence for a Reciprocal Relationship. *PLoS ONE*, *10*(8), e0136705. https://doi.org/10.1371/journal.pone.0136705

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

de Dreu, C. K. W., & Kret, M. E. (2016). Oxytocin Conditions Intergroup Relations Through Upregulated In-Group Empathy, Cooperation, Conformity, and Defense. *Biological Psychiatry*, *79*(3), 165–173. https://doi.org/10.1016/j.biopsych.2015.03.020

de Waal, F. B. M. (2008). Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology*, *59*(1), 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Deacon, T. (1997). *The symbolic species: the co-evolution of language and the human brain*. (London: Allen Lane).

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Devereux, P. G., & Ginsburg, G. P. (2001). Sociality Effects on the Production of Laughter. *The Journal of General Psychology*, *128*(2), 227–240. https://doi.org/10.1080/00221300109598910

Di Pellegrino, G., Fadiga, L., Fogassi, L. Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180. https://doi.org/10.1007/BF00230027

Dillard, J. P., Solomon, D. H., & Palmer, M. T. (1999). Structuring the concept of relational communication. *Communication Monographs*, *66*(1), 49–65. https://doi.org/10.1080/03637759909376462

Dillard, J. P., Solomon, D. H., & Samp, J. A. (1996). Framing Social Reality: The Relevance of Relational Judgements. *Communication Research*, *23*(6), 703-723. https://doi.org/10.1177/009365096023006004

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., et al. (2015). Universal Principles in the Repair of Communication Problems. *PLoS ONE*, *10*(9), e0136100. https://doi.org/10.1371/journal.pone.0136100

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory and Cognition*, *33*, 727–733. https://doi.org/10.3758/BF03195338

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469–493. https://doi.org/10.1016/0047-2484(92)90081-J

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(4), 681–735. https://doi.org/10.1017/S0140525X00032325

Dunbar, R. I. M. (1996). *Grooming, gossip and the evolution of language*. London: Faber.

Dunbar, R. (2004). 'Language, Music and Laughter in Evolutionary Perspective', in D. Kimbrough & U. Griebel (Eds.) *Evolution of Communication Systems: A Comparative Approach* (pp. 257-273). London: MIT Press.

Dunbar, R.I.M. (2007) The Social Brain Hypothesis and its Relevance to Social Psychology, in J. P. Forgas, M. G. Haselton, & W. von Hippel (Eds). *Evolution and the Social Mind*, pp. 21-31. Hove: Psychology Press.

Dunbar, R. I. M. (2010). The social role of touch in humans and primates: Behavioural function and neurobiological mechanisms. *Neuroscience & Biobehavioral Reviews*, *34*(2), 260–268. https://doi.org/10.1016/j.neubiorev.2008.07.001

Dunbar, R. I. M. (2018). The Anatomy of Friendship. *Trends in Cognitive Sciences*, *22*(1), 32–51. https://doi.org/10.1016/j.tics.2017.10.004

Dunbar, R. I. M., Baron, R., Frangou, A., Pearce, E., van Leeuwen, E. J. C., Stow, J., et al. (2011). Social laughter is correlated with an elevated pain threshold. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 279*(1731). https://doi.org/10.1098/rspb.2011.1373

Dunbar, R. I. M., Duncan, N. D. C., & Nettle, D. (1995). Size and structure of freely forming conversational groups. *Human Nature*, *6*(1), 67–78. https://doi.org/10.1007/BF02734136

Dunbar, R. I. M., Kaskatis, K., MacDonald, I., & Barra, V. (2012). Performance of Music Elevates Pain Threshold and Positive Affect: Implications for the Evolutionary Function of Music. *Evolutionary Psychology*, *10*(4), 688–702.

Dunbar, R. I. M., Korstjens, A. H., & Lehmann, J. (2009). Time as an ecological constraint. *Biological Reviews*, *84*(3), 413–429. https://doi.org/10.1111/j.1469-185X.2009.00080.x

Dunbar, R. I. M. & Gowlett, J. A. J. (2014). 'Fireside Chat: The Impact of Fire on Hominin Socioecology', in R. I. M. Dunbar, C. Gamble, & J. A. J. Gowlett (Eds.) *Lucy to language: the benchmark papers* (pp. 277-296), Oxford: Oxford University Press.

Dunbar, R. I. M., & Lehmann, J. (2013). Grooming and social cohesion in primates: a comment on Grueter et al. *Evolution and Human Behavior*, *34*(6), 453–455. https://doi.org/10.1016/j.evolhumbehav.2013.08.003

Duncan Jr, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283–292. http://dx.doi.org/10.1037/h0033031

Durkheim, É. (2005). The Dualism of Human Nature and Its Social Conditions. *Durkheimian Studies, 11*(1), 35-45. https://doi.org/10.3167/175223005783472211

Durrheim, K., Quayle, M., Tredoux, C. G., Titlestad, K., & Tooke, L. (2016). Investigating the Evolution of Ingroup Favoritism Using a Minimal Group Interaction Paradigm: The Effects of Inter- and Intragroup Interdependence. *PLoS ONE*, *11*(11), 1–26, e0165974.. https://doi.org/10.1371/journal.pone.0165974

Edelmann, R. J. (1987) *The psychology of embarrassment*. Chichester: Wiley.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, *115*(2), 268– 287.

Ellsworth, P. C.. & Carlsmith. J. M. (1968). Effects of eye contact and verbal content on affective response to a dyadic interaction. *Journal of Personality and Social Psychology*, *10*(1), 15-20. https://doi.org/10.1037/h0026385

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Everett, D. (2012). *Language: The Cultural Tool*. London: Profile Books.

Exline, R. V. (1963). Explorations in the process of person perception: visual interaction in relation to competition, sex, and need for affiliation. *Journal of Personality*, *31*(1), 1-20. https://doi.org/10.1111/j.1467-6494.1963.tb01836.x

Fairhurst, M. T., Janata, P., & Keller, P. E. (preprint). Distinguishing "self" from "other" in a dynamic synchronization task with an adaptive virtual partner. *bioRxiv,* 625061. https://doi.org/10.1101/625061

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, *27*(4), 491–503. https://doi.org/10.1017/S0140525X04000111

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Fawcett, C., & Tunçgenç, B. (2017). Infants' use of movement synchrony to infer social affiliation in others. *Journal of Experimental Child Psychology*, *160*, 127–136. https://doi.org/10.1016/j.jecp.2017.03.014

Fay, N., Garrod, S., & Carletta, J. (2000). Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science*, *11*(6), 481–486. https://doi.org/10.1111/1467-9280.00292

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*(1), 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–791. https://doi.org/10.1038/nature02043

Feld, S. (1982). *Sound and sentiment: birds, weeping, poetics, and song in Kaluli expression*. Philadelphia: University of Pennsylvania Press.

Feld, S. (1990). Wept Thoughts: The Voicing of Kaluli Memories. *Oral Tradition*, *5*(2–3), 241–266.

Feldman, R. (2007). Mother-Infant Synchrony and the Development of Moral Orientation in Childhood and Adolescence: Direct and Indirect Mechanisms of Developmental Continuity. *American Journal of Orthopsychiatry*, *77*(4), 582-597.

Feldman, R. (2012). Oxytocin and social affiliation in humans. *Hormones and Behavior*, *61*(3), 380–391. https://doi.org/10.1016/j.yhbeh.2012.01.008

Feldman, R. (2017). The Neurobiology of Human Attachments. *Trends in Cognitive Sciences*, *21*(2), 80–99. https://doi.org/10.1016/j.tics.2016.11.007

Fessler, D. M. T., & Holbrook, C. (2014). Marching into battle: synchronized walking diminishes the conceptualized formidability of an antagonist in men. *Biology Letters*, *10*(8). https://doi.org/10.1098/rsbl.2014.0592

Fino, E., Menegatti, M., Avenanti, A., & Rubini, M. (2019). Unfolding political attitudes through the face: facial expressions when reading emotion language of left- and right-wing political leaders. *Scientific Reports*, *9*(15689). https://doi.org/10.1038/s41598-019-51858-7

Fischer, R., Callander, R., Reddish, P., & Bulbulia, J. (2013). How Do Rituals Affect Cooperation? An Experimental Field Study Comparing Nine Ritual Types. *Human Nature*, *24*(2), 115–125. https://doi.org/10.1007/s12110-013-9167-y

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*(1), 173–215. https://doi.org/10.1016/j.cognition.2005.11.009

Fitch, T. (2012). 'The biology and evolution of rhythm: unravelling a paradox', in P. Rebuschat, M. Rohrmeier, J. A. Hawkins, & I. Cross (Eds.) *Language and Music as Cognitive Systems* (pp. 73-95), Oxford: Oxford University Press.

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0091

Foley, R. A., Martin, L., Mirazo, M., Stringer, C., & Foley, R. A. (2016). Major transitions in human evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1698). https://doi.org/10.1098/rstb.2015.0229

Franco, F. (2005). 'Infant Pointing: Harlequin, Servant of Two Masters', in N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.) *Joint Attention: Communication and Other Minds: Issues in Philosophy and Psychology*, Oxford: Oxford University Press. (Oxford Scholarship Online) https://doi.org/10.1093/acprof:oso/9780199245635.003.0007

Freeman, W. J. (2000). 'A neurobiological role of music in social bonding', in N. L. Wallin, B. Merker, & S Brown (Eds.) *The Origins of Music* (pp. 411-424), Cambridge: MIT Press.

