https://doi.org/10.35662/unine-thesis-2818

UNDERSTANDING TOGETHERNESS: JOINT ACTION CAPACITIES IN GREAT APES

PhD thesis submitted to the Faculty of Science

Institute of Biology

University of Neuchâtel

By

Raphaela HEESEN

Viva: November 8th 2019

Expert committee composition:

Thesis co-director – Prof. Klaus Zuberbühler, University of Neuchâtel, Switzerland

Thesis co-director – Prof. Adrian Bangerter, University of Neuchâtel, Switzerland

Internal examiner – Dr . Emilie Genty, University of Neuchâtel, Switzerland

External examiner - Prof. Carel van Schaik, University of Zürich, Switzerland

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Madame Raphaela HEESEN

Titre:

"UNDERSTANDING TOGETHERNESS: JOINT ACTION CAPACITIES IN GREAT APES"

sur le rapport des membres du jury composé comme suit:

- Prof. Adrian Bangerter, co-directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Klaus Zuberbühler, co-directeur de thèse, Université de Neuchâtel
- Prof. Carel van Schaik, Université de Zürich, Suisse
- Dr Emilie Genty, Université de Neuchâtel, Suisse

Neuchâtel, le 5 décembre 2019

Le Doyen, Prof. P. Felber

P. Fel

Abstract

Social animals cooperate in manifold ways towards a common outcome, for example by performing complementary and reciprocal roles, acting pro-socially, and adjusting their own actions in concert with those of their partners. Despite this, researchers have claimed that the cooperative interactions of non-human animal species –notably apes- remain egoistic in nature, insofar that the participants lack the cognitive capacity to share intentions for the establishment of a joint commitment toward a shared goal. Animals supposedly fail to engage in joint action via *shared intentionality*.

This theory incurs several problems, however. It assumes a suite of high-level cognitive capacities, which may not be necessary for joint action coordination. Moreover, the evidence on which it is predominantly built relies on comparative laboratory experiments between captive apes and human children, which lack ecological relevance and comparability of cross-species designs, and mostly involve mostly chimpanzees, known to be the most competitive of all apes. Recent evidence furthermore provides new insights into joint action coordination in apes, warranting the need for revisiting the theory.

In that regard, my thesis comprised two major goals– the first is of theoretical and the second of empirical nature. My first goal was to address the aforementioned shortcomings by re-conceptualizing the theory of shared intentionality theoretically. I did so by implementing an alternative framework, inspired by human joint action research, allowing for the study of how two participants co-create, co-maintain, and co-dissolve a state of *joint commitment* in so-called joint action phases of entry, main body and exit. In addition, when humans engage in joint action, the participants are constantly concerned with *face management* – i.e., using politeness to mitigate face threatening acts, with face being defined as a form of publicly manifest self-esteem. The behaviors and communicative signals used by participants in the joint action phases serve as yardsticks for their sensitivity to joint commitment and face management. The proposed framework thus allows for studying those behavioral prerequisites minimally required to achieve a higher-level state of shared intentionality through the process of joint action coordination.

My second goal was to empirically apply the framework to the study of great apes' (here chimpanzees and bonobos) natural interactions, more specifically social and social grooming and play. My findings revealed that both chimpanzees and bonobos exhibit recognizable entry and exit phases in grooming and play interactions, and display recognizable reengagement attempts after

interruptions of social activities. Additionally, the communication used by bonobos during joint action phases depended on partner-specific social attributes such as social bond and rank differences. These findings suggest that apes indeed exhibit several of the behavioral evidence for joint commitment understanding and face management, thus indicating that apes are capable of collaboratively constructing a state of shared intentionality through the process of joint action coordination. My findings raise questions about the role of joint commitment and face management in primate evolution, and stress the idea that human uniqueness has been shaped by ancestral primate roots. I hope my thesis will revive the debate on whether shared intentionality constitutes a uniquely human trait.

Key words • Joint action, communication, shared intentionality, politeness theory, great apes

Résumé

Les animaux sociaux coopèrent de multiples façons dans le but d'atteindre un objectif commun, par exemple en remplissant des rôles complémentaires et réciproques, en agissant de façon pro-sociale, et en adaptant leurs propres comportements à ceux de leurs partenaires. Malgré cela, une théorie influente suggère que les interactions coopératives des animaux non-humains, notamment les grands singes, sont de nature égoïste dans la mesure où les participants ne possèdent pas les capacités cognitives nécessaires pour partager leurs intentions et s'investir conjointement dans un objectif commun. En d'autres mots, les animaux ne sembleraient pas capables de s'engager dans une action commune par le biais d'un sentiment d'*intentionnalité partagée*. Cependant, cette théorie pose plusieurs problèmes. D'une part, elle présuppose la la possession d'un certain nombre de capacités cognitives très sophistiquées (mais pas forcément nécessaires à la réalisation d'une entreprise commune). D'autre part, cette théorie s'appuie sur les résultats d'études comparatives menées en laboratoire avec des chimpanzés captifs et des enfants humains. Ces études manquent ainsi de pertinence écologique et d'équité inter-espèce au niveau des protocoles expérimentaux. De plus, elles ont été réalisées principalement avec des chimpanzés, connus pour être l'espèce de grands singes la plus compétitive. Des études récentes ont cependant apportés de nouvelles connaissances sur la coordination d'actions jointes chez les grands singes qui, associées aux raisons mentionnées ci-dessus, nous poussent à remettent en question cette théorie prédominante.

Ma thèse avait deux objectifs principaux – le premier de nature théorique et le second de nature empirique. L'objectif théorique était de re-conceptualiser la théorie sur l'intentionnalité partagée. Dans ce but, j'ai développé un nouveau cadre théorique pour l'étude de la coordination d'actions jointes, inspiré par la recherche sur les interactions humaines, et qui permet d'étudier, chez toutes les espèces animales, comment deux participants co-créent, co-maintiennent et co-rompent un état *d'investissement commun* au cours des différentes phases d'une action jointe (entrée, principale et sortie). Lors du déroulement d'actions jointes, les participants sont constamment concernés par la nécessité de *préserver la « face »* – une forme de manifestation publique d'estime de soi. J'ai suggéré que les comportements et les signaux de communication déployés par les participants pour coordonner les actions jointes représentaient de bons indicateurs pour déterminer la présence d'une certaine responsabilité envers un investissement commun

d'une part, et d'une certaine sensibilité envers la nécessité de préserver « face » de chacun, d'autre part.

Mon second objectif était d'appliquer de manière empirique ce cadre théorique aux interactions sociales des grands singes (chimpanzés et bonobos), plus précisément aux interactions de toilettage et de jeu social.

Les résultats de ma thèse indiquent que les bonobos et les chimpanzés présentent plusieurs des comportements révélateurs d'une responsabilité envers un investissement commun, mais aussi d'une certaine forme de gestion de la « face». Ces résultats suggèrent ainsi que ces deux espèces de grands singes sont capables d'atteindre un état d'intentionnalité partagée au travers du processus de coordination d'actions jointes. Par exemple, elles présentent des phases d'entrée et de sortie identifiables (i.e., les partenaires produisent des efforts de communication pour débuter, ou conclure, une interaction) lorsqu'ils se toilettent ou jouent, et s'efforcent de réinstaurer l'activité à la suite d'une interruption. De plus, mes résultats montrent que le type de communication employé par les bonobos et les chimpanzés pour coordonner les différentes phases d'une action jointe varie en fonction des liens sociaux qu'ils entretiennent avec leur partenaire, c'est-à-dire leur liens d'affinité et leur différence hiérarchique. Ceci fait écho aux modèles théoriques de gestion de la « face» chez les humains.

Pris ensemble, les résultats de ma thèse posent la question du rôle de la compréhension d'un investissement commun et de la gestion de la face dans l'évolution des primates. J'espère que mes conclusions renforceront l'idée que les caractères uniques à l'Homme ont probablement été façonnés à partir de racines ancestrales dans l'échelle phylogénétique des primates, et qu'elles relanceront le débat sur l'exclusivité humaine dans la capacité à créer un sentiment d'intentionnalité partagée.

Mots clés • Action conjointe, communication, intentionnalité partagée, théorie de la politesse, grands singes

Acknowledgements

I don't even know where to start, when thinking about whom to thank or how to summarize this PhD experience. First of all, I wish to thank my supervisors, Emilie, Klaus, and Adrian, which have made this experience the most inspiring, fun, and perhaps biggest learning curve a person could ever be able to make! Honestly, I have learnt **so** much, that I can now stop learning anything for … let's say at least 30 years I guess. Or, alternatively, I could remove the previous 28 years of my life, then it would equal itself out. But then, these previous years don't count because I haven't been that smart yet… I don't know!!! It is hard to describe. But really, I thank you all *so much*, for being you, inspiring, curious, and excellent! Emilie, I wish to thank you especially, because you have always been there, and you have guided me closely and have been a friend at the same time. Additionally, I would like to thank Federico, who is as much of a member of the team, and has been a really great advisor for me during my time in San Diego. I've gained much support, and I had really inspiring discussions with you; it has been a pleasure to work with you on this project!

A large amount of gratitude especially go to Katia Iglesias, who supported me both with statistical advice but most importantly also with mental support; I always felt safe to ask anything, supported and given freedom to create my own knowledge, while at the same time gaining the advice I needed to learn statistics – thank you *so* much! Special thanks also go to Christof Neumann and Martin Götz, who have inspired my work with Bayesian statistics, thank you both for your input and discussions! Niel, thank you for your support with Python scripts and having been there in San Diego – I appreciated it.

I also *especially* wish to thank all the people who have been with me these three years of my PhD time, both in the field, at Uni, and in my personal life. A really big thank you goes to the people at La Vallée, who I met and who made me feel home during my stays in France. Not only have I received huge support for my research at site, but also have I felt home and with friends, thank you so much Franck, Carole, Lolita, Stephen, Raph, Thomas, JP, Bastien, Magaly, and everyone else from La Valléee. A very big thank you also goes to the team at Lola Ya Bonobo, have I have worked with in DRC, including Raphaël, Fanny, Claudine, as well as Alain Miteti Muyamba, Patrick Kinkani Kinkani, Ivonne Vela Ntona, C'arrive Nsimba Kiasi, Niclette Bonyoka Mboyo, Micheline Nzozi Katiayi – and *most* of all Héritier Izansone for being my friend and the best collaborator one could have! I'm also thankful to the staff at San Diego Zoo, who have made my stay in San Diego most efficient and great – thank you Michael, Kim, and Janet! Much thanks

also goes to my colleagues at the IPTO, who have – in the last three months I guess, or always (?) – seen me in a crazy PhD state! Garance, thank you for being so wonderful and full of empathy, you've given me much support during the last few month. Julie, thank you for having been my office-mate-in-crime; Alessandro (my current office-mate-in-crime), Cyril, Daniela, and *everyone* in the team, thank you; and really – Franziska! You are an idol in a way! Thank you for your powerful presence, you have given me much support throughout, bringing me to "blossom" and feel emancipated and empowered with my strengths.

And then… There are of course my loveliest friends and my family. A huge thanks goes to my mum, my dad, my lovely sister, Sabine, and my love Ludo… I could have never possibly been the same without you guys. You have made me feel *home*, as loved as a human can possibly feel, and enthusiastic about the future ways to come. Ludo we have been going through this all together, and now we're both Drs – weird right, because in a way we're just the same, no magic comes with Dr-Hood I guess :-) I cannot wait for our future ways to come. My close friends Babsi, Milli, Sophia, Simone, Jasmin, *Alice*, Anne, *Moe*, *Adrienne*, Paulina, Feli, Judith, who I cannot even describe. I don't have the words to say how much I thank you all for your support, love, friendship, and the most hilarious moments, laughing, and just being us, thank you for understanding me in these past few years and always. You've been so wonderful, that I could have not been the same right now if you hadn't shared this experience with me. Alice, living with you was the most family-like, and funniest time for me. Thank you for your love, kindness, funniness, craziness, adventure, and individualism. Moe! The best times, and always felt so good to be around you. We're so similar, that it almost sometimes feels like I see myself; you understand me in different ways, and I admire you for your kindness, intelligence, generousness and most of all for your "folie"! Adrienne, how could I have produced any smart thoughts in this time without your creativity, and beauty ?– I miss you already! Eithne, Thibaud, and Marlen, you have become my friends and my favorite colleagues/collaborators. I'm proud to share this wonderful interest with you and I'm impressed by you.

Thanks to all of you for being you!

Table of content

Chapter 1. General introduction

A thesis outlook

Many social animal species interact cooperatively (Boesch, 2002; Pitman & Durban, 2012; Vail, Manica, & Bshary, 2013). For example, the chimpanzees of the Taï forest in Côte d'Ivoire cooperate in groups to hunt for small mammals, and exhibit elaborate levels of collaboration with complementary hunting roles (Boesch, 2005). Yet, scholars have claimed that human cooperation has reached an unmatched level of complexity (Call, 2009; Levinson, 2006), insofar as only humans supposedly have the abilities and motivations to engage in joint action by sharing psychological states - a cognitive ability variously termed as *shared intentionality* (Bratman, 1993; Tomasello & Carpenter, 2007), collective intentionality (Searle, 1990), or "we" intentionality (Tuomela, 2005, 2007). Joint actions involving shared intentionality feature participants' coordination of individual actions towards a shared goal (Clark, 1996), with all participants commonly knowing that they are jointly committed (Clark, 2006; Gilbert, 2014).

However, there are several problems with the current definition of what it means to share intentions. The classical philosophical account of shared intentionality presumes that joint actions feature a suite of high-level cognitive capacities always present in joint action coordination (Tomasello, Carpenter, Call, Behne, & Moll, 2005), but new accounts argue that in fact such high-level capacities like sharing intentions and understanding of joint commitments may actually be achieved via lower-level alignment processes and perception instead (e.g., see Knoblich, Butterfill, & Sebanz, 2011; Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012; Vesper, Butterfill, Knoblich, & Sebanz, 2010). Tollefsen and Dale (2012) particularly emphasize the need for a more empirical account of joint action - one that considers not only the higher-level cognitive processes, but also the lower-level cognitive processes characterizing joint agency (p. 390): "It is only by integrating both kinds of account that we will be able to adequately explain the capacity to engage in joint activities".

Moreover, the previous theory of shared intentionality is largely supported by findings of laboratory experiments in which apes were tested in species-irrelevant settings with unfair cross-species designs (Warneken, Chen, & Tomasello, 2006), and in which experimental confounds have not been sufficiently removed (see Leavens, Bard, & Hopkins, 2017). The findings also mainly rely on chimpanzees, who are supposedly less socially tolerant than their closest cousins, the bonobos (Hare, Melis, Woods, Hastings, & Wrangham, 2007). Lastly, the theory's attempt to

reconstruct the evolutionary scenario in which shared intentionality would have evolved is unwarranted, insofar that other, previously unrecognized events (i.e., collective predator defense) likely triggered an advent of shared intentionality far earlier than previously assumed at 400kya (van Schaik & Burkart, 2018).