Fritz, T. (2013). The Dock-in Model of Music Culture and Cross-cultural Perception. *Music Perception: An Interdisciplinary Journal*, *30*(5), 511–516. DOI:10.1525/mp.2013.30.5.511

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), R724-R732). https://doi.org/10.1016/j.cub.2007.05.068

Fuchs, T. (2019). Empathy, Group Identity, and the Mechanisms of Exclusion: An Investigation into the Limits of Empathy. *Topoi*, *38*, 239–250. https://doi.org/10.1007/s11245-017-9499-z

Fuentes, A. (2006). Review of *The Origin and Evolution of Cultures* and *Not by Genes Alone*. *American Anthropologist*, *108*(3), 547–549.

Gabrielsson, A. (2001). 'Emotions in strong experiences with music', in P. N. Juslin & J. A. Sloboda (Eds.), *Series in affective science. Music and emotion: Theory and research* (pp. 431-449), New York, NY, US: Oxford University Press.

Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the Closeness of Relationships: A Comprehensive Evaluation of the "Inclusion of the Other in the Self" Scale. *PLoS ONE*, *10*(6), e0129478. https://doi.org/10.1371/journal.pone.0129478

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The Common Ingroup Identity Model: Recategorization and the Reduction of Intergroup Bias. *European Review of Social Psychology*, *4*(1), 1–26. https://doi.org/10.1080/14792779343000004

Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology*, *59*(4), 692–704. https://doi.org/10.1037/0022-3514.59.4.692

Gaigg, S. B. (2012). The interplay between emotion and cognition in autism spectrum disorder: implications for developmental theory. *Frontiers in Integrative Neuroscience*, *6*(113). https://doi.org/10.3389/fnint.2012.00113

Galbusera, L., Finn, M. T. M., Tschacher, W., & Kyselo, M. (2019). Interpersonal synchrony feels good but impedes self-regulation of affect. *Scientific Reports*, *9*(14691). https://doi.org/10.1038/s41598-019-50960-0

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social Neuroscience*, *3*(3–4), 317–333. https://doi.org/10.1080/17470910701563608

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403. https://doi.org/10.1016/j.tics.2004.07.002

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

Gamble, C. (1998). Palaeolithic society and the release from proximity: A network approach to intimate relations. *World Archaeology*, *29*(3), 426–449. https://doi.org/10.1080/00438243.1998.9980389

Gardner, A., & Grafen, A. (2009). Capturing the superorganism: A formal theory of group adaptation. *Journal of Evolutionary Biology*, *22*(4), 659–671. https://doi.org/10.1111/j.1420-9101.2008.01681.x

Gardner, W. L., Pickett, C. L., & Brewer, M. B. (2000). Social Exclusion and Selective Memory: How the Need to Belong Influences Memory for Social Events. *Personality and Social Psychology Bulletin*, *26*(4), 486–496. https://doi.org/10.1177/0146167200266007

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Garrod, S., & Pickering, M. J. (2009). Joint Action, Interactive Alignment, and Dialog. *Topics in Cognitive Science*, *1*(2), 292–304. https://doi.org/10.1111/j.1756-8765.2009.01020.x

Gebauer, L., Witek, M. A. G., Hansen, N. C., Thomas, J., Konvalinka, I., & Vuust, P. (2016). Oxytocin improves synchronisation in leader-follower interaction. *Scientific Reports*, *6*(38416). https://doi.org/10.1038/srep38416

George, N., & Conty, L. (2008). Facing the gaze of others. *Clinical Neurophysiology*, *38*(3), 197–207. https://doi.org/10.1016/j.neucli.2008.03.001

Gervais, M., & Wilson, D. S. (2005). The Evolution and Functions of Laughter and Humor: A Synthetic Approach. *The Quarterly Review of Biology*, *80*(4), 395–430.

Giles, H., Denes, A., Hamilton, D. L., & Hajda, J. M. (2009). Striking a Chord: A Prelude to Music and Intergroup Relations Research. *Group Processes & Intergroup Relations*, *12*(3), 291–301. https://doi.org/10.1177/1368430209102840

Giles, H. & Johnson, P. (1981). 'The Role of Language in Ethnic Group Relations', in J. C. Turner & H. Giles (Eds.) *Intergroup behaviour* (pp. 199-243), Oxford: Basil Blackwell.

Gintis, H. (2011). Gene-culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1566), 878–888. https://doi.org/10.1098/rstb.2010.0310

Goehr, L. (1992). *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press (Oxford Scholarship Online, 2003, DOI: 10.1093/0198235410.001.0001).

Goffman, E. (1956). Embarrassment and Social Organization. *American Journal of Sociology*, *62*(3), 264–271. https://doi.org/10.1086/222003

Goffman, E. (1959, 1990 reprint). *The Presentation of Self in Everyday Life*. London: Penguin.

Goffman, E. (1966). *Behavior in public places: notes on the social organization of gatherings*. New York: Free Press.

Goffman, E. (1967). *Interaction Ritual: Essays on face-to-face behaviour*. New York: Anchor Books.

Goffman, E. (1972). *Encounters: Two Studies in the Sociology of Interaction*. London: Allen Lane.

Golding, W. (1988). *Lord of the flies*. J. R. Baker and A. P. Ziegler (Eds). New York: Berkley Pub. Group.

Gonzalez-Liencres, C., Shamay-Tsoory, S. G., & Brüne, M. (2013). Towards a neuroscience of empathy: Ontogeny, phylogeny, brain mechanisms, context and psychopathology. *Neuroscience and Biobehavioral Reviews*, *37*(8), 1537–1548. https://doi.org/10.1016/j.neubiorev.2013.05.001

Good, A., Choma, B., & Russo, F. A. (2017). Movement Synchrony Influences Intergroup Relations in a Minimal Groups Paradigm. *Basic and Applied Social Psychology*, *39*(4), 231–238. https://doi.org/10.1080/01973533.2017.1337015

Good, A., & Russo, F. A. (2016). Singing Promotes Cooperation in a Diverse Group of Children. *Social Psychology*, *47*(6), 340–344. https://doi.org/10.1027/1864-9335/a000282

Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, *50*(3–4), 272–302. https://doi.org/10.1111/j.1475-682X.1980.tb00023.x

Graham, G. (2019). 'Behaviorism', in E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. California: Stanford University Press (Spring 2019 edition, https://plato.stanford.edu/archives/spr2019/entries/behaviorism/, accessed 1st July 19).

Gratier, M. (2008). Grounding in musical interaction: Evidence from jazz performances. *Musicae Scientiae*, *12*(1_suppl), 71–110. https://doi.org/10.1177/1029864908012001041

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Grice, H. P. (1969). Utterer's Meaning and Intentions. *The Philosophical Review*, *78*(2), 147–177. DOI:10.2307/2184179

Groppe, S. E., Gossen, A., Rademacher, L., Hahn, A., Westphal, L., Gründer, G., & Spreckelmeyer, K. N. (2013). Oxytocin Influences Processing of Socially Relevant Cues in the Ventral Tegmental Area of the Human Brain. *Biological Psychiatry*, *74*(3), 172–179. https://doi.org/10.1016/j.biopsych.2012.12.023

Grueter, C. C., Bissonnette, A., Isler, K., & van Schaik, C. P. (2013). Grooming and group cohesion in primates: implications for the evolution of language. *Evolution and Human Behavior*, *34*(1), 61–68. https://doi.org/10.1016/j.evolhumbehav.2012.09.004

Gutsell, J. N., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology*, *46*(5), 841–845. https://doi.org/10.1016/j.jesp.2010.03.011

Gutsell, J. N., & Inzlicht, M. (2012). Intergroup differences in the sharing of emotive states: neural evidence of an empathy gap. *Social Cognitive and Affective Neuroscience*, *7*(5), 596–603. https://doi.org/10.1093/scan/nsr035

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51. https://doi.org/10.1007/s12110-003-1015-z

Hagen, E. H. & Hammerstein, P. (2009). Did Neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves. *Musicae Scientiae*, *13*(2), 291-320. https://doi.org/10.1177/1029864909013002131

Hannon, E. E., & Trehub, S. E. (2005). Metrical Categories in Infancy and Adulthood. *Psychological Science*, *16*(1), 48–55. https://doi.org/10.1111/j.0956-7976.2005.00779.x

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*, 243-267. https://doi.org/10.1146/annurev-psych-010416-044046

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. *Neuron*, *88*(1), 181-193. https://doi.org/10.1016/j.neuron.2015.09.022

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups. *Psychological Science*, *17*(10), 847–853.

Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology*, *25*(2), 179–193. https://doi.org/10.1002/ejsp.2420250205

Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social Identity Salience and the Emergence of Stereotype Consensus. *Personality and Social Psychology Bulletin*, *25*(7), 809–818. https://doi.org/10.1177/0146167299025007004

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1693). https://doi.org/10.1098/rstb.2015.0366

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: a comparative perspective. *Nature Neuroscience*, *6*(7), 663–668. https://doi.org/10.1038/nn1080

Healey, P. G. T., Leach, J. & Bryan-Kinns, N. (2005). 'Inter-play: Understanding group music improvisation as a form of everyday interaction', in *Proceedings of Less is More – Simple Computing in an Age of Complexity*. Cambridge: Microsoft Research

Heine, S. J., & Norenzayan, A. (2006). Toward a Psychological Science for a Cultural Species. *Perspectives on Psychological Science*, *1*(3), 251–269. https://doi.org/10.1111/j.1745-6916.2006.00015.x

Henrich, J., & Gil-White, F. J. (1980). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196. https://doi.org/10.1016/S1090-5138(00)00071-4

Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five Misunderstandings About Cultural Evolution. *Human Nature*, *19*(2), 119–137. https://doi.org/10.1007/s12110-008-9037-1

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioural and Brain Sciences*, *33*(2-3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hess, U., & Fischer, A. (2013). Emotional Mimicry as Social Regulation. *Personality and Social Psychology Review*, *17*(2), 142–157. https://doi.org/10.1177/1088868312472607

Hetherington, C., Hendrickson, C., & Koenig, M. (2014). Reducing an in-group bias in preschool children: the impact of moral behavior. *Developmental Science*, *17*(6), 1042–1049. https://doi.org/10.1111/desc.12192

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology*, *53*(1), 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science*, *17*(5), 647–659. https://doi.org/10.1111/desc.12148

Hodgson, D. (2017). Costly signalling, the arts, archaeology and human behaviour. *World Archaeology*, *49*(4), 446-465. https://doi.org/10.1080/00438243.2017.1281757

Hodgson, D., & Verpooten, J. (2014). The Evolutionary Significance of the Arts: Exploring the By-product Hypothesis in the Context of Ritual, Precursors, and Cultural Evolution. *Biological Theory*, *10*(1), 73–85. https://doi.org/10.1007/s13752-014-0182-y

Hogg, M. A., & Abrams, D. (1988). *Social identifications: a social psychology of intergroup relations and group processes*. London: Routledge.

Hogg, M. A. & Abrams, D. (1990). 'Social motivation, self-esteem and social identity', in D. Abrams & M. A. Hogg (Eds.) *Social identity theory: constructive and critical advances* (pp. 28-47), London: Harvester Wheatsheaf.

Hogg, M. A., & Reid, S. A. (2006). Social Identity, Self-Categorization, and the Communication of Group Norms. *Communication Theory*, *16*(1), 7–30. https://doi.org/10.1111/j.1468-2885.2006.00003.x

Hogg, M. A., & Turner, J. C. (1985). Interpersonal attraction, social identification and psychological group formation. *European Journal of Social Psychology*, *15*(1), 51–66. https://doi.org/10.1002/ejsp.2420150105

Holler, J. & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: optimizing recipiency. *Frontiers in Psychology*, *6*(98), 76-89. https://doi.org/10.3389/fpsyg.2015.00098

Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, *27*(6), 949–961. https://doi.org/10.1521/soco.2009.27.6.949

Huberth, M., Dauer, T., Nanou, C., Román, I., Gang, N., Reid, W., et al. (2019). Performance monitoring of self and other in a turn-taking piano duet: A dual-EEG study. *Social Neuroscience, 14*(4), 449-461. DOI: 10.1080/17470919.2018.1492968

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Razzolatti, G. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System, *PLoS Biology*, *3*(3), e79. https://doi.org/10.1371/journal.pbio.0030079

Ingold, T. (1998). From complementarity to obviation: on dissolving the boundaries between social and biological anthropology, archaeology and psychology. *Zeitschrift für Ethnologie*, *123*(1), 21–52.

Insel, T. R., & Fernald, R. D. (2004). How The Brain Processes Social Information: Searching for the Social Brain. *Annual Review of Neuroscience*, *27*(1) 697–722. https://doi.org/10.1146/annurev.neuro.27.070203.144148

Insel, T. R., & Young, L. J. (2001). The neurobiology of attachment, *Nature Reviews Neuroscience, 2*, 129-136. doi:10.1038/35053579

Inzlicht, M., Gutsell, J. N., & Legault, L. (2012). Mimicry reduces racial prejudice. *Journal of Experimental Social Psychology*, *48*(1), 361–365. https://doi.org/10.1016/j.jesp.2011.06.007

Iriki, A., & Taoka, M. (2012). Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1585), 10–23. https://doi.org/10.1098/rstb.2011.0190

Isabella, R. A., & Belsky, J. (1991). Interactional Synchrony and the Origins of Infant-Mother Attachment: A Replication Study. *Child Development*, *62*(2), 373–384. https://doi.org/10.1111/j.1467-8624.1991.tb01538.x

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews*, *33*(6), 843-863. https://doi.org/10.1016/j.neubiorev.2009.02.004

Jablonka, E., & Lamb, M. J. (2007). Précis of *Evolution in Four Dimensions*. *Behavioral and Brain Sciences*, *30*(4), 353–365. https://doi.org/10.1017/S0140525X07002221

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, *9*(1), 21–25. https://doi.org/10.1016/j.tics.2004.11.003

Jakobson, R. (1960). Closing statement: Linguistics and poetics, in T. A. Sebeok (Ed.), *Style in language* (pp. 350-377). New York & London: John Wiley & Sons

Janis, I. L. (1982). *Groupthink: psychological studies of policy decisions and fiascoes*. Boston: Houghton Mifflin (2nd edition).

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). 'Musical Gestures: Concepts and Methods in Research', in R. I. Godøy & M. Leman (Eds.) *Musical Gestures: Sound, Movement, and Meaning* (pp. 12-35), Abingdon: Routledge.

Joffe, T. H. (1997). Social pressures have selected for an extended juvenile period in primates. *Journal of Human Evolution*, *32*(6), 593–605. https://doi.org/10.1006/jhev.1997.0140

Johnson, J. C. (1994). 'Anthropological contributions to the study of social networks: a review', in S. Wasserman & J. Galaskiewicz (Eds.) *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences* (pp. 113-151), Thousand Oaks: SAGE Publications. https://dx.doi.org/10.4135/9781452243528

Johnson-George, C., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology*, *43*(6), 1306–1317. https://doi.org/10.1037/0022-3514.43.6.1306

Jones, M. R. (2016). 'Musical time', in S. Hallam, I. Cross, & M. Thaut (Eds.) *The Oxford Handbook of Music Psychology* (pp. 125-141), Oxford: Oxford University Press (2nd edition).

Jordan, P., & Shennan, S. (2003). Cultural transmission, language, and basketry traditions amongst the California Indians. *Journal of Anthropological Archaeology*, *22*(1), 42–74. https://doi.org/10.1016/S0278-4165(03)00004-7

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., & Laukka, P. (2003). Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kalkhoff, W., Dippong, J., & Gregory Jr, S. W. (2011). The Biosociology of Solidarity. *Sociology Compass, 5*(10), 936-948. doi:10.1111/j.1751-9020.2011.00412.x

Kampe, K. K. W., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Reward value of attractiveness and gaze. *Nature*, *413*, 589. doi:10.1038/35098149

Katz, M. (2010). *Capturing sound: how technology has changed music*. London: University of California Press (ACLS Humanities E-Book).

Kavanagh, L. C., & Winkielman, P. (2016). The functionality of spontaneous mimicry and its influences on affiliation: An implicit socialization account. *Frontiers in Psychology*, *7*(458). https://doi.org/10.3389/fpsyg.2016.00458

Kendal, J. R. (2011). Cultural Niche Construction and Human Learning Environments: Investigating Sociocultural Perspectives. *Biological Theory*, *6*(3), 241–250. https://doi.org/10.1007/s13752-012-0038-2

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Kendon, A. (1970). Movement coordination in social interaction: Some examples described. *Acta Psychologica*, *32*, 101–125. https://doi.org/10.1016/0001-6918(70)90094-6

Kendon, A. (1990). *Conducting interaction: patterns of behaviour in focused encounters*. Cambridge: Cambridge University Press.

Kercher, K. (1992). Assessing Subjective Well-Being in the Old-Old: The PANAS as a Measure of Orthogonal Dimensions of Positive and Negative Affect. *Research on Aging*, *14*(2), 131-168. https://doi.org/10.1177/0164027592142001

Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology*, *14*(1–2), 155–161. https://doi.org/10.1016/0306-4530(89)90065-6

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41. https://doi.org/10.1016/S0010-0277(03)00064-7

Keysers, C. (2009). Mirror neurons. *Current Biology*, *19*(21), R971-R973. https://doi.org/10.1016/j.cub.2009.08.026

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, *153*(4), 628–636. https://doi.org/10.1007/s00221-003-1603-5

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behaviour in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. https://doi.org/10.1016/j.evolhumbehav.2010.04.004

Kleinke, C. L. (1986). Gaze and Eye Contact: A Research Review. *Psychological Bulletin*, *100*(1), 78–100. https://doi.org/10.1037/0033-2909.100.1.78

Knight, C., & Lewis, J. (2017). Wild voices: Mimicry, Reversal, Metaphor, and the Emergence of Language. *Current Anthropology*, *58*(4), 435–453. https://doi.org/10.1086/692905

Knijnik, J. (2018). Social agency and football fandom: the cultural pedagogies of the Western Sydney ultras. *Sport in Society*, *21*(6), 946–959. https://doi.org/10.1080/17430437.2017.1300394

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

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, *40*(5), 419–435. https://doi.org/10.1006/jhev.2001.0468

Koelsch, S., Gunter, T. C., V. Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach Speaks: A Cortical "Language-Network" Serves the Processing of Music. *NeuroImage*, *17*(2), 956–966. https://doi.org/10.1016/S1053-8119(02)91154-7

Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized Drumming Enhances Activity in the Caudate and Facilitates Prosocial Commitment - If the Rhythm Comes Easily. *PLoS ONE*, *6*(11), e27272. https://doi.org/10.1371/journal.pone.0027272

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*, *63*(11), 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*, 673–676. https://doi.org/10.1038/nature03701

Kret, M. E., & De Dreu, C. K. W. (2017). Pupil-mimicry conditions trust in partners: moderation by oxytocin and group membership. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 284*(1850). https://doi.org/10.1098/rspb.2016.2554

Kreutz, G. (2014). Does Singing Facilitate Social Bonding? *Music and Medicine*, *6*(2), 51–60.