In this thesis, I revisit the shared intentionality hypothesis in both theoretical and empirical ways. The theoretical objective is to re-conceptualize the philosophical account of shared intentionality, by looking at it from an ethological perspective, i.e., by acknowledging that higher-level cognitive states such as shared intentionality can be achieved through lower-level processes in joint action coordination, such as action alignment, communication and gaze exchanges, and perception (Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012). I do so by implementing a joint action framework (Genty et al., submitted; Heesen, Genty, Rossano, Zuberbühler, & Bangerter, 2017a), drawn from evidence of previous joint action research in humans (Clark, 1996), by which we may study the joint action coordination of our closest relatives, the chimpanzees and bonobos, in natural settings with conspecifics. In this framework, shared intentionality becomes visible in the process of joint action coordination, by which participants collaborate to establish, maintain, and dissolve *joint commitments*, leading to a visible sequential joint action phases with an *entry*, a *main body*, and an *exit* phase. The framework also takes into account the socio-affiliational imperatives intrinsic to joint action coordination with social partners of differing social relationships and status (Enfield, 2008; Stivers, Mondada, & Steensig, 2011), and allows for the study of potential biological precursors of politeness as face management (Brown, 2015; Brown & Levinson, 1987) in apes.

My empirical objective is then to apply the joint action framework, in order to assess whether we may find evidence for the ability of how apes to create a higher-level state of shared intentionality via lower-level processes in joint action. I will do so, by first studying how our closest relatives, the bonobos and chimpanzees, get into and out of social interactions, and whether the coordinative efforts constituting these phases reflect partners' social relationships. Second, I study how apes cope with interruptions of their interactions, by looking at how they suspend and reinstate social activities, as well as how they may adjust communicative efforts to partners' social relationships and personal responsibilities. Bonobos and chimpanzees provide pivotal models for the reconstruction of when and how shared intentionality may have evolved, as they are our closest living relatives (Prüfer et al., 2012; Wildman, Uddin, Liu, Grossman, & Goodman, 2003), and differ in social organization and ecology (see Gruber & Clay, 2016 for a review). My findings

will shed new light on the joint action capacities in apes, and on how natural selection may have shaped the evolution of shared intentionality from abilities minimally required for joint action, notably joint commitment understanding and face management.

In this general introduction, I will start by familiarizing the reader with the classical definition of shared intentionality (i.e., *Tomasello's definition*, 2016), as it has been portrayed in the previous literature, and by reviewing the relevant empirical data from respective experiments. Next, I will point out possible problems of the definition, and introduce the 'joint action framework', to address whether some animal behavior may qualify as evidence for their ability to co-create a state of shared intentionality. While introducing the joint action framework, I introduce relevant definitions from human joint action research, notably joint action phases, and politeness as face management, and the state of the art regarding joint action coordination in apes. Finally, I will state my empirical goals.

The classical definition of shared intentionality

A central feature of human sociality is the understanding of being part of a joint action (Bratman, 1992). Knowing that one is doing something with someone else, and sharing the same goal with that person, is crucial for the successful coordination of joint action. As Searle noted (1990, p. 401), shared (or collective) intentions are expressed in the form "we are doing such-and-such". Human joint action thus is special in the animal kingdom, as it features individuals' ability to share their psychological states– or in what is variously defined as to engage in "shared intentionality" (Tomasello & Carpenter, 2007), "collective intentionality" (Searle, 1990), or interacting in the "we-mode" (Gallotti & Frith, 2013; Tuomela, 2007). The classical definition of shared intentionality hence relies on the presence of higher-level cognitive abilities endowing us to *share psychological states* (Bratman, 1992).

According to Tomasello, the cognitive capacities related to shared intentionality would have evolved in two phases within human evolution, and supposedly distinguishes joint action in humans from interactions of other animal species (Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012). The first transition supposedly occurred about 400kya, when early humans shifted from an ape-like individual intentionality (*me* wanting you to do things to achieve my personal goal) to a "second-personal morality" or joint intentionality at the dyadic level (*we* acting together in a joint action towards a shared goal) (Tomasello, 2016, p. 86). According to Tomasello (2016, p. 51, 76-77), joint intentionality already counts as shared intentionality, yet does not govern the cooperative hunting in chimpanzees (and so in any other animal species), because these

activities do not involve the agents' understandings that they are pursuing a shared goal. This is because, so he argues, during cooperative hunting the chimpanzees supposedly do not, prior to the hunt, explicitly negotiate on roles and neither do they perform complementary roles, or share the spoils without trying to monopolize the prey (but see Boesch, 2005). As Tomasello puts it (2016, p.51):"[Chimpanzees, while hunting,] do not have a joint goal of capturing the monkey; rather, each has the personal goal of capturing the monkey for itself (as evidenced by their attempts to monopolize the carcass afterward." Joint intentionality remains therefore a purely human characteristic, that presumably had been triggered by the need for obligate collaborative foraging, making it a strategic solution at first, and allowing individuals to increase their own survival benefit via selective partner choice (Tomasello, 2016; Tomasello et al., 2012). The second transition presumably occurred about 150kya when human group size and competition arose, driving individuals to coordinate with strangers via a more "objective morality" (i.e., *us* knowing how to interact due to our cultural common ground) (Tomasello, 2016, p. 86) – leading to a so-called collective intentionality and modern human morality.

Tomasello and colleagues (Tomasello & Carpenter, 2007; Tomasello et al., 2005; Tomasello & Moll, 2010) list several socio-cognitive abilities that characterize shared intentionality, and which are believed to be human-unique (but see evidence below suggesting that apes may also express some of these abilities). These involve joint attention, perspective-taking, complementarity of roles (understanding that roles are interchangeable, and viewing the interaction from a birds-eye perspective), cooperative communication, provision of help, cultural (instructed) learning, and joint commitment. In these next paragraphs, I will review the literature and the current evidence on which of these cognitive skills are present in apes and humans; I will go through each of the skills in sequential order to update the reader on what – according to Tomasello - applies as human-unique and what does not.

Joint attention

Although apes are aware about what others see (Tomasello & Call, 2006), follow gaze (Call, Hare, & Tomasello, 1998; Tomasello, Call, & Hare, 1998), and recognize when their own gaze is exploited by others (Hall et al., 2017; Hare, Call, Agnetta, & Tomasello, 2000), they do not know that they and their partner see something *together* (Tomasello, 1995). As opposed to children, apes presumably lack the ability to engage in *joint attention* (Tomasello, 1995) - describing a triadic engagement by which partners coordinate their attention towards a target of shared interest, while both partners know it together (Carpenter & Call, 2013). Establishing joint

attention often involves declarative communication. Human infants, by around the age of 1 year, communicate *cooperatively*, with the interest to inform others about things, often to declare interest in a certain thing (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006).

Cooperative communication

Cooperative communication – as another cognitive correlate of shared intentionality - seemingly constitutes a further difference between humans and apes. Although apes can point to request things from experimenters (Leavens, Hopkins, & Bard, 1996), to date there is only weak evidence that they point to inform another conspecific, let alone to declare something (but see Hobaiter, Leavens, & Byrne, 2014). To illustrate the difference, when an experimenter helpfully points to indicate a hidden object at a certain location, one-year old infants, but not chimpanzees, perceive the pointing gesture as helpful cue to locate a target (Behne, Carpenter, & Tomasello, 2005; Itakura, Agnetta, Hare, & Tomasello, 1999; but see Hopkins, Russell, McIntyre, & Leavens, 2013). Yet, recent evidence challenges the idea that apes do not communicate cooperatively by informing others. Vocalizations in chimpanzees, for instance, can be used to inform the recipient about a relevant event (Crockford, Wittig, Mundry, & Zuberbühler, 2012), and like in humans, some gestures of great apes can also have iconic (Genty & Zuberbühler, 2014, 2015), or deictic content (Hobaiter et al., 2014; Leavens et al., 2015), although still such cases reported remain relatively scarce.

Furthermore, it is worth mentioning that research from the past 60 years has revealed a detailed array of findings concerning the communicative behavior of bonobos and chimpanzees, and brought about insights regarding similarities and possible differences to human communication (Byrne et al., 2017; Liebal, Waller, Slocombe, & Burrows, 2013). Like other non-human great apes, bonobos and chimpanzees communicate multi-modally via different sensory channels (Arbib, Liebal, & Pika, 2008), and also by combining different signal components, such as gestures (Graham, Hobaiter, Ounsley, Furuichi, & Byrne, 2018), facial expressions (Demuru, Ferrari, & Palagi, 2015; Parr, Waller, Vick, & Bard, 2007), and vocalizations (Clay & Zuberbühler, 2009; Crockford, Gruber, & Zuberbühler, 2018). Most signaling displays, including some facial expressions, like the play face, vocalizations and gestures, can be voluntarily controlled and adapted to the audience (Cartmill & Byrne, 2007; Flack, Jeannotte, & de Waal, 2004; Slocombe & Zuberbühler, 2007; Waller, Caeiro, & Davila-Ross, 2015).

Great apes have goals and intentions, and they do recognize these in other individuals (Buttelmann, Schütte, Carpenter, Call, & Tomasello, 2012; Call, Hare, Carpenter, & Tomasello, 2004; Warneken & Tomasello, 2006; Yamamoto, Humle, & Tanaka, 2012). Gestures, and to some also extent vocalizations (Crockford, Wittig, & Zuberbühler, 2015; Schel, Townsend, Machanda, Zuberbühler, & Slocombe, 2013), are used intentionally to pursue social goals. This involves response waiting (Call & Tomasello, 2007), flexible adjustments to the recipient's attentive state (Genty, Neumann, & Zuberbühler, 2015a; Leavens, Hostetter, Wesley, & Hopkins, 2004; Liebal, Call, Tomasello, & Pika, 2004), persistence in the case of goal-failure (Cartmill & Byrne, 2007), and directedness to a specific audience (Genty, Clay, Hobaiter, & Zuberbühler, 2014; Genty & Zuberbühler, 2014).

Vocal development seems to be largely genetically hard-wired (Arbib et al., 2008; Seyfarth & Cheney, 1997), with only little evidence for flexibility of learning (but see for instance, Crockford, Herbinger, Vigilant, & Boesch, 2004). By contrast, gestures have more potential regarding developmental acquisition and social learning (Halina, Rossano, & Tomasello, 2013; Pika, Liebal, Call, & Tomasello, 2005; Tomasello et al., 1997). Yet still, since a large part of great ape gesture types overlap in both meaning (Graham et al., 2018) and form (Byrne et al., 2017), scholars assume ape gesture repertoire to involve a high degree of innateness. For example, bonobos and chimpanzees overlap in about 90% of their gestural repertoire (Graham, Furuichi, & Byrne, 2017). Nonetheless, even those gestures presumed to be innate undergo a form of "repertoire tuning" process (Hobaiter & Byrne, 2011a), by which signalers come to learn the functional specificities of signals over time (Genty, 2019), probably through repetitive interactions with various partners (Pika & Fröhlich, 2019). Recently, researchers started to emphasize the importance of studying different signal modalities in combination (Fröhlich, Sievers, Townsend, Gruber, & van Schaik, 2019; Liebal et al., 2013), integrating all possible sensory channels, such as audible, tactile, and visual channels

Complementarity of roles and perspective-taking

Through the engagement in joint attention and shared activities, children develop so-called "perspectival cognitive representations" (Tomasello & Moll, 2010, p. 344). By internalizing one's own and the other person's intentional behaviors, children come to perceive the interaction as a result of both the perspective of the self and the other (Barresi & Moore, 1996). This leads children to acquire a birds-eye view of the interaction, with the understanding that their own roles are complementary and reversible with those of the partner – such a birds-eye understanding is

supposedly not present in chimpanzees (Fletcher, Warneken, & Tomasello, 2012). This form of perspective taking also later, around the age of four, paves the way for the child's understanding that a partner's perspective can be different from reality ("false belief" understanding, Wimmer & Perner, 1983). However, although false belief understanding was long thought to be uniquely human (Call & Tomasello, 2008), recent findings provide preliminary evidence for this ability in apes (Buttelmann, Buttelmann, Carpenter, Call, & Tomasello, 2017; Krupenye, Kano, Hirata, Call, & Tomasello, 2016).

Helping

Helping a partner to reach a common goal is a crucial aspect of shared intentionality, and it has been shown for both humans and apes. For example, chimpanzees help conspecifics to access food or other items (Melis & Tomasello, 2013; Melis et al., 2011), and offer spontaneous help to humans and conspecifics, both with or without the request for help (Greenberg, Hamann, Warneken, & Tomasello, 2010; Warneken & Tomasello, 2006; Yamamoto, Humle, & Tanaka, 2009), and without reward (Warneken, Hare, Melis, Hanus, & Tomasello, 2007). More so, bonobos extent such spontaneous helping behaviors to unfamiliar conspecifics (Tan, Ariely, & Hare, 2017; Warneken et al., 2007), showing that, at least in bonobos, pro-social behaviors and sympathetic concern are not necessarily restricted to kin (Clay & de Waal, 2013; see also Tan & Hare, 2013). Human children offer help for partners in social tasks, even without the partners' request for help (Warneken et al., 2007). They do so altruistically, even when it is costly for themselves (Warneken et al., 2007; Warneken & Tomasello, 2008), and show signs of physiological arousal when others are being helped (Hepach, Vaish, & Tomasello, 2012).

Cultural (instructed) learning

Cultural learning constitutes yet another cognitive correlate of shared intentionality that supposedly distinguishes humans from apes. Chimpanzees can socially learn from each other, but they do so exploitatively and "product-oriented", mainly focusing on success and outcome (Call, Carpenter, & Tomasello, 2005; Whiten, Custance, Gomez, Teixidor, & Bard, 1996). Children, by contrast, do so in "process-oriented" fashion by reliably imitating instructor's actions (Tomasello, Kruger, & Ratner, 1993). Children also give feedback to the instructor and recognize when someone is acting in a pedagogical way (Gergely & Csibra, 2006). This endows humans with a special form of cultural learning, which allows for cumulative culture - the process by which previous cultural knowledge is transmitted into the next generation by a high level of fidelity, and then modified and improved, leading to complex technological advances (Dean, Vale, Laland,

Flynn, & Kendal, 2014; Tennie, Call, & Tomasello, 2009). While humans learn by imitating, then, apes presumably learn by observing social partners yet focusing on the environmental result, rather than the actor's action itself (Hopper, Lambeth, Schapiro, & Whiten, 2008).