Kytö, M. (2011). "We are the rebellious voice of the terraces, we are Çarşi": Constructing a football supporter group through sound. *Soccer and Society*, *12*(1), 77–93. https://doi.org/10.1080/14660970.2011.530474

Lakens, D., & Stel, M. (2011). If They Move in Sync, They Must Feel in Sync: Movement Synchrony Leads to Attributions of Rapport and Entitativity. *Social Cognition*, *29*(1), 1–14. https://doi.org/http://dx.doi.org/10.1521/soco.2011.29.1.1

Lakin, J. L., & Chartrand, T. L. (2003). Using Nonconscious Behavioral Mimicry to Create Affiliation and Rapport. *Psychological Science*, *14*(4), 334–339. https://doi.org/10.1111/1467-9280.14481

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The Chameleon Effect as Social Glue : Evidence for the Evolutionary Significance of Nonconscious Mimicry. *Journal of Nonverbal Behavior*, *27*(3), 145–162. https://doi.org/10.1023/A:1025389814290

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *Behavioural and Brain Sciences*, *23*(1), 131–175.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2001). Cultural niche construction and human evolution. *Journal of Evolutionary Biology*, *14*(1), 22–33. https://doi.org/10.1046/j.1420-9101.2001.00262.x

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology*, *127*, 191–197. https://doi.org/10.1016/j.biopsycho.2017.06.001

Lansing, J. S., & Fox, K. M. (2011). Niche construction on Bali: the gods of the countryside. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1566), 927–934. https://doi.org/10.1098/rstb.2010.0308

Launay, J., Dean, R. T., & Bailes, F. (2014). Synchronising movements with the sounds of a virtual partner enhances partner likeability. *Cognitive Processing*, *15*(4), 491–501. https://doi.org/10.1007/s10339-014-0618-0

Le Bomin, S., Lecointre, G., & Heyer, E. (2016). The Evolution of Musical Diversity: The Key Role of Vertical Transmission. *PLoS ONE*, *11*(3), e0151570. https://doi.org/10.1371/journal.pone.0151570

Lee, D. H., & Anderson, A. K. (2017). Reading What the Mind Thinks From How the Eye Sees. *Psychological Science*, *28*(4), 494-503. https://doi.org/10.1177/0956797616687364

Lee, R. M., & Robbins, S. B. (1995). Measuring belongingness: The Social Connectedness and the Social Assurance scales. *Journal of Counseling Psychology*, *42*(2), 232–241. https://doi.org/10.1037/0022-0167.42.2.232

Leigh, E. G. (2010). The group selection controversy. *Journal of Evolutionary Biology*, *23*(1), 6–19. https://doi.org/10.1111/j.1420-9101.2009.01876.x

Leman, M. (2008). *Embodied music cognition and mediation technology*. London: MIT Press.

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-Categorization and Bystander Non-Intervention: Two Experimental Studies. *Journal of Applied Social Psychology*, *32*(7), 1452–1463. https://doi.org/10.1111/j.1559-1816.2002.tb01446.x

Levinson, S. C. (2006). 'On the Human "Interaction Engine"', in N. J. Enfield & S. C. Levinson (Eds.) *Roots of human sociality: culture, cognition and interaction* (pp. 39-69), Oxford: Berg.

Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1651). https://doi.org/10.1098/rstb.2013.0302

Lewis, D. K. (1969, 2002 reprint). *Convention: a philosophical study*. Oxford: Blackwell Publishers.

Lewis, D. K. (1975). 'Languages and Language', in K. Gunderson (Ed.) *Minnesota Studies in the Philosophy of Science* (pp. 3-35). Minneapolis: University of Minnesota Press.

Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1599), 2171–2180. https://doi.org/10.1098/rstb.2012.0119

Lewis, J. (2013). A Cross-Cultural Perspective on the Significance of Music and Dance to Culture and Society: Insight from BaYaka Pygmies. In M. A. Arbib (Ed.), *Language, Music, and the Brain: a mysterious relationship* (pp. 45–65). Cambridge: MIT Press.

Lewis, J. (2014). 'BaYaka Pygmy multi-modal and mimetic communication traditions', in D. Dor, C. Knight, & J. Lewis (Eds.), *The Social Origins of Language*, Oxford University Press. DOI:10.1093/acprof:oso/9780199665327.003.0007

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78*(2), 223–246. https://doi.org/10.1037/0022-3514.78.2.223

Likowski, K. U., Muhlberger, A., Seibt, B., Pauli, P., & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, *44*(4), 1065–1072. https://doi.org/10.1016/j.jesp.2007.10.007

Lim, M. M., & Young, L. J. (2006). Neuropeptidergic regulation of affiliative behavior and social bonding in animals, *50*(4), 506–517. https://doi.org/10.1016/j.yhbeh.2006.06.028

Locksley, A., Ortiz, V., & Hepburn, C. (1980). Social categorization and discriminatory behavior: Extinguishing the minimal intergroup discrimination effect. *Journal of Personality and Social Psychology*, *39*(5), 773–783.

London, J. (2002). Cognitive Constraints on Metric Systems: Some Observations and Hypotheses. *Music Perception: An Interdisciplinary Journal*, *19*(4), 529–550. DOI: 10.1525/mp.2002.19.4.529

Lonsdale, A. J., & North, A. C. (2009). Musical Taste and Ingroup Favouritism. *Group Processes & Intergroup Relations*, *12*(3), 319–327. https://doi.org/10.1177/1368430209102842

Lucas, G., Clayton, M., & Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musical Review*, *6*(2), 75–102. doi:10.18061/1811/51203

MacDonald, K., & MacDonald, T. M. (2010). The Peptide That Binds: A Systematic Review of Oxytocin and its Prosocial Effects in Humans. *Harvard Review of Psychiatry*, *18*(1), 1-21. https://doi.org/10.3109/10673220903523615

MacDonald, R., Wilson, G., & Miell, D. (2012). Improvisation as a creative process within contemporary music. *Musical Imaginations: Multidisciplinary Perspectives on Creativity, Performance and Perception* (pp. 242-255). Oxford: Oxford University Press.

Machin, A. J., & Dunbar, R. I. M. (2011). The brain opioid theory of social attachment : a review of the evidence. *Behaviour*, *148*(9-10), 985–1025. https://doi.org/10.1163/000579511X596624

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, *27*(3), 405-416. https://doi.org/10.1016/S0191-8869(98)00251-7

Macrae, C. N., Duffy, O. K., Miles, L. K., & Lawrence, J. (2008). A case of hand waving: Action synchrony and person perception. *Cognition*, *109*(1), 152–156. https://doi.org/10.1016/j.cognition.2008.07.007

Maher, P. J., Van Tilburg, W. A. P., & Van Den Tol, A. J. M. (2013). Meaning in music: Deviations from expectations in music prompt outgroup derogation. *European Journal of Social Psychology*, *43*(6), 449-454. https://doi.org/10.1002/ejsp.1969

Malinowski, B. (1923). 'The problem of meaning in primitive languages', A supplement to C. K. Ogden & I. A. Richards, *The Meaning of Meaning: a study of the influence of language upon thought and of the science of symbolism* (pp. 451-510). London: Kegan Paul.

Manson, J. H., Wrangham, R. W., Boone, J. L., Chapais, B., Dunbar, R. I. M., Ember, C. R., et al. (1991). Intergroup Aggression in Chimpanzees and Humans [and Comments and Replies]. *Current Anthropology*, *32*(4), 369–390. https://doi.org/10.1086/203974

Marean, C. W. (2015). An Evolutionary Anthropological Perspective on Modern Human Origins. *Annual Review of Anthropology, 44*(1), 533-556. https://doi.org/10.1146/annurev-anthro-102313-025954

Marean, C. W. (2016). The transition to foraging for dense and predictable resources and its impact on the evolution of modern humans. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, *371*(1698). https://doi.org/10.1098/rstb.2015.0239

Mariath, L. M., Silva, A. M., Kowalski, T. W., Gattino, G. S., Araujo, G. A., Figueiredo, F. G., et al. (2017). Music genetics research: Association with musicality of a polymorphism in the AVPR1A gene. *Genetics and molecular biology*, *40*(2), 421–429. doi:10.1590/1678-4685-GMB-2016-0021

Maroun, M., & Wagner, S. (2016). Oxytocin and Memory of Emotional Stimuli: Some Dance to Remember, Some Dance to Forget. *Biological Psychiatry*, *79*(3), 203–212. https://doi.org/10.1016/j.biopsych.2015.07.016

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835. https://doi.org/10.1073/pnas.211415698

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*, 547–550. https://doi.org/10.1038/nature18635

McNeill, W. H. (1995). *Keeping Together in Time: Dance and Drill in Human History*. Cambridge: Harvard University Press.