Collaboration via joint commitment

Apes are able to cooperate in collaborative tasks requiring joint efforts (Melis, Hare, & Tomasello, 2006a; Melis & Tomasello, 2013; Rosati, Stevens, Hare, & Hauser, 2007). Chimpanzees tactically select partners for cooperative tasks who have previously proven to be good collaborators (Melis et al., 2006a) –a skill suggesting active collaboration with intentional selection of unrelated individuals, usually common in humans. Nonetheless, social tolerance promotes cooperative tendencies (Melis, Hare, & Tomasello, 2006b), and since it is more common in bonobos than in chimpanzees (Hare et al., 2007), it leads to differing levels of success in both species during cooperative tasks (Hare et al., 2007; Melis et al., 2006b). In chimpanzees, a lack of social tolerance has been shown to constrain cooperation, insofar that individuals who are generally less tolerant tend to perform more poorly in cooperative tasks with social partners (Melis et al., 2006b). Elevated tolerance levels in bonobos facilitate bonobos' cooperative abilities, and allow them to outperform chimpanzees in cooperative settings (Hare et al., 2007; but see Cronin, De Groot, & Stevens, 2015). While bonobos generally perform better in the social domain (perspective-taking and social causalities), chimpanzees show greater skills in the physical domain (tool use and physical causalities) (Herrmann, Hare, Call, & Tomasello, 2010).

Although great apes engage in complex collaborative group activities, Tomasello argues that each apes act according to personal intentions, rather than shared intentions with joint goals and plans, with participants' knowing that they are jointly commitment (Bratman, 1992). Joint commitment, to date, has been assumed to be a unique aspect of *human* shared intentionality (Tomasello & Carpenter, 2007; Tomasello & Moll, 2010), and "essential to all true joint activities" (Clark, 2006, p. 126). Bratman (1992, p. 328) states that human joint action should involve *mutual responsiveness, commitment to the joint activity,* and *commitment to mutual support.* Mutual responsiveness involves partner's attention and mutual feedback towards intentions and actions, knowing that the other seeks to do the same. Joint commitment to the joint activity means that partners both share an "intention in favor of the joint activity" (Bratman, 1992, p. 329). Commitment to mutual support involves partners' motivation to help each other to fulfill their roles, ensuring joint success. Children do this by helping or encouraging a social partner who is unable to continue the collaborative activity in their own role (Warneken, Gräfenhain, &

Tomasello, 2012). Moreover, the children stay committed to a collaborative interaction until the end, even remaining to help the partner after they had already received their own reward (Gräfenhain, Carpenter, & Tomasello, 2013; Greenberg et al., 2010; Hamann, Warneken, & Tomasello, 2012). They also stay longer if they had formed an explicit joint commitment (verbally), and acknowledge their leave-taking if they have to break the commitment (Gräfenhain, Behne, Carpenter, & Tomasello, 2009).

Evidence suggests that joint commitment in humans develops gradually throughout ontogeny. Its basis is set by toddlers first developing interest to engage in coordinated games more generally (Eckerman & Whitehead, 1999; Ross & Lollis, 1987). It then builds up to the child's recognition of obligations linked to explicitly formulated joint commitment at the age of three years (Gräfenhain et al., 2009, 2013; Kachel, Svetlova, & Tomasello, 2017). And finally, around the age of five years, it reaches a more profound recognition even without explicit formulation - by the mere understanding of interdependency (Kachel & Tomasello, 2019). Warneken et al. (2006) launched an experimental paradigm to compare the joint commitment understanding of apes and human children. They had 18-24-months old toddlers and apes play cooperative social games with a human experimenter, and then the experimenter suddenly dropping from the game without further acknowledgement. The toddlers (though more so the 24-months old ones) readily reengaged the experimenter by performing game-relevant actions, pointing, and vocalizing. Yet, the chimpanzees *never* reengaged, and instead most often attempted to resume the previous game individually.

Altogether, these lines of evidence resulting from laboratory experiments led researchers to assume that the interactions of apes are shaped by competition, rather than cooperation, and that apes do not have share intentionality (Call, 2009; Tomasello, 2014). The issue is that most results from natural observations are ignored (Boesch, 2005), and that consecutive studies today reveal controversial results. They showed that, when minoring the degree of the games' complexity, and when testing more socially tolerant species such as bonobos (Hare et al., 2007), apes can engage in shared activities with social partners (MacLean & Hare, 2013; Pika & Zuberbühler, 2008). In the next section, I will expand on this point, and discuss the problems that may lead to such controversial findings.

Problems with the classical definition of shared intentionality

The evidence addressing the question of whether apes have shared intentionality according to Tomasello's definition (Tomasello, 2016) is controversial. Tomasello's top-down approach to

studying shared intentionality implies that individuals are only able to shared intentions and goals if they possess the necessary cognitive abilities (listed in section above). As outlined above, apes have some of the abilities necessary for shared intentionality, but yet they lack others, which leads scholars to believe that they must lack shared intentionality (Tomasello, 2016). In what follows, I will attempt to explain why it still has been so difficult to show that apes have shared intentionality, and stress that shared intentionality should be re-conceptualized and studied differently.

First, the cognitive concepts constructing the term *shared intentionality* are highly affected by theoretical presuppositions of mentalization and thus difficult to test empirically in any kind of species (Leavens et al., 2017). This may easily lead to inaccurate assumptions about higher cognitive process involved in joint action, although simpler alignment processes may actually be responsible for the behavioral patterns found (Knoblich & Sebanz, 2008; Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012). As Rosas and Bermúdez have made the point for joint attention (2018), I argue that certain cognitive abilities like joint commitment understanding can also be non-conceptual representations instead, insofar that participants may not actually have any concept of the cognitive state by which their joint commitment is governed. In that way, joint commitment, like joint attention (Rosas & Bermúdez, 2018), may not *need to be conscious* and to rely on cognitive processes, but perhaps on motor synchrony or perceptual processes instead. As Bermúdez and Cahen phrase it in the Stanford Encyclopedia of Philosophy (2015): "[…]Some mental states can represent the world even though the bearer of those mental states need not possess the concepts required to specify their content". I argue that this may help to explain, then, why we may find evidence for sophisticated forms of joint action in agents that presumably lack shared intentionality, notably young children (Gräfenhain et al., 2009; Kachel & Tomasello, 2019) and apes (Fröhlich, Kuchenbuch, et al., 2016; MacLean & Hare, 2013; Pika & Zuberbühler, 2008; Rossano, 2013). The current top-down definition of shared intentionality presumes high-level cognitive processes in each and every social encounter – which is virtually impossible and probably not necessary in everyday fast-paced human interactions (Vesper et al., 2010). The current theory of shared intentionality mainly rests on philosophical grounds that over-intellectualize the conditions needed for joint action coordination – e.g., it concentrates on the planning of shared goals and joint commitment prior to interacting (Gilbert, 2014). However, in reality participants often engage in joint action spontaneously, and come to the common perception that they are sharing a joint goal throughout their joint engagement, through lower-level processes of alignment, including mutual gaze exchanges (Rosas & Bermúdez, 2018), communicative turn

taking (Levinson, 2016), motor synchronization (Reddish, Fischer, & Bulbulia, 2013; Sebanz, Bekkering, & Knoblich, 2006), and perception (Knoblich et al., 2011). Therefore, the theory of shared intentionality must become naturalized, and extended towards an empirical, process-based account of joint action (Tollefsen & Dale, 2012). As Tollefsen and Dale put it (2012, p. 393): "[the philosophical accounts lack] detail regarding how actions during joint activity unfold (execution problem), what cognitive or other processes are employed in the activity (implementation problem), and what starts the activity (multiple initiation problem). In this light, surface alignment processes can serve as a kick-start for the emergence of higher-level cognitive states such as shared intentions and goals".

Second, recent empirical findings challenge the reconstruction of the evolution of shared intentionality provided by Tomasello (2012; 2016). The challenge particularly concerns the assumption that shared intentionality has evolved 400kya, first starting at the dyadic level (as consequence of obligate collaborative foraging) and then transitioning to the group level (as consequence of competition between groups and increase of group size). Van Schaik and Burkart (2018) argue for another evolutionary event that most likely preceded obligate collaborative foraging: collective predator defense. This vital activity must have involved *a strong group commitment*, most likely featuring shared intentionality (Boehm, 2018). As this event was dated to about 3 Mya (Ferraro et al., 2013), shared intentionality has most likely emerged far earlier than previously assumed at 400kya, in combination with both group and dyadic components (Boehm, 2018; van Schaik & Burkart, 2018).

Third, the way shared intentionality has previously been assessed may be problematic due to inadequacy in experimental testing procedures, unfair species comparison, a main focus on only one species that is less competitive than others (the chimpanzee), and ecological irrelevance (Kingstone, Smilek, & Eastwood, 2008; Leavens et al., 2017). Leavens and colleagues (2017) pointed out at least *nine* studies with experimental confounds, designed for testing socio-cognitive traits presumably governed by higher-level cognitive states, e.g., shared intentionality. These confounds corresponded to differences in the group's rearing environments, prior task preparation, sampling protocols, testing procedures and age (Leavens et al., 2017, p. 489). Moreover, it is questionable whether the previous experiments used to test shared intentionality were relevant regarding the apes' ecology, as they involve testing conditions with human partners and often complex, species-atypical social games (see Warneken et al., 2006). Indeed, when conditions are kept more natural and involve less complicated social games, apes can engage in shared activities

and presumably recognize joint commitment with a human partner (MacLean & Hare, 2013; Pika & Zuberbühler, 2008). Moreover, observations of natural interactions in apes reveal the existence of certain joint action capacities in apes – for example the ability to create participation frameworks (Rossano, 2013) and engagement in communicative turn-taking (Fröhlich, Kuchenbuch, et al., 2016), with signals flexibly adjusted to interpersonal relationships (Fröhlich, Wittig, & Pika, 2016a).

Fourth, cooperative hunting in chimpanzees (Boesch, 2002) and in lions (Stander, 1992) actually ***do*** fulfill the criteria that Tomasello had set for the ability to share intentions and pursuing shared goals; hunters take individual roles, which they can flexibly adapt and adjust depending on the progress of the hunt and other individuals' movements. Additionally, chimpanzees divide the spoils after the hunt, according to age, rank and hunting responsibilities (Boesch, 1994, 2005). It is really puzzling why these activities are not considered as joint action with shared goals (Tomasello, 2016). Again, part of the reason for why it is impossible to "proof" shared intentions in animals is because such higher-level cognitive abilities are unfalsifiable when testing behaviors (Leavens et al., 2017), which is why such a *top-down approach is just not helpful* for finding valid species differences regarding shared intentionality.

Fifth, the current definition of shared intentionality fails to account for the need to manage interpersonal relationships, so for the relevance of face work in human social encounters (Goffman, 1955). Face is defined as a form of publicly manifest self-esteem (Goffman, 1967), and it concerns peoples' right to be approved by others, as respected and deserving person (positive face want), as well as the right to be autonomous and unrestricted in choices (negative face want). Face management, as I will detail in the upcoming sections below, is a universal phenomenon in humans, and may have an ethological basis, by which it may be studied in other animal species beyond humans. As face management is constantly apparent in joint action (Arundale, 1999; Brown, 2015; Brown & Levinson, 1987), it intrinsically belongs to joint actions in humans, and should therefore be taken into the equation when one is interested to unveil human uniqueness in ethological ways. Tomasello's definition of shared intentionality does not take this phenomenon into account, which further warrants for a revision of the theory.

A last important point is that shared intentionality has thus far mainly been assessed in chimpanzees, despite the fact that they are considered to have lower levels of social tolerance than our second closest relative, the bonobo (Hare et al., 2007). Bonobos have been shown to perform better in tasks related to the social world than chimpanzees (Herrmann, Hare, et al., 2010), and

exhibit a high level of pro-socially, evidenced by their heightened levels of social tolerance, motivation to help other's (even unrelated individuals) (Tan et al., 2017), and their sympathetic concern for distressed individuals (Clay & de Waal, 2013). It may well be that bonobos are better candidates for studying joint action coordination capacities linked to shared intentionality, than are chimpanzees, and that the results to be drawn from bonobos' findings may challenge the current hypothesis that only humans possess shared intentionality. The differing dominance styles of the two species and their different ecology (see for details Chapter 2) may help to generate predictions about the kinds of conditions that might have promoted the evolution of joint action capacities.

To summarize, shared intentionality has previously been investigated as the *ability* to share psychological states (Tomasello et al., 2005), which has made it difficult to study this ability empirically, let alone in ecologically relevant settings. This has made species comparisons between apes and humans particularly difficult, and – since experiments addressing the theory have focused on testing apes with human-like games (Warneken et al., 2006) - given possible advantages to humans. Here, I intend to stress that shared intentionality in apes may be demonstrated by taking a different approach: a joint action approach, which visualizes shared intentionality in the process of joint action. It is a bottom-up approach, actually, where joint action is coordinated via communicative signal and gaze exchanges, and then leading to the establishment of shared intentionality as an "interactional achievement". In what follows, I will discuss how such a joint action approach may be implemented.

Theoretical goal: Implementing a joint action framework to study shared intentionality as interactional achievement

Shared intentionality, in its previous top-down definition, represents an umbrella term comprising a suite of different high-level cognitive correlates (Tomasello, 2016), and thus far only humans have been shown to exhibit all of those. I argue however, that by drawing on a more process-based, joint action approach (Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012), shared intentionality can be studied differently. Here, I introduce a reverse approach – a bottom-up approach - to study the necessary prerequisites of joint action which allow for the achievement of shared intentionality as a collaborative state that emerges through simpler processes such as gaze and signal exchanges when participants interact jointly. Achieving a state of togetherness in interaction involves the establishment, maintenance and dissolution of joint commitment, as well as the management of face – a publicly manifest self-esteem, which I will introduce in detail below (Goffman, 1955). Joint commitment and face management may be considered as the minimal

ingredients to achieve a higher state of shared intentionality in joint action, and are publicly visible in the joint actions of participants, in the form of certain behavioral and communicational patterns, which I will also come to detail below. In other words, I here focus on how participants come to achieve a state of shared intentionality as the result of the coordinative process of joint action. I argue that the joint action framework, by which participants coordinate on establishing, maintaining and dissolving a joint commitment to interact, allows for the re-assessment of shared intentionality in our closest relatives in ways that avoid assumptions about higher-level cognitive processes. In what follows, I will first introduce the reader to relevant concepts of joint action, joint action phases, and face management in humans. Next, I will review the current state of the art regarding the potential abilities of apes to coordinate joint action. I will then present the joint action framework, for studying shared intentionality as the process by which participants coordinate joint action, and state the empirical goals of my thesis.