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental Science*, *10*(1), 126–134. https://doi.org/10.1111/j.1467-7687.2007.00574.x

Merker, B. (2000). 'Synchronous Chorusing and Human Origins', in N. L. Wallin, B. Merker & S. Brown (Eds.) *The Origins of Music*, (pp. 315-327), Cambridge: MIT Press.

Merker, B., Morley, I., & Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0095

Merriam, A. P (1964). *The Anthropology of Music*. Evanston: Northwestern University Press.

Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, C. N. (2011). Do birds of a feather move together? Group membership and behavioral synchrony. *Experimental Brain Research*, *211*(3-4), 495–503. https://doi.org/10.1007/s00221-011-2641-z

Miles, L. K., Nind, L. K., Henderson, Z., & Macrae, C. N. (2010). Moving memories: Behavioral synchrony and memory for self and others. *Journal of Experimental Social Psychology*, *46*(2), 457–460. https://doi.org/10.1016/j.jesp.2009.12.006

Miller, G. F. (2000). 'Evolution of human music through sexual selection', in N. L. Wallin, B. Merker & S. Brown (Eds.) *The Origins of Music*, (pp. 271-300), Cambridge: MIT Press.

Modigliani, A. (1971). Embarrassment, facework, and eye contact: Testing a theory of embarrassment. *Journal of Personality and Social Psychology*, *17*(1), 15–24. https://doi.org/10.1037/h0030460

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian Intelligence Hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1480), 639–648. https://doi.org/10.1098/rstb.2006.2000

Molleman, L., Quiñones, A. E., & Weissing, F. J. (2013). Cultural evolution of cooperation: The interplay between forms of social learning and group selection. *Evolution and Human Behavior*, *34*(5), 342–349. https://doi.org/10.1016/j.evolhumbehav.2013.06.001

Molleson, K. (2013). 'A quiet word about classical concert etiquette' in *The Guardian* (https://www.theguardian.com/music/2013/sep/10/noise-classical-concert-etiquette-kate-molleson) accessed 27 June 2019.

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons : from motion to 'e'motion. *Social Cognitive and Affective Neuroscience*, *1*(3), 235–241. https://doi.org/10.1093/scan/nsl029

Moran, N. (2010). 'Improvising musicians' looking behaviours: Duration constants in the attention patterns of duo performers', in S. Demorest, S. Morrison, & P. S. Campbell (Eds.) *Proceedings of the 11th international conference on Music Perception and Cognition (ICMPC11)*.

Moran, N. (2013). 'Social Co-Regulation and Communication in North Indian Duo Performances', in M. Clayton, B. Dueck, & L. Leante (Eds.) *Experience and Meaning in Music Performance* (pp. 40-61), New York: Oxford University Press.

Moran, N. (2014). Social implications arise in embodied music cognition research which can counter musicological "individualism". *Frontiers in Psychology*, *5*(676). https://doi.org/10.3389/fpsyg.2014.00676

Morley, I. (2012). 'A grand gesture: vocal and corporeal control in melody, rhythm, and emotion', in P. Rebuschat, M. Rohrmeier, J. A. Hawkins, & I. Cross (Eds.) *Language and Music as Cognitive Systems* (pp. 110-120), Oxford: Oxford University Press.

Morley, I. (2013). *The prehistory of music: human evolution, archaeology, and the origins of musicality*. Oxford: Oxford University Press.

Mu, Y., Guo, C., & Han, S. (2016). Oxytocin enhances inter-brain synchrony during social coordination in male adults. *Social Cognitive and Affective Neuroscience*, *11*(12), 1882-1893. https://doi.org/10.1093/scan/nsw106

Mullen, B., Migdal, M. J., & Hewstone, M. (2001). Crossed categorization versus simple categorization and intergroup evaluations: a meta-analysis. *European Journal of Social Psychology*, *31*(6), 721–736. https://doi.org/10.1002/ejsp.60

Murdoch, I. (2004) *The Bell*. London: Vintage Books.

Naiman, R. J., Johnston, C. A., & Kelley, J. C. (1988). Alteration of North American Streams by Beaver. *BioScience*, *38*(11), 753–762. https://doi.org/10.2307/1310784

Nave, G., Camerer, C., & McCullough, M. (2015). Does Oxytocin Increase Trust in Humans? A Critical Review of Research. *Perspectives on Psychological Science*, *10*(6), 772-789. https://doi.org/10.1177/1745691615600138

Nettle, D. (1999). 'Language variation and the evolution of societies', in R. Dunbar, C. Knight & C. Power (Eds.) *The Evolution of Culture* (pp. 214-227), Edinburgh: Edinburgh University Press.

Nettle, D., & Dunbar, R. I. M. (1997). Social Markers and the Evolution of Reciprocal Exchange. *Current Anthropology*, *38*(1), 93–99. https://doi.org/10.1086/204588

Nichols, A. L., & Webster, G. D. (2013). The single-item need to belong scale. *Personality and Individual Differences*, *55*(2), 189–192. https://doi.org/10.1016/j.paid.2013.02.018

Norenzayan, A., & Heine, S. J. (2005). Psychological Universals: What Are They and How Can We Know? *Psychological Bulletin*, *131*(5), 763–784. https://doi.org/10.1037/0033-2909.131.5.763

North, A. C., & Hargreaves, D. J. (1999). Music and adolescent identity. *Music Education Research*, *1*(1), 75–92. https://doi.org/10.1080/1461380990010107

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

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation, *Science*, *314*(5805), 1560-1563. https://doi.org/10.1126/science.1133755

Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of neonates to parents' and others' voices. *Early Human Development*, *18*(1), 27–36. https://doi.org/10.1016/0378-3782(88)90040-0

Odling-Smee, F. J., Laland, K. N. & Feldman, M. W. (2003) *Niche construction: the neglected process in evolution*. Oxford: Princeton University Press.

Onishi, K. K., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Over, H., & Carpenter, M. (2013). The Social Side of Imitation. *Child Development Perspectives*, *7*(1), 6–11. https://doi.org/10.1111/cdep.12006

Overy, K., & Molnar-Szakacs, I. (2009). Being Together in Time: Musical Experience and the Mirror Neuron System. *Music Perception: An Interdisciplinary Journal*, *26*(5), 489–504. doi:10.1525/mp.2009.26.5.489

Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). 'Introduction: What Is Developmental Systems Theory?' in R. D. Gray, P. E. Griffiths, & S. Oyama (Eds.) *Cycles of Contingency: Developmental Systems and Evolution* (pp. 1-11), London: MIT Press.

Papoušek, H. (1996a). 'Musicality in infancy research: biological and cultural origins of early musicality', in I. Deliège & J. A. Sloboda (Eds.) *Musical beginnings: origins and development of musical competence* (pp. 37-55), Oxford: Oxford University Press.

Papoušek, M. (1996b). 'Intuitive parenting: a hidden source of musical stimulation in infancy', in I. Deliège & J. A. Sloboda (Eds.) *Musical beginnings: origins and development of musical competence* (pp. 88-112), Oxford: Oxford University Press.

Parsons, T., & Shils, E. A. (1951). 'Values, Motives, and Systems of Action', in T. Parsons & E. A. Shils (Eds.), *Toward a General Theory of Action* (pp. 47-273), Cambridge, MA: Harvard University Press.

Patel, A. D. (2010). 'Music, biological evolution, and the brain', in M. Bailar, C. Field, & C. J. Henry (Eds.) *Emerging Disciplines*, (pp. 91-144), Houston: Rice University Press

Paxton, A., & Dale, R. (2013). Multimodal Networks of Interpersonal Interaction and Conversational Contexts. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*, 1121–1126.

Pearce, E. (2014). Modelling mechanisms of social network maintenance in hunter-gatherers. *Journal of Archaeological Science*, *50*, 403–413. https://doi.org/10.1016/j.jas.2014.08.004

Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: singing mediates fast social bonding. *Royal Society Open Science*, *2*(10). https://doi.org/10.1098/rsos.150221

Pearce, E., Launay, J., MacCarron, P., & Dunbar, R. I. M. (2016b). Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. *Psychology of Music*, *14*(4), 496-512. https://doi.org/10.1177/0305735616667543

Pearce, E., Launay, J., van Duijn, M., Rotkirch, A., David-Barrett, T., & Dunbar, R. I. M. (2016a). Singing together or apart: The effect of competitive and cooperative singing on social bonding within and between sub-groups of a university Fraternity. *Psychology of Music*, *44*(6), 1255–1273. https://doi.org/10.1177/0305735616636208

Pepitone, A. (1981). Lessons from the history of social psychology. *American Psychologist*, *36*(9), 972-985.