Joint action

Joint action (Clark, 1996) implies that two or more participants coordinate their actions with each other toward a shared goal. Participants coordinate their actions step-by-step and in concert with one another, involving both mutual responsiveness and joint commitment (Clark, 2006). Joint actions take many forms, some of which are small-scaled and occur daily, such as having dinner or even having a conversation (Clark, 1996), and some of which are large-scale projects enduring in time, such as a collaborative research project. The concept of joint action is widely applied by various disciplines, including linguistics (Sperber & Wilson, 1986), philosophy (Bratman, 1992; Grice, 1975; Tuomela, 2007) pragmatics (Sacks, Schegloff, & Jefferson, 1974; Schegloff & Sacks, 1973; Sidnell & Stivers, 2012), and psychology (Bangerter & Clark, 2003; Bangerter, Clark, & Katz, 2004). The results provide a detailed understanding of how agents coordinate joint action, both via higher-level and lower-level cognitive processes.

Low-level cognitive processes in joint action represent processes by which humans come to coordinate at the behavioral and motoric level (Vesper et al., 2016), and these processes emerge with or without a necessary "deeper" cognitive contract between the agents (Garrod & Pickering, 2009; Tollefsen & Dale, 2012; Vesper et al., 2010). Non-cognitive motor and alignment processes lead to automatic, emergent coordination, and involve gaze exchanges, turn-taking, entrainment (synchronization), perception-action matching and action simulation (Knoblich et al., 2011). A gaze exchange between two people standing in close proximity may serve as a kick-start for interaction, and can be a cue that leads to the exchange of linguistic utterances and consequently

the establishment of an interaction and joint commitment toward a shared goal (Tollefsen & Dale, 2012). Turn-taking - i.e., an exchange of communicative turns with an average modal response gap in humans of 0.2s (Levinson, 2016) – creates a platform for a rapid exchange of ideas and feedback, allowing for fast-paced interactions and efficient coordination. Entrainment is a self-organizing mechanism leading to an automatic synchronization of peoples' behavior, as in the example of an audience's applause turning into synchronized clapping (Neda, Ravasz, Brechet, Vicsek, & Barabasi, 2000). Perception-action matching triggers the matching of observed actions of others into one's own action repertoire, thereby often leading to action mimicry (Chartrand & Bargh, 1999). Action simulation permits motor predictions in the observer by the perception of another agent's actions (Kourtis, Sebanz, & Knoblich, 2010).

High-level cognitive processes in joint actions comprise "deeper" cognitive processes, by which people recognize others as mental agents with personal intentions, thoughts and beliefs, distinguishing thereby their own self from other's, and acknowledging other's believes can deviate from reality (Frith & Frith, 2005; Tomasello & Rakoczy, 2003). The ability of attributing mental states endows humans to imagine other agents' perspectives of an interactive scenario (Gallotti & Frith, 2013), leading to a shared understanding that the agents are doing something together (Gallotti & Frith, 2013; Tuomela, 2007). Shared intentionality, in its current account, represents such a high-level cognitive ability, allowing for the meshing of intentional states during joint action, and giving rise to the agents' mutual understanding that they are pursuing a common goal (Tomasello et al., 2005).

Much research has been devoted to how humans naturally engage in spontaneous joint activities through micro-analytical methods (Sidnell & Stivers, 2012), which has led to a close analysis providing detailed insights into the sequential structure of joint action. In what follows, I will review the state of the art on what is known about this sequential structure, laying out the behavioral and communicative techniques of would-be participants to perform in a joint action.

Joint action phases

Joint actions are orderly processes, in which participants collaborate in a step-by-step fashion to achieve their shared goal (Clark, 1996). These processes pose generic coordination problems, which participants have to solve in order to achieve successful coordination. Specifically, initiators of an interaction need to select would-be participants, establish a joint focus of attention, make their intentions comprehensible, and establish a joint commitment to interact (Genty et al., submitted; Heesen et al., 2017). The participants need to make each other understand

what the interaction content will be about, who is to take what role, and where and when the interaction takes place. Once the participants have jointly committed to interact and decided on the details of the interaction, joint actions involve restrictions on participants' autonomy. Because the participants committed to invest time and effort to reach the shared goal, they give up freedom to pursue potential other activities. The recruitment of would-be participants and committing to a joint action therefore involve face threats (Brown & Levinson, 1987; Goffman, 1955), which participants constantly attempt to mitigate. If the integrity of the interaction is threatened, for example by an interruption of a third party (Bangerter, Chevalley, & Derouwaux, 2010), participants need to agree on suspending it, and then need to re-establish the joint commitment once they are ready. Finally, participants need to come to the mutual conviction that they are ready to end the joint action, and to coordinate on disengaging.

These coordination demands lead to three observable macro-level phases in joint action – the opening, main body, and closing (Clark, 1996), which entail publicly visible behavioral and communicative components. Each of these macro-level phases comprise several sub-steps, which participants need to coordinate on, either via speech (e.g., on the telephone, Clark & French, 1981), or via multi-modal communication involving tactile, visual, and audible channels (e.g., Modaff, 2003; L. Mondada, 2006). According to Goffman, a stretch of interaction comprises (1981b, p. 130): *"[…]all that relevantly goes on from the moment two (or more) individuals open such dealings between themselves and continuing until they finally close this activity out. The opening will typically be marked by the participants turning from their several disjointed orientations, moving together and bodily addressing one another; the closing by their departing in some physical way from the prior immediacy of copresence. Typically, ritual brackets will also be found, such as greetings and farewells, these establishing and terminating open, official, joint engagement, that is, ratified participation."*

In interaction openings, participants first have to build up participation frameworks (Goffman, 1981b; Kendon, 1990; Lorenza Mondada, 2009), by first selecting and ratifying participants, with whom they intend to interact and to ensure that their prospective partners are available to engage in a yet unspecified interaction. Participation frameworks entail a state of shared focus of attention among participants (Goffman, 1981a; Kendon, 1976, 1990), and hence demarcate who is (and who is not) a ratified participant (Goffman, 1981a). Participants then have to negotiate on what their joint commitment entails, including the nature of the commitment, its timing, content, and location (Clark, 1996, 2006). Openings are thus comprised of a *pre-entry*

(ratification of participants and establishment of joint focus of attention) and *entry* phase (jointly committing to content, location etc.).

In the main body, participants carry out the intended business of their encounter. They update each other on the progress and provide mutual communicative feedback to transition from one task to the next (Bangerter & Clark, 2003). If the joint action is interrupted, for example by a third party, participants have to coordinate on suspending the joint commitment, addressing the interruption, and then reinstating the joint commitment once they are ready (Chevalley & Bangerter, 2010). To collaboratively accomplish suspension, the participants have to negotiate and agree that they are ready to suspend the interaction. Initiating a suspense may involve the use of politeness expressions, such as requesting for permission for suspension (Chevalley & Bangerter, 2010). To resume the interaction, participants need to first check their partner's availability, and signal their readiness to resume the interaction. Next, participants may use politeness behaviors, such as apologizing for letting partners wait (e.g., "I'm sorry to keep you waiting"), justifications as to why the interaction had to be suspended (e.g., "I had a call waiting"), and asking partners for help in reconstructing the previous topic (Bangerter et al., 2010; Chevalley & Bangerter, 2010).

In interaction closings, participants accomplish several sub-steps to progressively coordinate the termination of their encounter and, at the same time, manage relationship continuity beyond the encounter (Albert & Kessler, 1976). Participants first need to signal their readiness to end the encounter, in order to come to the mutual conviction that the encounter can end. Participants may exchange turns such as "Okay?" - "Alright" (Schegloff & Sacks, 1973, p. 307), ensuring that so far unaddressed topics can still be raised if necessary. They then further progress through the closing via several steps, such as summarizing the encounter, displaying the wish for relationship continuity, and justifying why the encounter has to end (Albert & Kessler, 1978; Knapp, Hart, Friedrich, & Shulman, 1973). After the participants have signaled to each other that they are ready to end the interaction, they have to take leave of each other, for example by exchanging well-wishing expressions such as *good-byes*, hanging up the phone, or walking away (Broth & Mondada, 2013; Clark & French, 1981). Leave-taking is a consequence of having come to the mutual conviction to be ready to end the encounter. Communicating in exit phases also serves to manage interpersonal relationships beyond the encounter, and reflects the degree of affiliation between participants, or "reaffirms acquaintances" (Clark & French, 1981, p. 1). In endings of telephone calls, people are likely to exchange goodbyes if they had exchanged information, or had exchanged personal information. This means, the more service / personal

information participants exchange, the more they "owe" to each other in the end, and "good byes" or "thank you" is one way to depict such a state (Clark & French, 1981). The closing is thus comprised of a *pre-exit* (coming to the mutual conviction to be ready for terminating the encounter) and *exit* phase (taking leave).

As previously noted, joint actions involve joint commitments, and these commitments rely on the individuals' personal contributions, time and efforts for the successful achievement of the joint goal. The commitments thus entail obligations (Clark, 2006) that restrict the participants' momentary freedom to pursue other activities. Breaking these obligations therefore carries face threats. In what follows, I will detail what these face threats entail and explain the concept of politeness as face management.

Politeness as face management

Imagine that a person, who went to dinner with friends, suddenly decides to leave without informing the others. The friends would soon start wondering where the person would be, most likely worry, and - when finding out that she left without letting anyone know - have all reasons to be puzzled or mad. Failing to accommodate the obligations linked to joint commitments is therefore face threatening, and can have negative consequences for social relationships.

Goffman's notion of "face" is "(...) the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact" (Goffman, 1967, p. 5). It can be described as a publicly manifest self-esteem, and rests on the universal assumption by all members of a society that all members have face (see also Brown, 2015). Levinson and Brown (1987), in their seminal series of politeness theory, stated two kinds of face wants (resulting from the public identity or self-image, which every person claims for her/himself). The desire for one's own self-image or identity to be approved of and appreciated by others (defined as positive face), and the desire to keep personal space, one's own territory, freedom of choice, and the want not to be imposed upon by others (defined as negative face) (Brown & Levinson, 1987). When one person threatens another person's face wants, for instance by disapproving their possessions as valuable (threatening positive face), or by asking them to do something unreasonably weighty (threatening negative face), both risk to lose face (Brown & Levinson, 1978; Goffman, 1967; Holtgraves, 2002). Loosing face means feeling embarrassed or humiliated, which is thus a public emotion (Goffman, 1967). As face is fragile and emotionally invested, people must *constantly* manage each other's face (i.e., by being aware and adhering to protect the above mentioned face wants) (Brown & Levinson, 1987).

With face always being potentially at stake is meant that *any act* that is being made by an individual towards a person is *always* face-threatening (yet varying in degree of threat), and therefore requires mutual cooperation by participants (Brown, 2015). A face-threatening act (FTA) then is one that threatens a person's positive or negative face, by operating in opposition towards the face requirements of the other person. To mitigate FTAs, people use politeness (Brown & Levinson, 1978). Politeness represents an interpersonal ritual, or "a concept designating 'proper' social conduct, rules for speech and behavior stemming generally from high-status individuals or groups" (Brown, 2015, p. 326). Politeness is predominantly a linguistic phenomenon, expressed via speech acts in face-to-face settings (Brown, 2015), but may also be accompanied by non-verbal acts (Stadler, 2006). It is thus far considered to distinguish humans from other animal species, because it is in itself a feature of language, which represents the cooperative nature of humans by their primary ways of communicating (Brown, 2015).

Politeness applies as a form of ***face management***, by which individuals regulate each other's face (Brown, 2015, p. 327). Levinson and Brown (1987) recognized that politeness also constitutes as a universal phenomenon, that appears across cultures, and implies that participants orient language acts to their partners' face, or as Brown puts it (2015, p. 326) their "public persona". When individuals ask others to do things or recruit partners for interaction (i.e., when they make an imposition on the partner in any way), their politeness strategies can vary depending on face threats involved; generally, politeness use increases with the weightiness of an FTA – greater impositions require operations with more politeness involved. Politeness can thus be taken as behavioral evidence for a person's ability to take into account another person's face (i.e., their public self-image, feelings and the way a person should be treated interactionally, with respect and right for autonomy) (Brown, 2015).

This involves behaving in ways that reflect recognition of the social relationship that one may carry with a certain partner, as well as a partner's relative social status to oneself (Brown & Levinson, 1987). Face threats are thus dependent on different social factors. They increase with participant's social distance (as symmetric relation) and relative power difference (as asymmetric relation), and the culture-dependent ranking of an imposition as face-threatening (Brown & Levinson, 1987). So, individuals are polite by taking into account these persona aspects of others; they are more polite with higher-status individuals, and more polite to unfamiliar individuals. For example, FTAs increase and more politeness is required if a lower status individual (e.g., student) recruits a higher status individual (e.g., professor), as opposed to when a higher status individual

recruits a lower status individuals. Additionally, FTAs increase and more politeness is required when an individual recruits a socially distant individual (who they never met before), as opposed to when an individual recruits a familiar individual (someone living in the same village). As face is universally managed, people have universal expectations on how people should be treated, depending on their social status and social affiliation; it is therefore that people have rights to sanction others when they disrespect face requirements (Limberg, 2009).

There are five different *politeness strategies* in how speakers may minimize the risk of losing face (fig 1). If the risk of losing face is large, for example because an imposition that is being made is unreasonably weighty (e.g., asking your professor to write your thesis), then it would be best not to make the FTA at all. If the FTA has to be done, and the risk of losing face is less than in the example before, then there are two choices. The choice less face threatening is to go "off record", which means that one may say things indirectly (e.g., a person wanting someone to close the window may "it is cold in here" instead of demanding imperatively). The more face threatening choice would be to go "on record", and this may be done by either doing the FTA with or without redressive action. Doing the FTA without redressive action, baldly, is mostly done in situations where the risk to lose face is lowest. This can be in situations where superiors demand things in urgency from subordinates, in a setting where face work may not be the highest priority, for instance a surgeon asking the assistant to "hand the scalpel". Doing the FTA with redressive action allows for two options of politeness use. "Positive, or solidarity politeness" (Brown, 2015, p. 327) orients to the positive face of a person, and implies that one says things in a way that approve the partner's personality, thereby promoting feelings of closeness and affiliation (e.g., "What a beautiful vase this is! Where did it come from?", Brown & Levinson, 1987, p. 103). "Negative, or respect politeness" (Brown, 2015, p. 327) is used when the estimated risk of face loss is greater than for positive politeness, and orients to the negative face of a person. It implies avoidance-based acts, by which the speaker expresses intentions to not or minimally interfere with a hearer's autonomy (e.g., "May I borrow your car, please?", Brown & Levinson, 1987, p. 143).