Pettigrew, T. F. (1998). Intergroup Contact Theory. *Annual Review of Psychology*, *49*(1), 65–85. https://doi.org/10.1146/annurev.psych.49.1.65

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the Beat: Movement Influences Infant Rhythm Perception. *Science*, *308*(5727), 1430. https://doi.org/10.1126/science.1110922

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a Cue: The Need to Belong and Enhanced Sensitivity to Social Cues. *Personality and Social Psychology Bulletin*, *30*(9), 1095–1107. https://doi.org/10.1177/0146167203262085

Pinker, S. (1997). *How the mind works*. London: Allen Lane.

Platow, M. J., Grace, D. M., & Smithson, M. J. (2012). Examining the Preconditions for Psychological Group Membership: Perceived Social Interdependence as the Outcome of Self-Categorization. *Social Psychological and Personality Science*, *3*(1), 5–13. https://doi.org/10.1177/1948550611407081

Platow, M. J., Haslam, S. A., Both, A., Chew, I., Cuddon, M., Goharpey, N., et al. (2005). "It's not funny if *they're* laughing": Self-categorization, social influence, and responses to canned laughter. *Journal of Experimental Social Psychology*, *41*(1), 542–550. https://doi.org/10.1016/j.jesp.2004.09.005

Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology*, *80*(6), 918–930. https://doi.org/10.1037/0022-3514.80.6.918

Prentice, D. A., Miller, D. T., & Lightdale, J. R. (1994). Asymmetries in Attachments to Groups and to Their Members: Distinguishing Between Common-Identity and Common-Bond Groups. *Personality and Social Psychology Bulletin*, *20*(5), 484–493. https://doi.org/10.1177/0146167294205005

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1-20. https://doi.org/10.1017/S0140525X02000018

Rabbie, J. M., & Horwitz, M. (1988). Categories versus groups as explanatory concepts in intergroup relations. *European Journal of Social Psychology*, *18*(2), 117–123. https://doi.org/10.1002/ejsp.2420180204

Rabinowitch, T.-C., & Knafo-Noam, A. (2015). Synchronous Rhythmic Interaction Enhances Children's Perceived Similarity and Closeness towards Each Other. *PLoS ONE*, *10*(4), e0120878. https://doi.org/10.1371/journal.pone.0120878

Rabinowitch, T.-C., & Meltzoff, A. N. (2017). Synchronized movement experience enhances peer cooperation in preschool children. *Journal of Experimental Child Psychology*, *160*, 21–32. https://doi.org/10.1016/j.jecp.2017.03.001

Rabinowitch, T-C., Cross, I., & Burnard, P. (2012a). 'Musical Group Interaction, Intersubjectivity and Merged Subjectivity', in D. Reynolds & M. Reason (Eds.) *Kinesthetic Empathy in Creative and Cultural Practices* (pp. 109-120), Bristol: Intellect Press.

Rabinowitch, T.-C., Cross, I., & Burnard, P. (2012b). Long-term musical group interaction has a positive influence on empathy in children. *Psychology of Music*, *41*(4), 484–498. https://doi.org/10.1177/0305735612440609

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology*, *28*(4), 785–798. https://doi.org/10.1348/026151009X479178

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The Sources of Normativity : Young Children's Awareness of the Normative Structure of Games, *Developmental Psychology*, *44*(3), 875–881. https://doi.org/10.1037/0012-1649.44.3.875

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology*, *49*(2), 298–301. https://doi.org/10.1016/j.jesp.2012.10.017

Rauchbauer, B., Majdandžić, J., Hummer, A., Windischberger, C., & Lamm, C. (2015). Distinct neural processes are engaged in the modulation of mimicry by social group-membership and emotional expressions. *Cortex*, *70*, 49–67. https://doi.org/10.1016/j.cortex.2015.03.007

Reddish, P., Tong, E. M. W., Jong, J., Lanman, J. A., & Whitehouse, H. (2016). Collective synchrony increases prosociality towards non-performers and outgroup members. *British Journal of Social Psychology*, *55*(4), 722–738. https://doi.org/10.1111/bjso.12165

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A Meta-Analysis. *Zeitschrift Fur Psychologie*, *224*(3), 168–189. https://doi.org/10.1027/2151-2604/a000252

Revesz, R. (2016, 14 July). Black Republican senator says he was stopped seven times in one year by policy, in *The Independent*. https://www.independent.co.uk/news/world/americas/tim-scott-black-republican-senator-racism-floor-speech-stopped-seven-times-one-year-senate-a7137086.html (accessed 23/10/2019)

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

Richerson, P. J. & Boyd, R. (2005, 2006 edition). *Not by genes alone: how culture transformed human evolution*. Chicago: University of Chicago Press.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A Neural Basis for Social Cooperation. *Neuron*, *35*(2), 395–405. https://doi.org/10.1016/S0896-6273(02)00755-9

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, *21*(5), 188–194. https://doi.org/10.1016/S0166-2236(98)01260-0

Roberts, S. G. B., & Dunbar, R. I. M. (2011). Communication in social networks: Effects of kinship, network size, and emotional closeness. *Personal Relationships*, *18*(3), 439–452. https://doi.org/10.1111/j.1475-6811.2010.01310.x

Roederer, J. G. (1984). The Search for a Survival Value of Music. *Music Perception*, *1*(3), 350–356.

Rohlf, Michael, "Immanuel Kant", *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/kant/>.

Roth, T. L., Lubin, F. D., Sodhi, M. & Kleinman, J. E. (2009). Epigenetic mechanisms in schizophrenia. *Biochimica et Biophysica Acta*, *1790*(9), 869-877. https://doi.org/10.1016/j.bbagen.2009.06.009

Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist*, *26*(5), 443-452.

Rowley-Conwy, P., & Layton, R. (2011). Foraging and farming as niche construction: stable and unstable adaptations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366* (1566), 849–862. https://doi.org/10.1098/rstb.2010.0307

Rozin, P. (2001). Social Psychology and Science: Some Lessons From Solomon Asch. *Personality and Social Psychology Review*, *5*(1), 2–14. https://doi.org/10.1207/S15327957PSPR0501_1

Rubin, M., & Hewstone, M. (1998). Social Identity Theory's Self-Esteem Hypothesis: A Review and Some Suggestions for Clarification. *Personality and Social Psychology Review*, *2*(1), 40–62. https://doi.org/10.1207/s15327957pspr0201_3

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. *Psychological Bulletin*, *115*(1), 102-141.

Rutchick, A. M., Hamilton, D. L., & Sack, J. D. (2008). Antecedents of entitativity in categorically and dynamically construed groups. *European Journal of Social Psychology*, *38*(6), 905–921. https://doi.org/10.1002/ejsp.555

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109. https://doi.org/10.1177/0305735607068889

Sabini, J., Siepmann, M., Stein, J., & Meyerowitz, M. (2000). Who is Embarrassed by What? *Cognition and Emotion*, *14*(2), 213–240. https://doi.org/10.1080/026999300378941

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928. DOI: 10.1126/science.274.5294.1926

Savage, P. E. (2018). Alan Lomax's Cantometrics Project: A comprehensive review. *Music & Science, 1*, 1-19. https://doi.org/10.1177/2059204318786084

Sawyer, R. K. (2005). 'Music and conversation', in D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.) *Musical communication* (pp. 45-60), Oxford: Oxford University Press.

Scheele, D., Wille, A., Kendrick, K. M., Stoffel-Wagner, B., Becker, B., Güntürkün, O. et al. (2013). Oxytocin enhances brain reward system responses in men viewing the face of their female partner. *Proceedings of the National Academy of Sciences*, *110*(50), 20308–20313. https://doi.org/10.1073/pnas.1314190110

Schellenberg, E. G., Corrigall, K. A., Dys, S. P., & Malti, T. (2015). Group Music Training and Children's Prosocial Skills. *PLoS ONE*, *10*(10), e0141449. https://doi.org/10.1371/journal.pone.0141449

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(4), 393–414. https://doi.org/10.1017/S0140525X12002452

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G. et al. (2010). Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry. *Journal of Cognitive Neuroscience*, *22*(12), 2702–2715. https://doi.org/10.1162/jocn.2009.21401

Schlegoff, E. A. (2000). Overlapping Talk and the Organization of Turn-Taking for Conversation. *Language in Society*, *29*(1), 1–63.

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

Schober, M. F. & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211-232. https://doi.org/10.1016/0010-0285(89)90008-X

Schubert, T. W., & Otten, S. (2002). Overlap of Self, Ingroup, and Outgroup: Pictorial Measures of Self-Categorization. *Self and Identity*, *1*(4), 353–376. https://doi.org/10.1080/152988602760328012

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*, 199-207. doi:10.1038/35044563

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

Seeger, A. (1987). *Why Suyá sing: a musical anthropology of an Amazonian people*. Cambridge: Cambridge University Press.

Seyfarth, R. M., & Cheney, D. L. (2003). Meaning and Emotion in Animal Vocalizations. *Annals of the New York Academy of Sciences*, *1000*(1), 32–55. https://doi.org/10.1196/annals.1280.004

Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. *Biological Psychiatry*, *79*(3), 194–202. https://doi.org/10.1016/j.biopsych.2015.07.020

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*(3), 617–627. https://doi.org/10.1093/brain/awn279

Shaw, M. E. (1976). *Group dynamics: the psychology of small group behaviour*. New York: McGraw-Hill (2nd edition).

Sherif, M. (1967). *Group conflict and co-operation: their social psychology*. London: Routledge and Kegan Paul.

Shevy, M. (2008). Music genre as cognitive schema: Extramusical associations with country and hip-hop music. *Psychology of Music*, *36*(4), 477–498. https://doi.org/10.1177/0305735608089384

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and Coordinative Structures. *Topics in Cognitive Science*, *1*, 305–319. https://doi.org/10.1111/j.1756-8765.2009.01021.x

Simmel, G., & Hughes, E. C. (1949). The Sociology of Sociability. *American Journal of Sociology, 55*(3), 254-261. https://doi.org/10.1086/220534

Singer, T., & Lamm, C. (2009). The Social Neuroscience of Empathy. *Annals of the New York Academy of Sciences*, *1156*(1), 81–96. https://doi.org/10.1111/j.1749-6632.2009.04418.x

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *189*, 260–274. https://doi.org/10.1016/j.cognition.2019.03.019

Smaldino, P. E. (2014). The cultural evolution of emergent group-level traits. *Behavioural and Brain Sciences*, *37*(3), 243–254. https://doi.org/10.1017/S0140525X13001544

Small, C. (1998). *Musicking: the meanings of performing and listening*. Connecticut: Wesleyan University Press.

Smith, E. R., Murphy, J., & Coats, S. (1999). Attachment to Groups: Theory and Measurement. *Journal of Personality and Social Psychology*, *77*(1), 94–110.

Smith, S. M. (1986). Environmental context-dependent recognition memory using a short-term memory task for input. *Memory & Cognition*, *14*(4), 347–354. https://doi.org/10.3758/BF03202513

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Soley, G. (2019). What Do Group Members Share? The Privileged Status of Cultural Knowledge for Children. *Cognitive Science*, *43*(e12786). https://doi.org/10.1111/cogs.12786

Sommerville, J. A., & Hammond, A. J. (2007). Treating another's actions as one's own: children's memory of and learning from joint activity. *Developmental Psychology*, *43*(4), 1003–1018. https://doi.org/10.1037/0012-1649.43.4.1003

Song, S., Miller, K. & Abbott, L. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience 3*, 919–926. https://doi.org/10.1038/78829

Sperber, D., & Wilson, D. (2012). Introduction. In D. Wilson & D. Sperber, *Meaning and Relevance* (pp. 1-28). Cambridge: Cambridge University Press (accessed online). doi:10.1017/CBO9781139028370.002

Spoor, J. R., & Kelly, J. R. (2004). The Evolutionary Significance of Affect in Groups: Communication and Group Bonding. *Group Processes & Intergroup Relations*, *7*(4), 398–412. https://doi.org/10.1177/1368430204046145

Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development*, *7*(4), 349–367. https://doi.org/10.1080/14616730500365928

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*(6), 1467–1478. https://doi.org/10.1037/0022-3514.48.6.1467

Steiner, I. D. (1974). Whatever happened to the group in social psychology? *Journal of Experimental Social Psychology*, *10*(1), 94–108. https://doi.org/10.1016/0022-1031(74)90058-4

Stephan, K. M. , Thaut, M. H., Wunderlich, G., Schicks, W., Tian, B., Tellmann, L., et al. (2002). Conscious and Subconscious Sensorimotor Synchronization - Prefrontal Cortex and the Influence of Awareness. *NeuroImage*, *15*(2), 345-352. https://doi.org/10.1006/nimg.2001.0929

Sterelny, K. (2011). From hominins to humans: how *sapiens* became behaviourally modern. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(166), 809–822. https://doi.org/10.1098/rstb.2010.0301

Sterne, J. (2006). The mp3 as cultural artifact. *New Media & Society*, *8*(5), 825-842. DOI: 10.1177/1461444806067737

Stevens, C. J. & Byron, T. (2016). 'Universals in music processing: Entrainment, acquiring expectations, and learning', in S. Hallam, I. Cross, & M. Thaut (Eds.) *The Oxford Handbook of Music Psychology* (pp. 19-31), Oxford: Oxford University Press (2nd edition).

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Heinemann, T., Hoymann, G., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592. DOI: 10.1073/pnas.0903616106

Stobart, H. (1994). Flourishing horns and enchanted tubers: Music and potatoes in highland Bolivia. *British Journal of Ethnomusicology*, *3*(1), 35–48. https://doi.org/10.1080/09681229408567224

Stobart, H. (1996). The Llama's Flute: Musical Misunderstandings in the Andes. *Early Music*, *24*(3), 470–482.

Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1585), 75–87. https://doi.org/10.1098/rstb.2011.0099

Streeck, J., Goodwin, C., & LeBaron, C. D. (2011). 'Embodied Interaction in the Material World: An Interaction', in J. Streeck, C. Goodwin, & C. D. LeBaron (Eds.) *Embodied interaction: language and body in the material world* (pp. 1-26), Cambridge: Cambridge University Press.

Stroebe, K., Lodewijkx, H. F. M., & Spears, R. (2005). Do Unto Others as They Do Unto You: Reciprocity and Social Identification as Determinants of Ingroup Favoritism. *Personality and Social Psychology Bulletin*, *31*(6), 831–845. https://doi.org/10.1177/0146167204271659

Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-Motivated Helping: The Moderating Role of Group Membership. *Personality and Social Psychology Bulletin*, *32*(7), 943–956. https://doi.org/10.1177/0146167206287363

Sutcliffe, A., Dunbar, R., Binder, J., & Arrow, H. (2012). Relationships and the social brain: Integrating psychological and evolutionary perspectives. *British Journal of Psychology*, *103*(2), 149–168. https://doi.org/10.1111/j.2044-8295.2011.02061.x

Swann, W. B., Gómez, Á., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*(5), 995–1011. https://doi.org/10.1037/a0013668

Tajfel, H., & Turner, J. C. (1979, 1986 edition). 'An integrative theory of intergroup conflict', in W. G. Austin & S. Worchel (Eds.) *Psychology of intergroup relations* (pp. 33–47), Chicago: Nelson-Hall. (2nd edition)

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: "self-other" merging and neurohormonal mechanisms. *Frontiers in Psychology*, *5*(1096). https://doi.org/10.3389/fpsyg.2014.01096

Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, *37*(5), 343–349. https://doi.org/10.1016/j.evolhumbehav.2016.02.004

Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10). https://doi.org/10.1098/rsbl.2015.0767

Tëmkin, I., & Eldredge, N. (2007). Phylogenetics and Material Cultural Evolution. *Current Anthropology*, *48*(1), 146–154. https://doi.org/10.1086/510463

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Thagard, P. (2018). 'Cognitive Science', in in E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. California: Stanford University Press (Spring 2019 edition, https://plato.stanford.edu/archives/spr2019/entries/cognitive-science/, accessed 1st July 19).

Thibaut, J. W. & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tolbert, E. (1994). 'The voice of lament: female vocality and performative efficacy in the the Finnish-Karelian *itkuvirsi*', in L. C. Dunn & N. A. Jones (Eds.) *Embodied voices: representing female vocality in Western culture* (pp. 179-194), Cambridge: Cambridge University Press.

Tolbert, E. (2001). Music and Meaning: An Evolutionary Story. *Psychology of Music*, *29*(1), 84–94. https://doi.org/10.1177/0305735601291006

Tomasello, M. (2008). *Origins of Human Communication*. London: MIT Press

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125. https://doi.org/10.1111/j.1467-7687.2007.00573.x

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. https://doi.org/10.1016/j.jhevol.2006.10.001

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps in the Evolution of Human Cooperation. *Current Anthropology*, *53*(6), 673–692. https://doi.org/10.1086/668207

Tomlinson, G. (2015). *A million years of music: the emergence of human modernity*. New York: Zone Books.

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0089

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is Infant-Directed Speech Prosody a Result of the Vocal Expression of Emotion? *Psychological Science*, *11*(3), 188–195. https://doi.org/10.1111/1467-9280.00240

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*(1), 73–99. https://doi.org/10.1016/j.cognition.2005.11.006

Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*:20140096. https://doi.org/10.1098/rstb.2014.0096

Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae*, *3*(1_suppl), 155–215. https://doi.org/10.1177/10298649000030S109

Trezza, V., Baarendse, P. J. J., & Vanderschuren, L. J. M. J. (2010). The pleasures of play: pharmacological insights into social reward mechanisms. *Trends in Pharmacological Sciences*, *31*(10), 463–469. https://doi.org/10.1016/j.tips.2010.06.008

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Tropp, L. R., & Wright, S. C. (2001). Ingroup Identification as the Inclusion of Ingroup in the Self. *Personality and Social Psychology Bulletin*, *27*(5), 585–600. https://doi.org/10.1177/0146167201275007

Tunçgenç, B., & Cohen, E. (2016). Movement Synchrony Forges Social Bonds across Group Divides. *Frontiers in Psychology*, *7*(782). https://doi.org/10.3389/fpsyg.2016.00782

Turino, T. (2008). *Music as social life: the politics of participation*. Chicago: University of Chicago Press.

Turino, T. (2009). Four Fields of Music Making and Sustainable Living. *The World of Music, 51*(1), 95-117.