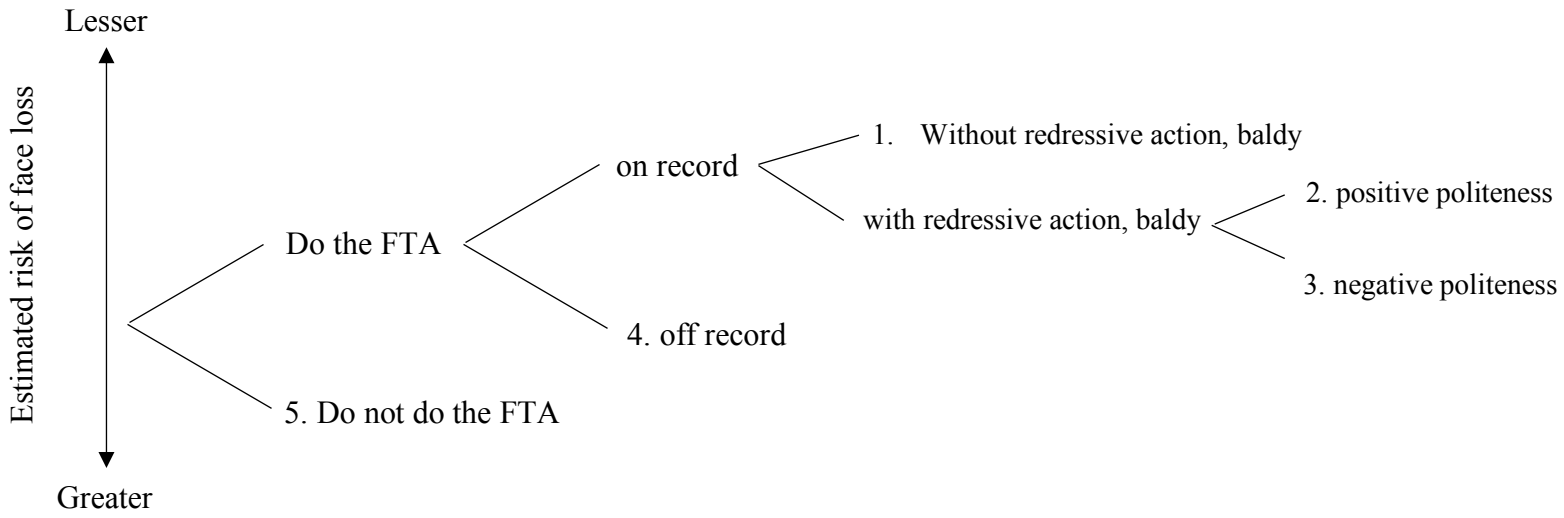


Figure 1 Circumstance of estimated risks of face loss that will determine the choice of politeness strategy that signalers will deploy. Figure taken from Levinson and Brown (1987, p.60). If the estimated risk of face loss is minimal, signalers may do the FTA on record and without redressive action (baldly). If the estimated risk of face loss is greatest, they should at best not do the FTA. The numbers (1-5) indicate politeness strategies deployed in circumstances where risks of face loss are lowest (1) to highest (5).

Face management in joint action

So far, I have introduced the concepts of joint action, joint action phases and politeness as face management. Before proceeding, I attempt to create a connection between these concepts, which I hope will provide a groundwork for the upcoming joint action framework, as well as prepare the reader for upcoming predictions in the empirical chapters.

Some scholars have made the important notion that engaging in joint action actually not only affords people's efforts to track shared knowledge, but also their efforts to manage the socio-relational dimensions inherent to joint action (Enfield, 2006; Heritage & Atkinson, 1984; Stivers et al., 2011). Enfield (Enfield, 2006, p. 412) emphasized that, alongside an informational imperative (i.e., tracking of common ground), social interaction also features a socio-affiliational imperative, which requires that participants engage with each other by an "appropriate level of intensity or intimacy", considering their current social relationship. Politeness is a pivotal means by which participants of an interaction may manage such socio-affiliational imperatives. As face is fragile and always at risk, joint action coordination constantly requires participants to deploy appropriate politeness strategies. Yet, politeness may be most evident at particular moments of the joint action, where the estimated risks of face loss are particularly *salient*.

This may be the case especially in phases of joint action, where joint commitments and socio-affiliative imperatives are most fragile. For example, as explained above in detail (section joint action phases), in the entry phase of a joint action, joint commitments are to be negotiated and brought into existence. The risk of losing face may be heightened, as there is room for refusal or denial, or for not appropriately addressing a prospective partner according to his/her social status and the social relationship of the two participants. In the suspension phase of a joint action, after an interruption has occurred, participants may need to ask for permission to suspend the joint action, as well as to put the commitment "on hold", and this may threaten the integrity of their interaction and quality of relationship, and incur risks for face loss. In the reengagement phase, after participants have dealt with the interruption, the risk of face loss is high as participants need to deploy adequate politeness strategies, which on the one hand allow to re-establish the broken joint commitment, and on the other to treat the partner with necessary respect. People may use apologies or justification as to why the partner had to wait. Additionally, face is at risk because there is a chance of refusal to reinstate the activity by the other partner. In the exit phase, where joint commitments are to be dissolved, face risks may especially increase, because participants

have to come to the mutual agreement that they are ready to end their encounter, as well as manage their relationship in ways that allows for positive future contacts. The exit phase may involve risks such as disagreement between the participants, or of deploying politeness strategies which inadequately reflect a current social relationship, by which future contacts may become embarrassing.

To conclude, politeness as face management allows participants to create, maintain and dissolve joint commitments in joint action *without threatening face*. Again, as politeness theory is a universal phenomenon that occurs across cultures, and because it is so intrinsic to the joint action coordination of humans, it may have an ethological basis, which warrants the need for exploring it in non-human animal species, notably apes.

Predictions for an ethological study of face management

I here provide a schema on how one may draw predictions on face management in joint action phases of humans (fig 2), which may then be useful to draw predictions of what we should find in apes to show evidence for precursors of face management. Joint action phases comprise participants' coordination efforts to establish, maintain (thus re-establish after interruptions), and dissolve joint commitments. When recruiting someone to interact and to do something (i.e., establishing a joint commitment), people need to deploy different kinds of politeness styles; for instance, one may be to approve the partner's personality (i.e., social distance and power status), and the other may be to minimize the weight of the imposition made (i.e., larger vs. small) (Brown, 2015). Face management, as the use of politeness to manage face, may be more or less evident, depending on the weightiness of an imposition, and on the social relationship / status difference between two participants of a joint action. With more or less evident I mean that participants deploy more or less politeness efforts, respectively, in their communicative acts. Following the schema (fig 2), one would expect that face management is less evident in interactions relying on joint commitments that imply less time/effort and between partners who are more familiar / similar in social status. On the contrary, face management may be more evident in interactions relying on joint commitments that imply much effort/time and between partners who are less familiar / different in social status.

Therefore, evidence for precursors of face management in the interactions of apes would be provided by identification of the following behavioral patterns:

- When initiating, suspending, reinstating or terminating an interaction, apes communicate less (e.g., use signals less frequently, in decreased amplitude, uni-modally) when interacting with partners who they have a stronger bond with, and with partners who are subordinate (lower-ranked) relative to themselves;
- When initiating, suspending, reinstating or terminating an interaction, apes communicate more (e.g., use signals more frequently, in greater amplitude, multi-modally combined), when interacting with partners who they have a weaker bond with, and with partners who are dominant (higher-ranked) relative to themselves.

If apes then produce communication when initiating, suspending, reinstating, or terminating interactions with social partners, we may be able to identify interaction phases similar as those in human joint action. We may call them entry phases, suspension phases, reengagement phases, and exit phases, and we may measure the phases' duration and likelihood of presence, as a measure of "effort" or "communicative costs" that signalers make when coordinating such phases with partners. Therefore, evidence for precursors of face management in interaction phases of apes would be provided by identification of the following behavioral patterns:

- The complexity (i.e., duration, presence) of the phases increases when initiated by a signaler who addresses a weakly bonded, or a dominant partner;
- The complexity (i.e., duration, presence) of the phases decreases when initiated by a signaler who addresses a strongly bonded, or a subordinate partner.

To sum up, face management in humans is universal, and may thus have a biological basis, thus warranting its study in apes. As humans are not the only species to collaborate in group activities with joint efforts and social partners of differing relationships (Boesch, 2005), apes may rely on similar strategies to manage both potential socio-affiliational imperatives and other coordination demands inherent to group activities (e.g., potential forms of joint commitment). It is therefore worth studying how apes may come to achieve coordinated activities with social partners, and how they, at the same time, manage social relationships. In other words, it is worth studying whether apes exhibit the patterns mirroring the aforementioned predictions drawn from politeness theory. Before coming to discuss how we may systematically implement a joint action framework that allows for such research in apes, I will first review the evidence suggestive for the ability of apes to engage in joint action via shared intentionality.

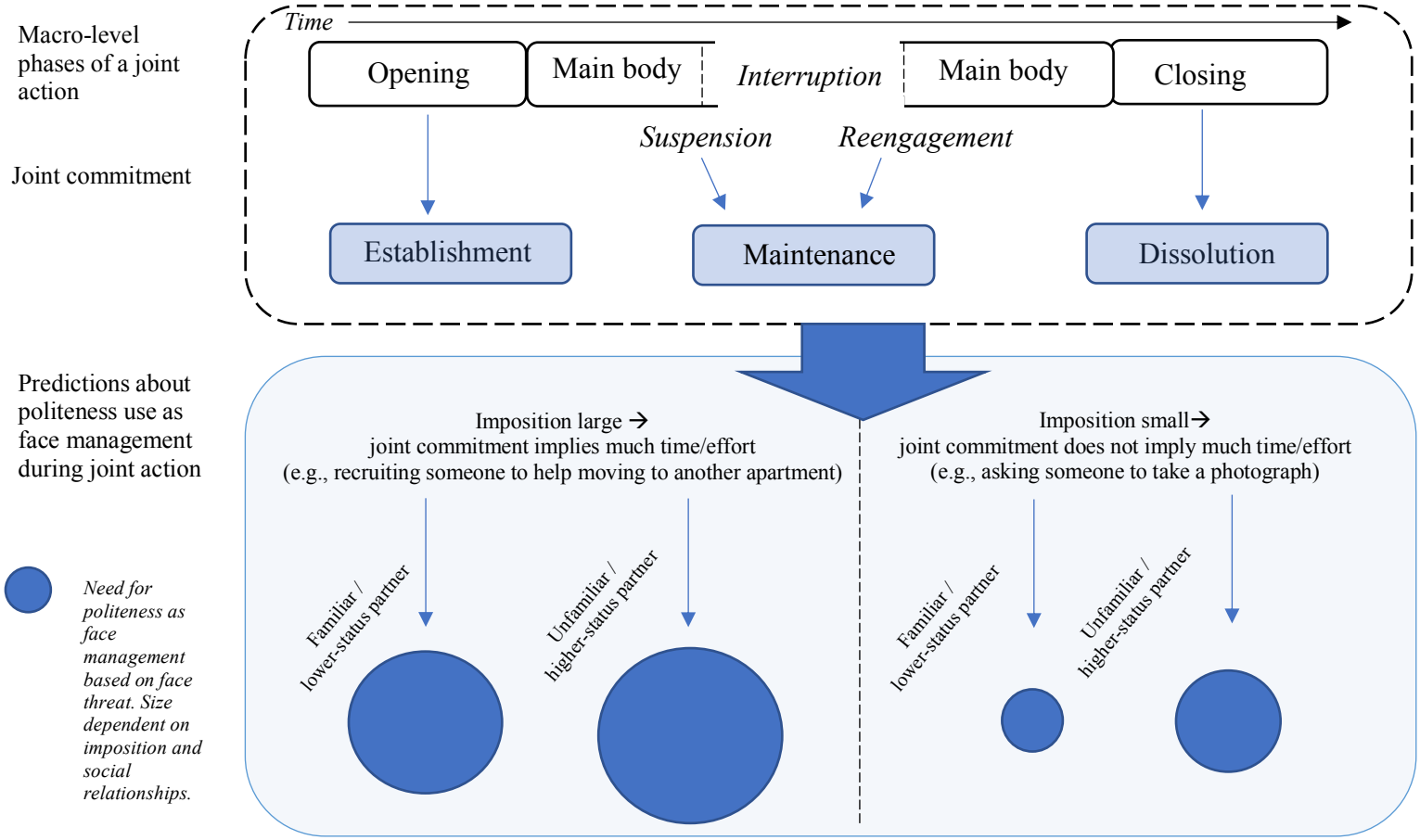


Figure 2 Graphical representation of the link between the concepts of joint action (Clark, 1996), joint commitment (Clark, 2006), and politeness theory (Levinson & Brown, 1987). Joint action phases comprise participants' coordination efforts to establish, maintain (i.e., re-establish after interruptions), and dissolve joint commitments. When making an imposition (i.e., establish a joint commitment), participants need to deploy two kinds of politeness styles: one is to approve the partner's personality (i.e., social distance or familiarity, and power status), and the other is to minimize the weight of the imposition made (i.e., larger vs. small) (see Brown, 2015).

Joint action in apes?

Apes regularly engage in joint activities with two or more partners that involve advanced skills of communication and intention reading, often toward an obvious shared outcome. For example, chimpanzees regularly engage in group hunting (Boesch, 2002) and boundary patrolling (Watts & Mitani, 2001), where individuals coordinate and adjust their own movements and precisely anticipate others' movements, and do not disengage from the activity until their common goal has been reached (e.g., to capture the prey, to control the boundaries of the home range). Moreover, by means of communication, apes routinely coordinate a variety of other joint activities, such as social grooming (Fedurek, Slocombe, Hartel, & Zuberbühler, 2015; Pika & Mitani, 2009), social play (Palagi, 2008), sex (Genty, Neumann, & Zuberbühler, 2015b), as well as coordinative problem-solving tasks in laboratory settings (Melis & Tomasello, 2019). In day-to-day activities like grooming, play, and mother-infant traveling, gestures are frequently exchanged via turn-taking sequences, giving rise to a sequential structure of interaction not unlike the one characterizing joint actions in humans (Fröhlich, Kuchenbuch, et al., 2016; Fröhlich, Wittig, & Pika, 2016b; Fröhlich, Wittig, et al., 2016a; Rossano, 2013; Rossano & Liebal, 2014). Individuals tailor their communication in relation to the attentional state of their recipients (Hostetter, Russell, Freeman, & Hopkins, 2006), as well as to the recipient's comprehension (Cartmill & Byrne, 2007; Crockford et al., 2012; Roberts, Vick, Roberts, & Menzel, 2014) and familiarity (Genty et al., 2015a).

Social play and grooming are good candidates to study joint action in apes, because they occur daily and frequently, and involve a mutual focus of attention as well as ample communication. For one individual to groom another, requires that the other one remains in a body position that ideally allows for grooming; for two individuals to play together, both have to move their bodies in a certain way that allows for playful action. Social play is especially cooperative, because it requires adjustment of strength and timing of individual acts, and anticipation of others' actions (Bekoff, 2004; Palagi, Antonacci, & Cordoni, 2007; Palagi, Cordoni, Demuru, & Bekoff, 2016). Play is also highly reciprocal, requires active engagement (Palagi, 2008; Palagi, 2006), and substantial communication (Demuru et al., 2015; Fröhlich, Wittig, et al., 2016a; Genty, Breuer, Hobaiter, & Byrne, 2009). Play involves risks of escalating into aggression due to intense body contact and often rough bodily movements. Participants mitigate such risks by using abundant communication to delineate from non-playful contexts; indeed, most of the gestural repertoires in apes contain gestures used for playful activities (Call & Tomasello, 2007; Hobaiter & Byrne,

2011b). Social grooming most often appears in lengthy, reciprocal bouts, in which partners frequently exchange grooming roles (Watts, 2016). Apart from a hygienic function, it serves many social functions (Dunbar, 1991), such as stress reduction (Crockford et al., 2013), tolerance enhancement (Port, Clough, & Kappeler, 2009), relationship establishment (Schino, di Sorrentino, & Tiddi, 2007), and social bonding (Fedurek & Dunbar, 2009).

There is also ample evidence that the sequential structure of these grooming and play activities resembles the one described for human joint action (see Genty et al., submitted). Apes coordinate on getting into, maintaining, and getting out of interactions via signal exchanges and multi-modal signal combinations. In the beginning, they orient their bodies in a way that demarcates who is to participate, and exchange gaze and gestures to determine roles and content – a form of participation framework (Fröhlich, Wittig, et al., 2016b, 2016a; Rossano, 2013). Social grooming can be initiated through specific postural, gestural or vocal signals (Fedurek et al., 2015; Goodall, 1986). Some gestures are even used to direct the groomer's attention to a body location (Pika & Mitani, 2009). Lip-smacking – a rapid opening and closing of the mouth common in chimpanzees–serves to coordinate grooming bouts, and increases especially when the risk of termination by the recipient is heightened (Fedurek et al., 2015). Generally, grooming styles are affected by partner's social bond and rank difference; for example, in chimpanzees, males groom more with other males than with females, groom more with closely bonded partners and with partners closer in rank to themselves and direct grooming up to higher ranking males (Arnold & Whiten, 2003). Grooming styles also reflect emotional preferences between unrelated partners; for example, in male bonobos, grooming is less reciprocated between partners who are more familiar than between less familiar partners (Surbeck & Hohmann, 2015).

To signal playful intentions, and to coordinate social play activities, apes deploy specific gestures, facial expressions and vocalizations. Play initiations seem to involve two steps. For example, apes often enlist play partners by using gestures to attract their attention (Tanner, 2004), such as the *drum object* gesture in chimpanzees (Hobaiter & Byrne, 2014). Once partners share attention, they may deploy specific gesture types that determine the type of play that follows; for example, in gorillas, *one-handed grab* gestures seem to be used for initiating contact play, *drum object* gestures seem to be used for initiating chase-play (Genty et al., 2009). Play solicitation gestures in chimpanzees are also flexibly adjusted to the partner's age, in that self-handicapping gestures are especially displayed for younger and same-aged partners (Fröhlich, Wittig, et al., 2016a). During the play-bout, apes frequently exhibit and calibrate play-faces and laughter to

facilitate progression of the bout (Demuru et al., 2015; Palagi, 2006; Palagi et al., 2007). For example, play bouts which involve full play faces (upper teeth exposed) are longer than those involving play faces (upper teeth covered), and full play faces are more frequently observed in intense than gentle play (Waller & Cherry, 2012). Gestures used within play may also serve to change the play type. For instance, while galloping in chimpanzees is used to decrease intensity and to switch from chase to contact play, hand shaking gestures typically increase the intensity and serve to switch from contact to chase play (Hobaiter & Byrne, 2014). Apes also use gestural and bodily displays, such as feet shaking and rolling over, respectively, to resume disrupted play bouts (Hobaiter & Byrne, 2014). To terminate play bouts, gorillas have been shown to frequently use hand-on gestures or body signals, such as pirouetting (Genty et al., 2009; Luef & Liebal, 2013).

Considering these lines of evidence, apes seem to engage in joint action by collaborating while getting into, maintaining, and getting out of interactions with peers. Yet, the evidence is not systematic, in that previous studies have not applied a controlled framework for studying whether these signals are serve to solve similar coordination problems as those in humans, nor have researchers systematically controlled for the social relationships between partners when studying the signals in the putative phases (but see Fröhlich, Wittig, et al., 2016a). A systematic joint action framework is needed, which would allow for the controlled study of joint action in humans and non-human animal species (notably apes), in order to study shared intentionality as interactional achievement.

The joint action framework

To overcome previous difficulties related to the top-down definition of shared intentionality, I here present an alternative approach to studying shared intentionality (see Genty et al., submitted; Heesen et al., 2017, Appendix S1). This approach relies on a more natural, process-based account of joint action coordination (Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012), builds on previous knowledge about human joint action (Clark, 1996), and acknowledges the possibility that shared intentionality may be achieved through the process of joint action coordination – so has an ethological basis. Its central questions are - *how do participants collaboratively create, maintain and dissolve a potential state of joint commitment through the process of joint action coordination, and do they manage socio-affiliational imperatives*? Looking at how individuals come to achieve coordinated joint action with partners would allow for a systematic, comparative assessment of shared intentionality in humans and non-human primates,

with more species-relevant criteria and an alternative view of shared intentionality as interactional achievement. The framework could be applied to the natural interactions of apes, and hence represents a more natural account on how apes may come to achieve shared intentionality; this would overcome issues related to uncomparable and artificial cross-species comparisons in the lab.

Shared intentionality, in this approach, is not looked at in the classical way - as the individual ability and motivation to share psychological states (Tomasello & Carpenter, 2007), by which it presumably relies on high-level cognitive abilities (Tomasello & Moll, 2010). Rather, I acknowledge the possibility that shared intentionality can exist as an ethological, observabl phenomenon, that can be achieved through the collaborative process of joint action coordination (Rosas & Bermúdez, 2018; Tollefsen & Dale, 2012). My approach assumes that joint commitment and face management, like joint attention (Rosas & Bermúdez, 2018), may *not necessarily* be represented mentally, but can be perceptional instead. This account looks at joint action in the same fashion as Arundale (1999, p. 125): "[…]as incrementally accomplished or achieved in action between persons, rather than as the meshing of two individuals' separate cognitive plans or schemas".

I argue that it is legitimate to study shared intentionality as an interactional achievement in the joint actions of young children and apes, where agents are concerned to come to the common perception that they are involved in a joint commitment over the course of their coordination efforts throughout the interaction. The agents have to work together in order to establish a common perception that they have a joint commitment, which they then attempt to maintain (and to re-establish if broken), and dissolve together once they are ready. In other words, the joint action framework represents a form of alignment process, which allows for studying the conditional behavioral preconditions by which two participants can reach shared intentionality (Tollefsen & Dale, 2012). If apes experience comparable coordination problems needed for the successful establishment of joint commitment as in humans, we should be able to observe comparable coordination efforts to solve such problems. These may be observed as publicly visible behaviors and communicative signal/gaze exchanges, in phases where joint commitment would naturally be established, maintained, suspended, reinstated, and dissolved (Genty et al., submitted; Heesen et al., 2017, see Appendix S1).

In fig. 3 I provide a detailed summary of the joint action framework, taken and adapted after both Genty et al. (submitted) and Heesen et al. (2017, see Appendix S1). The joint action

framework builds on previous empirical evidence from human interaction research (see section *joint action phases*), and may serve as a systematic assessment form that allows for the study of shared intentionality as interactional achievement in apes. Focusing on the coordination problems that emerge in joint action, it allows for a systematic comparison of whether apes may coordinate their actions with each other in naturally occurring joint activities, to solve coordination problems similar to those in human joint action. In humans, these phases imply both, the coordination demands to regulate joint commitments, as well as to mitigate face threats inherent to these commitments. Here, I focus on studying those joint action phases in apes, which in humans most prominently indicate joint commitment understanding– i.e., where joint commitment is brought into existence and publicly negotiated (**entry phases**), suspended and re-established (**suspension and reengagement phases**) and dissolved (**exit phases**), and where face management may be most salient.

To find evidence for the collaborative establishment of joint commitment, apes then would need to show evidence of some form of mutual perception, or recognition of their partner, in interaction beginnings. This may involve exchanges of gaze (look at each other's face at the same time), and exchanges of vocal or non-vocal communicative species-specific (and activity specific) communicative signals. These observable behavioral and communicative outputs may then serve as yardstick to assess the presence of an *entry* phase, and possibly suggest that the participants may jointly commit to interact (Clark, 1996, 2006; Lorenza Mondada, 2009).

To find evidence that apes are concerned with collaboratively suspending or reinstating a form of joint commitment, apes would need to again show some mutual recognition or attention toward another via mutual gaze or exchange of communicative signals at the time an interruption occurs, and when reinstating a social activity with the same partner. These behaviors would serve as yardstick to assess the presence of a *suspension* and *reengagement* phase, and possibly suggest that the participants may jointly agree to suspend their interaction, as well as jointly agree to resume the interaction, respectively. This would possibly hint at their ability to suspend and reinstate a potential form of joint commitment (Chevalley & Bangerter, 2010; Gräfenhain et al., 2009, 2013; Warneken et al., 2006).

To find evidence that apes collaboratively dissolve a form of joint commitment, apes would need to again show some mutual recognition or attention toward another via mutual gaze or exchange of communicative signals right before or very shortly after their interaction had ended. These behaviors would serve as yardstick to assess the presence of an *exit* phase, and possibly

suggest that the participants mutually accept that their encounter can end. This would possibly indicate their ability to dissolve a form of joint commitment (Albert & Kessler, 1978; Clark, 1996, 2006; Schegloff & Sacks, 1973).

As previously stated, politeness is a form of face management which serves to regulate the socio-affiliational imperatives inherent to joint action coordination (Arundale, 1999; Brown & Levinson, 1987; Enfield, 2008). It thus plays a key role when participants come to jointly interact, as they allow for the smooth coordination of joint action without threatening partners' relationships or reputation. As face management has a universal foundation in human joint action with a possible ethological basis, its study is warranted in apes. This can be done with the joint action framework, namely by studying whether the communicative signals used to coordinate interaction in apes (or the structure of the resulting phases) depend on social bond (as an inverse measure for social distance), and rank difference (as an inverse measure for power difference). If the patterns would follow the predictions made in section "*Predictions for an ethological study of face management studying",* we may be able to identify possible precursors to face management in our closest relatives.

I conclude, therefore, that shared intentionality may be studied as a collaborative achievement of joint action. The joint action framework serves as a yardstick to assess potential behaviors / communicative signals in apes, which may serve to solve similar coordination problems than in humans. These behaviors and communicative signals may then serve as evidence for joint commitment understanding and precursors of face management, in case they are flexibly exchanged, as well as placed at distinct moments of the bout, and if they depend on social relationships and status differences.

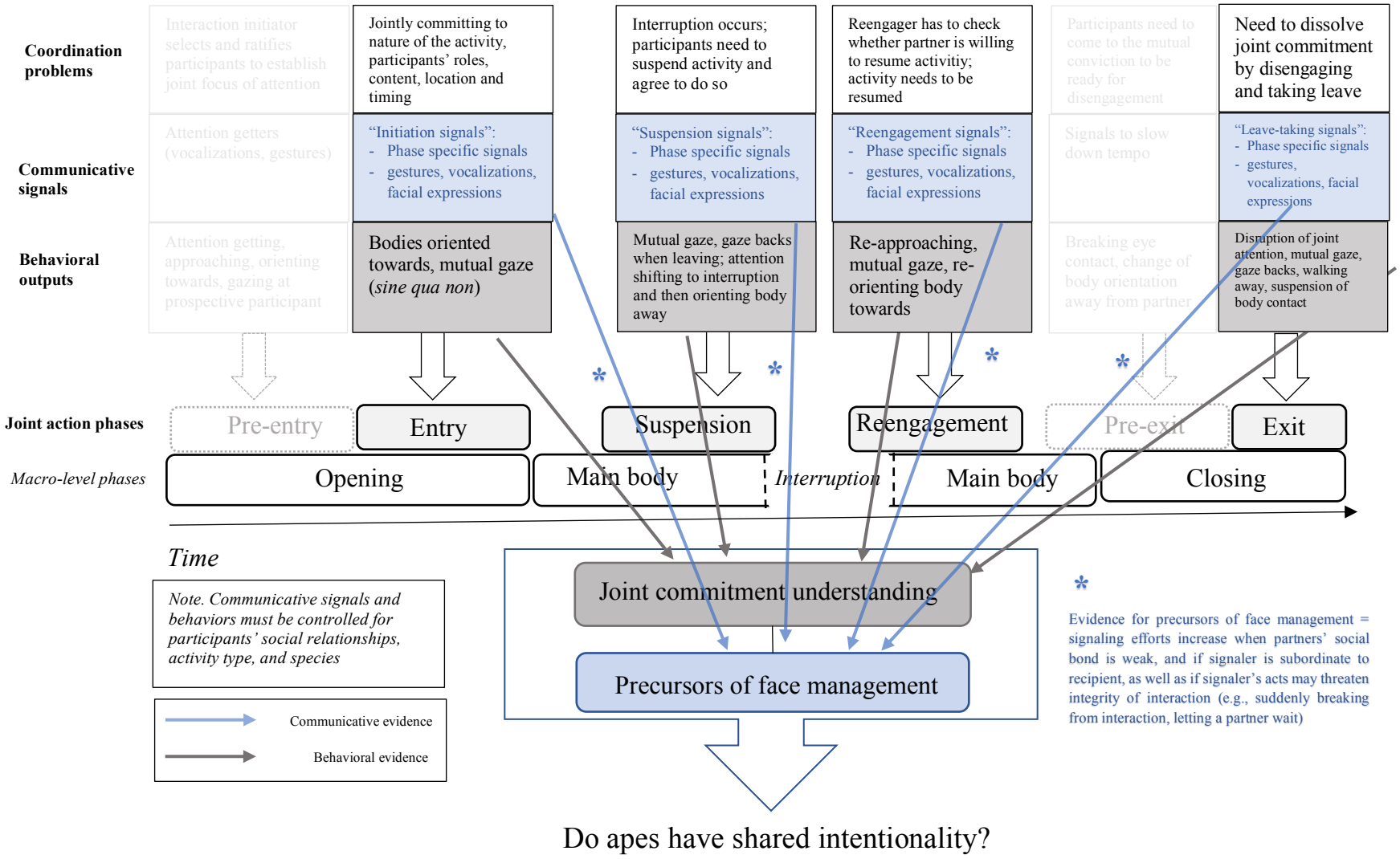


Figure 3 An alternative approach to studying shared intentionality as an interactional achievement in the natural interactions of apes. The figure was inspired by Emilie Genty. The framework implies that participants of an interaction face several distinct coordination problems when getting into, maintaining and getting out of joint actions. Participants' communicative and behavioral efforts to solve these coordination demands are publicly observable, and constitute suggestive evidence for precursors of face management and joint commitment understanding. The light grey, almost transparent phases indicate existing phases of the joint action, but are not studied in this thesis. The exchanges of behavioral and communicative outputs lead to visible "joint action phases", which characterizing an overall joint action sequence of opening, main body and closing. To show evidence for joint commitment in apes all phases marked in grey should be present.