Turner, J. C., & Bourhis, R. Y. (1996). 'Social identity, interdependence and the social group: A reply to Rabbie et al.', in W. P. Robinson (Ed.), *Social Groups and Identities: Developing the Legacy of Henri Tajfel* (pp. 25–63). Oxford, UK: Butterworth-Heinemann.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, *25*(3), 237–252. https://doi.org/10.1111/j.2044-8309.1986.tb00732.x

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: a self-categorization theory*. Oxford: Basil Blackwell.

Turner, J. C., Sachdev, I., & Hogg, M. A. (1983). Social categorization, interpersonal attraction and group formation. *British Journal of Social Psychology*, *22*(3), 227–239. https://doi.org/10.1111/j.2044-8309.1983.tb00587.x

Turner, J. S. (2000). *The extended organism: the physiology of animal-built structures*. London: Harvard University Press.

Turner, V. (1969, 2002 reprint). Liminality and Communitas. In M. Lambek (Ed.) *A Reader in the Anthropology of Religion* (pp. 358-374). Oxford: Blackwell.

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences*, *35*(4), 203-218. https://doi.org/10.1017/S0140525X11001452

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, *46*(4), 693–695. https://doi.org/10.1016/j.jesp.2010.03.004

van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and Prosocial Behavior. *Psychological Science*, *15*(1), 71–74. https://doi.org/10.1111/j.0963-7214.2004.01501012.x

van der Kleij, R., Schraagen, J. M., Werkhoven, P., & De Dreu, C. K. W. (2009). How Conversations Change Over Time in Face-to-Face and Video-Mediated Communication. *Small Group Research*, *40*(4), 355–381. https://doi.org/10.1177/1046496409333724

van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind: Exploring the possibilities of 4E cognition and dynamical systems theory. *Music & Science*, *1*, 1-18. https://doi.org/10.1177/2059204318792319

van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender Differences in Cooperation and Competition. *Psychological Science*, *18*(1), 19–23. https://doi.org/10.1111/j.1467-9280.2007.01842.x

Vandemoortele, S., Feyaerts, K., Reybrouck, M., De Bièvre, G., Brône, G. & De Baets, T. (2018). Gazing at the partner in musical trios: a mobile eye-tracking study. *Journal of Eye Movement Research*, *11*(2). https://doi.org/10.16910/jemr.11.2.6

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*(8–9), 998–1003. https://doi.org/10.1016/j.neunet.2010.06.002

Vesper, C., 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

von dem Hagen, E. A. H., Beaver, J. D., Ewbank, M. P., Keane, J., Passamonti, L., Lawrence, A. D., & Calder, A. J. (2009). Leaving a bad taste in your mouth but not in my insula. *Social Cognitive and Affective Neuroscience*, *4*(4), 379–386. https://doi.org/10.1093/scan/nsp018

von Zimmermann, J., Vicary, S., Sperling, M., Orgs, G., & Richardson, D. C. (2018). The Choreography of Group Affiliation. *Topics in Cognitive Science*, *10*, 80–94. https://doi.org/10.1111/tops.12320

Wade, M. J., Wilson, D. S., Goodnight, C., Taylor, D., Bar-Yam, Y., de Aguiar, M. A. M., et al. (2010). Multilevel and kin selection in a connected world. *Nature*, *463*, E8–E9. https://doi.org/10.1038/nature08809

Wai, M., & Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences*, *52*(7), 794–799. https://doi.org/10.1016/j.paid.2012.01.008

Wai-man Ip, G., Chiu, C., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology*, *90*(3), 368–381. https://doi.org/10.1037/0022-3514.90.3.368

Wang, Y., & Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Frontiers in Human Neuroscience*, *6*(153). https://doi.org/10.3389/fnhum.2012.00153

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158-177. https://doi.org/10.1037/h0074428

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (1967, 2014 edition) *Pragmatics of human communication: a study of interactional patterns, pathologies, and paradoxes*. New York: Norton

Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, *7*(8), 847–854. https://doi.org/10.1038/nn1276

Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, *37*(2), 152–158. https://doi.org/10.1016/j.evolhumbehav.2015.10.002

Wellman, H. M., Cross, D., & Watson, J. (2010). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, *72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

West, S. A., Griffin, A. S., & Gardner, A. (2007a). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, *20*(2), 415–432. https://doi.org/10.1111/j.1420-9101.2006.01258.x

West, S. A., Griffin, A. S., & Gardner, A. (2007b). Social semantics: How useful has group selection been? *Journal of Evolutionary Biology*, *21*(1), 374–385. https://doi.org/10.1111/j.1420-9101.2007.01458.x

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press (accessed online).

Whiten, A. (2017). Social Learning and Culture in Child and Chimpanzee. *Annual Review of Psychology*, *68*(1), 129-154. https://doi.org/10.1146/annurev-psych-010416-044108

Whiten, A., & Flynn, E. (2010). The transmission and evolution of experimental microcultures in groups of young children. *Developmental Psychology, 46*(6), 1694–1709. https://doi.org/10.1037/a0020786

Wiessner, P. W. (2014). Embers of society: Firelight talk among the Ju'hoansi Bushmen. *Proceedings of the National Academy of Sciences*, *111*(39), 14027–14035. https://doi.org/10.1073/pnas.1404212111

Wild, G., Gardner, A. & West, S. A. (2009). Adaptation and the evolution of parasite virulence in a connected world. *Nature* 459, 983–986. doi:10.1038/nature08071

Wilson, D. S. (2007). Social semantics: Toward a genuine pluralism in the study of social behaviour. *Journal of Evolutionary Biology*, *21*(1), 368–373. https://doi.org/10.1111/j.1420-9101.2007.01396.x

Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioural and Brain Sciences*, *17*(4), 585–608. https://doi.org/10.1017/S0140525X00036104

Wilson, D. S. & Sober, E. (1998).Multilevel selection and the return of group-level functionalism. *Behavioral and Brain Sciences*, *21*(2), 305-306. https://doi.org/10.1017/S0140525X98221194

Wilson, D. S., van Vugt, M., & Gorman, R. O. (2008). Multilevel Selection Theory and Major Evolutionary Transitions Implications for Psychological Science, *17*(1), 6-9. https://doi.org/10.1111/j.1467-8721.2008.00538.x

Wilson, G. B., & MacDonald, R. A. R. (2016). Musical choices during group free improvisation: A qualitative psychological investigation. *Psychology of Music*, *44*(5), 1029-1043. https://doi.org/10.1177/0305735615606527

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin and Review*, *12*(6), 957–968. https://doi.org/10.3758/BF03206432

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Woodburn, J. (1982). Egalitarian Societies. *Man*, *17*(3), 431–451.

Woolhouse, M. H., & Lai, R. (2014). Traces across the body: influence of music-dance synchrony on the observation of dance. *Frontiers in Human Neuroscience*, *8*(965). https://doi.org/10.3389/fnhum.2014.00965

Woolhouse, M. H., Tidhar, D., & Cross, I. (2016). Effects on Inter-Personal Memory of Dancing in Time with Others. *Frontiers in Psychology*, *7*(167). https://doi.org/10.3389/fpsyg.2016.00167

Wrightsman, L. S. (1991). 'Interpersonal trust and attitudes toward human nature', in J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes, Vol. 1. Measures of personality and social psychological attitudes* (pp. 373-412), London: Academic Press.

Wynne-Edwards, V. C. (1962). *Animal dispersion in relation to social behaviour.* Edinburgh: Oliver & Boyd.

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *Journal of Neuroscience*, *29*(26), 8525–8529. https://doi.org/10.1523/JNEUROSCI.2418-09.2009

Yabar, Y., Johnston, L., Miles, L., & Peace, V. (2006). Implicit Behavioral Mimicry: Investigating the Impact of Group Membership. *Journal of Nonverbal Behavior*, *30*(3), 97–113. https://doi.org/10.1007/s10919-006-0010-6

Yamagishi, T., & Hashimoto, H. (2016). Social niche construction. *Current Opinion in Psychology*, *8*, 119–124. https://doi.org/10.1016/j.copsyc.2015.10.003

Yamagishi, T., & Kiyonari, T. (2000). The Group as the Container of Generalized Reciprocity. *Social Psychology Quarterly*, *63*(2), 116-132. https://doi.org/10.2307/2695887

Yamagishi, T., & Mifune, N. (2008). Does Shared Group Membership Promote Altruism?: Fear, Greed, and Reputation. *Rationality and Society*, *20*(1), 5–30. https://doi.org/10.1177/1043463107085442

Zak, P. J., Kurzban, R., & Matzner, W. T. (2005). Oxytocin is associated with human trustworthiness. *Hormones and Behavior*, *48*(5), 522–527. https://doi.org/10.1016/j.yhbeh.2005.07.009

Ziv, N., Hoftman, M., & Geyer, M. (2011). Music and moral judgment: The effect of background music on the evaluation of ads promoting unethical behavior. *Psychology of Music*, *40*(6), 738-760.

'music, n. and adj.' *Oxford English Dictionary* (Third Edition, 2003, Oxford: Oxford University Press, www.oed.com/view/Entry/124108) accessed 13 June 2019.

Boersma, P. & Weenick, D. (2019). *Praat: doing phonetics by computer [computer program].* Retrieved March 2019 from http://www.praat.org.

ELAN [computer software]. (2019). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.