Empirical goal: Applying the joint action framework

Empirical goal 1: How do apes initiate and terminate social interactions?

My first empirical goal (Chapter 3) is to study how bonobos and chimpanzees coordinate naturally occurring social interactions (notably social play and social grooming) with conspecifics. Specifically, I aim at investigating whether the two species exchange gaze and communicate at the beginning and in the end of their encounters, presuming that, if apes are able of creating a state of shared intentionality in joint action, they should exhibit entry and exit phases comparable to those in human joint action. Additionally, I aim at exploring whether the structure of these phases (i.e., duration and likelihood of phase presence), depend on social bond strength and rank difference between partners, presuming that, if apes possess precursors to face management, phases should be shorter and fewer when initiated with strongly bonded / subordinate partners. In addition to these aims, I will investigate potential differences between species and activity types, as well as compare the differences in the use of mutual gaze in interaction endings between species and activity types. This study addresses the larger question of whether apes may be capable of engaging in joint action, and thus whether they may collaboratively achieve a state of shared intentionality.

Empirical goal 2: How do apes suspend and reinstate interrupted social interactions?

My second empirical goal comprises three empirical studies, and concerns the study of how bonobos and chimpanzees generally cope with interruptions of ongoing social activities, i.e., how they suspend and reinstate interrupted interactions. Presuming that apes are capable of creating a state of shared intentionality in joint action, apes should exchange gaze and communicative signals to suspend and re-establish interrupted social activities with social partners. In the first study (Chapter 4), I study if apes of differing age classes reengage a reluctant human partner to an interrupted social game, a "tug-of-war game", which relies on a back-and-forth-pulling of a garden hose. In the second study (Chapter 5), I investigate whether this behavioral pattern extends to natural interactions with conspecifics, and whether there are species differences in the motivation to reinstate activities. In my last study (Chapter 6), I explore whether bonobos are, apart from reinstating activities, also motivated to communicate when suspending social grooming interactions with peers. More specifically, I look at whether bonobos' communication to suspend and reengage depends on the social bond strength and rank differences between partners, as well as the signalers' responsibilities with respect to the cause of interruption, and their own interaction role. I also compare whether bonobos' motivation to reinstate a previous activity

is heightened if the activity had been social (i.e., social grooming) compared to solitary (self-grooming, solitary play). These studies will address the larger question of whether apes may be capable of engaging in joint action (here by their ability to coordinate on suspension and reinstatement of joint activities), and whether they may collaboratively create a state of shared intentionality.

Chapter 2. The study species: bonobos and chimpanzees

In what follows, I will briefly introduce the general ecology, and social organization of my study species – the bonobo and the chimpanzee. The chapter serves to portray the ecological differences and similarities between the two species, to prepare for the following empirical chapters.

General ecology

The genus *Pan* comprises two genetically distinct great ape species: the bonobo (*Pan paniscus*) and the common chimpanzee (*Pan troglodytes*, with four genetically distinct subspecies, see Becquet, Patterson, Stone, Przeworski, & Reich, 2007; Morgan et al., 2011 for details). The lineage of the genus *Pan* diverged from the human lineage about five to eight Mya (Prüfer et al., 2012), and so both species share about 99% of their genetic material with us (Prüfer et al., 2012; Wildman et al., 2003). The lineage of two species themselves diverged around 1-3 Mya (Langergraber et al., 2012; Prüfer et al., 2012).

Bonobos and chimpanzees are diurnal and semi-terrestrial primates. Bonobos are endemic to the equatorial rainforests of the Democratic Republic of Congo. Their habitats are composed of rainforest, swamp and dry forest, disturbed forest, and to some extent savannah islands (Hashimoto et al., 1998). These forest habitats are high in fruit abundance, herbaceous food resources, and characterized by low seasonality (Kano, 1992; White & Wrangham, 1988). The habitat of the chimpanzee covers 21 countries, ranging from the eastern to western forest belts of Africa, with areas of primary and secondary moist lowland and (sub)montane forests, dry and swamp forests, and savannah environments (Humle, Maisels, Oates, Plumptre, & Williamson, 2016). Like bonobos, chimpanzees feed on fruits, herbaceous vegetation and protein-rich foods such as eggs and small animal prey. Both species are omnivorous and opportunistic feeders, engaging in occasional hunting activities to prey on other primate species and smaller mammals (Boesch, 1994; Fruth & Hohmann, 2018), yet chimpanzees do so more frequently than bonobos (Humle et al., 2016). The inter-birth interval of bonobos is estimated to 4.8 years, and that of chimpanzees to 5.2-6.6 years (Gruber & Clay, 2016). In both species, females reach sexual maturity at about 8-10 years, with females dispersing from their natal group at around 6-10 years in bonobos, and 11 years in chimpanzees (Gruber & Clay, 2016). The age spans can be up to 50 years, with a generation time of about 25 years (Langergraber et al., 2012).

Social organization

Bonobos and chimpanzees live in multi-male, multi-female fission-fusion societies, averaging communities of about 30-35 members, in extreme cases up to 150 members, with home ranges of around 20-60km^2 (Fruth et al., 2016; Humle et al., 2016; Mitani & Watts, 2005). Both species' social societies are characterized by female migration and male philopatry (Gruber & Clay, 2016; Kano, 1992). Yet, the two species greatly differ in their social structure.

Chimpanzees form strong alliances between males (Goodall, 1986), but relatively low affiliative contacts between females – although female sociability varies greatly between sites (Williams, Liu, & Pusey, 2002). Within the group, male chimpanzees use aggression to assert their dominance status and to compete for sexual contacts with females (Boesch, 2009). The males are dominant over females, and form strictly linear dominance hierarchies, with an alpha male and his alliances on top of the hierarchy (Boesch, 2009). The males form coalitions to defend their home range against members from other groups, by engaging in cooperative border patrolling with severe inter-group aggression and, sometimes, lethal inter-group killings (Watts, Muller, Amsler, Mbabazi, & Mitani, 2006).

In contrast to chimpanzees, bonobo females are highly gregarious and form strong social alliances especially with one another (Furuichi, 1989). Bonobo societies are characterized by high levels of promiscuity due to increased female control over mating partners (Furuichi, 2011). The cohesive grouping behavior and the strong affiliative bonds between females allow females to obtain high dominance ranks within mixed social hierarchies with males (Surbeck, Deschner, Schubert, Weltring, & Hohmann, 2012). Contrary to the despotic dominance hierarchies in chimpanzees, the dominance style of bonobos tends to be non-linear and more egalitarian (but see Jaeggi, Stevens, & Van Schaik, 2010; Stevens, Vervaecke, de Vries, & van Elsacker, 2007). In bonobos, the sexes form amicable relationships and the dominance status of females depends on contextual variables, such as feeding and sexual swelling periods (Surbeck et al., 2012). While in chimpanzees, males obtain their ranks by means of physical aggression, males in bonobos obtain dominance ranks through amicable bonds with females, and through their close mother-son relationships (Surbeck et al., 2017). Bonobos are also highly playful even in adulthood (Palagi & Paoli, 2007), and exhibit intense sexual behaviors also with same-sex partners that is decoupled from reproduction (Hohmann & Fruth, 2000). The highly frequent socio-sexual contacts in bonobos promote social tolerance and cooperation (Moscovice et al., 2019), reconcile former

opponents after conflicts, regulate competition and promote food sharing (see Gruber & Clay, 2016 for a review). Contrary to chimpanzees, bonobos generally have increased levels of social tolerance (Hare et al., 2007), pro-social behaviors (Tan et al., 2017) and food sharing, even with strangers (Tan & Hare, 2013).

Chapter 3. Joint action phases in apes: Entries and exits

Heesen Raphaela[1], Bangerter Adrian[1], Iglesias Katia[2], Neumann Christof [1,3], Rossano Federico[4], Zuberbühler Klaus[1,5], Pajot Aude[1], Laura Perrenoud[1], Guéry Jean-Pascal[6], Genty Emilie[1]

[1]University of Neuchâtel, [2]HES-SO University of Applied Sciences and Arts of Western Switzerland, [3] Laboratoire de sciences cognitives et de psycholinguistique, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, Paris, France, [4]University of California San Diego, [5]University of St Andrews, [6]La Vallée Des Singes, France

Abstract

Humans are particularly motivated to engage in joint action involving shared goals. Entering and exiting from joint action poses several coordination problems, but also carries social risk and causes potential threats to partners' face that need to be managed. A universal strategy to avoid risks of losing face is the use of politeness, i.e., considerate speech and bodily acts calibrated according to cultural-specific norms, and to social distance and power difference between interacting partners. The greater the social distance or power difference, the higher the risk of losing face and the greater the use of politeness. In this study, we explored the way our closest relatives, chimpanzees and bonobos, coordinate joint action with conspecifics and found that the communicative efforts deployed to enter and exit joint action is affected by partners' power difference and social distance, as predicted by the politeness theory. Our finding suggest that apes may have the cognitive foundation to engage in joint action, notably an understanding of joint commitment and precursors to face management.

Key words
Politeness theory, shared intentionality, joint action, great apes, face management

This manuscript is in preparation, and will be submitted after the submission of the thesis.

Introduction

Many social animal species engage in cooperative activities (Boesch, 2002; Holekamp, Sakai, & Lundrigan, 2007; MacNulty, Tallian, Stahler, & Smith, 2014; Pitman & Durban, 2012; Vail et al., 2013), often by performing complementary roles, anticipating others' future actions, and collaborating to reach a shared goal (Boesch, 2005). Yet, the underlying socio-cognitive

abilities underpinning animal cooperative behavior supposedly differ to those of humans, *insofar as only humans have the abilities and motivations to engage in joint action – collaborative interactions by which participants share intentions to achieve a shared goal* (Clark, 1996). *In joint action,* participants are mutually aware that their own individual action contributes to the overall success of the joint venture (Clark, 1996). To successfully coordinate joint action, participants need to solve generic coordination problems linked to the establishment, maintenance and dissolution of so-called joint commitments (Clark, 2006) – the driving force of joint action.

The coordination problems that arise when participants interact jointly comprise both informational and socio-affiliational imperatives (Clark, 1996, 2006; Enfield, 2006; Heritage & Atkinson, 1984; Stivers et al., 2011). The informational imperatives are related to the need to track shared knowledge, or common ground, and collaboratively refer to it throughout the interaction (Clark & Wilkes-Gibbs, 1986; Enfield, 2006, 2008). The socio-affiliational imperatives require that participants take into account each other's relationship, and engage with each other by an "appropriate level of intensity or intimacy" (Enfield, 2006, p. 412), considering their current social relationship. Politeness – "a concept designating 'proper' social conduct, rules for speech and behavior stemming generally from high-status individuals or groups" (Brown, 2015, p. 326), is a pivotal means by which participants of an interaction manage relationships. In particular, what participants manage is each other's face, with face being best described as publicly manifest self-esteem (Goffman, 1955a).

In their seminal series on politeness theory, Levinson and Brown (1987) recognized that politeness use is a universal phenomenon that appears across cultures, and that can best be viewed as *being attentive to the face needs of others* (i.e., not just to the rules of a certain culture). Because of its universal foundation in human joint action, politeness seems to have an ethological basis, which makes it a particularly interesting phenomenon of study even possibly beyond humans. In joint actions, politeness may be most evident at particular moments where face threats are particularly salient; these may well be those where people are at risk to "loose face" – so to feel embarrassed or humiliated. For example, joint commitments involve face threats, because they afford people's time and effort, and thus restrict their autonomy and freedom to pursue other activities (Brown & Levinson, 1987; Goffman, 1967). Therefore, politeness may be most evident in entry and exit phases of a joint action, since it is that in entry phases, joint commitments are brought into existence and publicly negotiated on, and in exit phases, they are dissolved (Clark, 1996; Heesen et al., 2017). Participants of a joint action thus not only need to collaborate in order

to establish, maintain, and dissolve joint commitment, but they also need to constantly manage the socio-affiliational imperatives of joint action via politeness - as a form of face management (Brown, 2015). These processes are publicly visible and have been empirically studied in the face-to-face encounters of humans (Clark, 1996), and we here first briefly describe the ethology of such a processes, before detailing our research goals.

As a result of the coordinative efforts inherent to spontaneous ad-hoc joint actions, joint action unfolds into three publicly recognizable macro-level phases: the opening, main phase and closing. The opening comprises two phases, a *pre-entry* and an *entry* phase. In interaction openings, participants have to build up participation frameworks (Goffman, 1981b; Kendon, 1990), by first selecting and ratifying participants, with whom they intend to interact and to ensure that their prospective partners are available to engage in a yet unspecified interaction – this is the *pre-entry* phase. This leads to a behavioral pattern and spatial arrangement, the participation framework, that then demarcates who has (and who does not have) a participation status in the interaction (Goffman, 1981b). In what follows, the so-called *entry* phase, participants then jointly commit to the nature and content of their encounter, as well as where and when it should take place (Goffman, 1981b; Kendon, 1976, 1990; Lorenza Mondada, 2009). *Pre-entries* and *entries* also reflect participants' social relationships and status differences; for example, when subordinates recruit superiors they make utterances with more politeness, compared to when they themselves are superior, in order to minimize possibilities of conflict (Morand, 1996). Being polite towards superiors means, for instance, not choosing not to recruit the superior at all, or by using more indirect, ambiguous formulations that allow for denial (Duthler, 2006; Morand, 1996, 2000).

Once participants have jointly committed to what their joint action entails, the actual business of the interaction takes place in the *main body* phase (Clark, 1996); for instance, people eat lunch together, discuss their business plans, or watch a movie.

Finally, to terminate an interaction, participants first need to reach the mutual conviction that they are ready to end the encounter, then progress through several sub-steps (Albert & Kessler, 1976; Broth & Mondada, 2013; Schegloff & Sacks, 1973), and finally take leave (Broth & Mondada, 2013; Clark & French, 1981). The progressing through several sub-step before leave-taking leads to recognizable *pre-entry* phases, as it may involve participants' verbal and bodily efforts (Broth & Mondada, 2013; Heesen et al., 2017). Participants may, for instance, produce summarizing statements, justifications about why the encounter should end, affective positive statements, and expressions displaying the wish for relationship continuity (Albert & Kessler,

1978). Finally, leave-taking may involve an exchange of well-wishing or gratitude expressions, such as good-byes and thank you, respectively (Clark & French, 1981), as well as walking away (Broth & Mondada, 2013) and – if on the phone – hanging up the phone (Clark & French, 1981). Leave-taking is referred to as the closing or *exit* phase. At the affiliative level, closings involve relationship management as well. For instance, in pre-exits, strangers produce more external justifications (e.g., "I have another meeting") to justify the exit from an interaction, more well-wishing statements, and also more statements of positive affect, than do friends (Albert & Kessler, 1978), and friends produce less head-nodding and more looking away than do strangers (O'Leary & Gallois, 1985).

Levinson and Brown (1987) recognized that politeness use serves to mitigate face threats inherent to impositions made during an encounter. As face is fragile and emotionally invested, it requires *constant* management. It is always potentially at stake, meaning that any act that is being made by an individual (in a social interaction or social setting), is *always* face-threatening (yet varying in degree of threat), and therefore requires mutual cooperation by participants (Brown, 2015). A face-threatening act (FTA) then is one that threatens a person's face, by operating in opposition towards the face requirements of the other person. People are concerned to constantly mitigate FTAs, and use politeness as a form of face management. Face management involves behaving in ways that reflect recognition of the social relationship that one may carry with a certain partner, as well as a partner's relative social status to oneself (Brown & Levinson, 1987). More specifically, FTAs increase with participant's social distance (as symmetric relation) and relative power difference (as asymmetric relation), and the culture-dependent ranking of an imposition as face-threatening (Brown & Levinson, 1987). Politeness theory (Brown & Levinson, 1987) thus predicts that the greater the social distance (*D*) or power difference (*P*) between interacting partners, the higher is the risk of losing face (i.e., the higher is the weightiness of a FTA, W_X) - dependent on the seriousness of the imposition defined by a certain culture, R_X (1).

$$W_X = D\ (S, H) + P(H, S) + R_X \qquad (1)$$

In other words, people are more polite with higher-status individuals, and more polite to unfamiliar individuals. So, for example, more politeness efforts are required if a lower status individual (e.g., student) recruits a higher status individual (e.g., professor), as opposed to when a higher status individual recruits a lower status individuals. Additionally, more politeness efforts

are required if an individual recruits a socially distant partner (who they never met before), as opposed to when an individual recruits a familiar partner (someone living in the same village).

Humans are not the only social species to interact cooperatively with partners, and to experience problems inherent to the socio-affiliative imperatives inherent to the social life in a group. For example, our closest relatives, the bonobos and chimpanzees, live in large and complex social societies (Boesch, Hohmann, & Marchant, 2002), in which they experience many social challenges, like managing immediate and postponed intra-group aggression (Kaburu, Inoue, & Newton-Fisher, 2013), gaining access to food and reproductive partners (Surbeck et al., 2012), and securing support from other group members during conflicts (Koyama, Caws, & Aureli, 2006). Previous research has also demonstrated that chimpanzees are sensitive to their partner's and their own reputation based on past experience; they prefer to recruit, cooperate and exchange grooming with cooperative, tolerant and trustworthy individuals (Engelmann & Herrmann, 2016; Gomes, Mundry, & Boesch, 2009; Melis et al., 2006b; Melis, Hare, & Tomasello, 2008) and they seem to gain better access to resources if they behaved cooperatively in the past, since partners strategically chose cooperative individuals and punish competitive free-loaders (Suchak et al., 2016). Bonobos and chimpanzees represent pivotal models for the reconstruction of the cognitive abilities of our last common ancestor (Prüfer et al., 2012; Wildman et al., 2003), and apply as ideal candidates to comparatively study coordination of joint action due to their sophisticated and dynamic ways of interacting. Grooming and play serve as particularly adequate activity contexts for this research, as they are both carried out in lengthy bouts, occur daily, and involve role-reversals and ample communication (Bekoff, 2004; Palagi et al., 2007, 2016; Watts, 2016). Play is also highly reciprocal, requires active engagement (Palagi, 2008; Palagi, 2006), and substantial communication (Demuru et al., 2015; Fröhlich, Wittig, et al., 2016a; Genty et al., 2009).

It is thus worth studying whether great apes are sensitive to the problems inherent to the socio-affiliational imperatives of face management in humans when interacting cooperatively with partners. If they do, then they should exhibit recognizable entry and exit phases, which consist of communicative signals that are flexibly adjusted to the relationship (social bond and rank difference) between partners, in patterns as predicted by politeness theory . Demonstrating that the communicative efforts deployed to coordinate entries and exits of joint action increase with partners' rank difference or social bond strength would suggest that apes experience similar interactional imperatives for which politeness may have evolved in humans (Brown & Levinson, 1987). We suggest that by seeing joint action as an incremental process that requires mutual

coordination efforts by participants to be constructed, maintained, and terminated, applying the concept of joint action (Clark, 1996) and politeness theory (Brown and Levinson, 1987) to the study of great ape social interaction provides an ideal research framework for comparing the joint action coordination skills across different species and beyond humans (Heesen et al., 2017).

The study aims at exploring how chimpanzees and bonobos naturally coordinate interactions with conspecifics. Specifically, we aim at identifying whether entry and exit phases are present in bonobos' and chimpanzees' social play and grooming interactions, and to measure the mean duration by which they are produced, as well as describing the gaze and communication patterns in exit phases. We also aim at testing whether mutual exit types (relying on either mutual gaze exchanges, or on combination of unilateral gazes and tactile gestures) are affected by the species and type of activity. Moreover, we intend to analyze whether, following politeness theory (Brown and Levinson, 1987), the communicative efforts that constitute the entry and exit phases increase with partners' social distance (measured by social bond strength as inverse proxy) and power difference (measured by rank difference), and whether these effects differ between the two species. Since bonobos are generally considered to more tolerant and to more egalitarian compared to chimpanzees (Hare et al., 2007; Palagi, 2006; Tan & Hare, 2013), and egalitarian settings have been shown to attenuate the overall effect of power difference (Morand, 1996), we hypothesize that bonobos' coordination skills will be overall less influenced by rank differences, but more by social bond strength, than chimpanzees (and vice versa). If so, *bonobos should produce fewer and shorter entry and exit phases* compared to chimpanzees when partners' *social bond strength is high*. Comparatively, *chimpanzees should produce fewer and shorter entry and exit phases* compared to bonobos when partners' *rank difference is higher*.

We formulate the following six research questions:

1. Are entry and exit phases present**?**
2. What are the phases' mean durations?
3. What are the kinds of exit types that bonobos and chimpanzees engage in?
4. Are mutual exit types affected by the species and the type of activity?
5. Are the phases affected by social bond strength, and are there species differences in this effect?
6. Are the phases affected by rank difference, and are there species differences in this effect?

Methods

Study species and sites

We collected observational data of two groups of bonobos and three groups of chimpanzees (50 individuals) at four sites. R.H. observed one group of bonobos at San Diego Zoo, USA, from January 2017 to April 2017 (9 individuals, range 4-44 years; see detailed information on group composition in ESM, S2.1), and a second group of bonobos at La Vallée des Singes, France, from April 2017 to September 2017 (16 individuals, range 4-52 years). In this same time period, a trained research assistant collected data from a group of chimpanzees at La Vallée des Singes, France (7 individuals, range 9-23 years). LP observed a second group of chimpanzees at Basel Zoo, Switzerland, during an observation period from October 2016 to February 2017 (9 individuals, range 7-49 years). A.P. observed a third group of chimpanzees at La Réserve Africaine de Sigean, France, from June to August 2018 (9 individuals, range 5-45 years).

The apes at La Vallée des Singes are housed in large enclosures consisting of an outdoor island-type enclosure with a large forest area and climbing structures in grassy areas (8000m^2 bonobos, 7.500m^2 chimpanzees), and an indoor enclosure with enrichment and various climbing structures (600m^2 bonobos, 220m^2 chimpanzees). The food for the bonobos and chimpanzees is distributed five to six times a day and includes daily ratios of primate pellets, fruits, and vegetables (and a daily portion of rice for chimpanzees, and a weekly portion of rice for bonobos). The bonobos and chimpanzees are also occasionally fed with nuts, meat once a week, and eggs twice a week. Both species can also forage for natural vegetation in their outdoor enclosure (wild berries, herbaceous vegetation).

The bonobos at the San Diego zoo are housed in an enclosure including both an outdoor (560m^2) and an indoor enclosure (356m^2). The outdoor enclosure consists of numerous climbing structures, a fresh-water stream, and elements of enrichment. The group receives abundant food types 3-4 times a day (primate pellets and cereals, fruits, vegetables, nuts), and additional supplementary feeds several times a week (honey, peanut butter, popcorn, seeds), often provided through enrichment toys.

The chimpanzees at Basel zoo are housed in a building comprising of six lodges in the inside enclosure (total 233.3m^2) and two in the outside enclosure (total 477m^2). The lodges are provided with climbing structures, ropes, puzzle boxes and other enrichments, and are connected

to each other. The subjects can freely roam between them; two are not visible to visitors. Food is provisioned six times a day, including salad, vegetables, fruits, eggs and primate pellets.

The chimpanzees housed at La Réserve Africaine de Sigean are kept in a 200m^2 indoor enclosure and a 1ha outdoor island area. During the day, they are kept on the outdoor island that consists of natural vegetation, bushes, trees, sand, a surrounding water pond, climbing structures, bridges and shelters. The group is provided with food six times a day, including vegetables, fruits primate pellets and nuts, and chicken and eggs once a week.

Fresh water for all groups is always provided *ad libitum,* and all groups (except the chimpanzees from Basel zoo) additionally can drink from water sources in their outdoor enclosures, such as ponds and streams.

Data collection

Data were collected using continuous focal behavior sampling (Altmann, 1974), maintaining a balanced record of focal samples across individuals. All focal samples (15-30min, depending on the site and its visibility conditions) were recorded using two Panasonic HC-V770 Camcorders with two external Sennheiser unidirectional microphones. Overall, we collected 1298.55h of observation (600.25h for bonobos, 698.30h for chimpanzees), with on average 30.00h per bonobo at San Diego Zoo, on average (±SE) 20.64h (±0.52h) per bonobo at La Vallée des Singes, 37.01h (±0.45h) per chimpanzee at La Vallée des Singes, 30.00h per chimpanzee at Basel Zoo, and 18.08h (±0.42h) per chimpanzee at La Réserve Africaine de Sigean.

To create a variable representing the "social distance" (aka social bond strength) between partners, we recorded social behaviors during focal samples on an iPad 6. Prior to the study, we created individual templates for this purpose using the software FileMaker Pro 15 (v. 15.0.3.305). We recorded data on proximity, socially directed approaches, and social interactions (grooming and play). Proximity data were recorded using scan samples, every 5 minutes during the focal sample, determining the proximity of focal individuals to nearest neighbors and on which group member were present during a given scan. The direction and outcome of socially directed approaches were recorded continuously, determining if a focal individual approached another group member or was approached by one; we recorded only positive (resulting in a social interaction or contact-sitting) or neutral (resulting in proximate sitting) approaches. In addition, we continuously recorded the duration and directionality (who initiated the social interaction) of social grooming and play between a focal individual and a partner.

As a proxy for individuals' "power difference", we recording the outcome of agonistic interactions *ad libitum* (any time, in or outside of focal samples), noting the IDs of involved individuals and assigned winner and loser, or recorded whether the interaction ended undecided without clear winner and loser).

Social bond strength and rank difference

To compute the social bond between partners, we used an inverse proxy referring to the social affinity or social bond strength between partners, the Dyadic Composite Sociality Index (DSI) (Silk, Alberts, & Altmann, 2006; Silk, Cheney, & Seyfarth, 2013) using the socialindices package in R (Neumann, 2017, accessed via https://github.com/gobbios/socialindices). The DSI is defined as (2), where d comprises the numbers (or duration) of behaviors chosen to select the index, f_{ixy} denotes the rate (or duration) of behavior i for the dyad xy, and $\bar{f}_i$ constitutes the mean rate (or mean duration) of that behavior i considering all dyads of the sample (Silk et al., 2013, p. 216).

$$DSI_{xy} = \frac{\sum_{i=1}^{d} \frac{f_{ixy}}{\bar{f}_i}}{d} \quad (2)$$

The values of the DSI scale can range from $0 \rightarrow \infty$; larger DSI values of dyads indicate stronger social bonds, and lower DSI values of dyads indicate weaker social bonds (Silk et al., 2013). Our DSI computation included the measures of the duration of grooming and play between two partners, their proximity rates (i.e., having been in arm reach distance) and their approach rates (i.e., focal having approached the partner or the partner having approached the focal). Because the bonobos of San Diego Zoo were not always let out as one coherent group on the same day, we additionally controlled for observation time and co-residency of dyads for this group.

To compute partners' power differences, we Elo-ratings (Albers & de Vries, 2001; Neumann et al., 2011) for individuals of each site using the "EloRating" package in R (Neumann & Kulik, 2014). Elo-ratings are based on the outcomes of agonistic interactions between partners and depend on the sequence in which interaction take place over time. At the start of the rating procedure, each individual is assigned the same arbitrary value (e.g., 1000), and the rating of the two interacting individuals will be updated according to the outcome of their conflict; winners gain points, losers lose points, with the amount of point gain depending on the expected probability that the higher-rated individual wins (see for detailed explanations Neumann et al., 2011). We assessed

wins and losses based on the occurrences of the following behaviors: aggressing an individual physically or by showing aggressive behaviors, taking away resources, displacing or chasing an opponent, or by fleeing from the opponent, being displaced, showing submissive behaviors such as bared-teeth displays or whimpers, or giving away a resource. Outcomes were considered undecided if no clear winner or loser could be determined (e.g., both showing aggressive behaviors, chasing each other, hitting or slapping another, sharing the resource upon negotiation). Such undecided outcomes are incorporated as a disadvantage to higher-ranked individuals (a decrease of rating, but smaller than with a lost conflict) (Neumann et al., 2011). Since Elo-ratings typically reflect traditional ordinal dominance ranks very closely, for simplicity we refer to Elo-ratings as ranks and Elo-rating differences as rank differences (Neumann, McDonald, & Daizaburo, 2018).

To obtain power differences, we subtracted the rank of the partner from the rank of the phase initiator: negative rank differences then indicated that the initiator of the phase was lower-ranked than the partner, and positive rank differences indicated that the initiator of the phase was higher-ranking to the partner. Larger absolute values in the rank difference indicated that partners were more different in rank than smaller absolute values in the rank difference.

Joint action phases: Definition and coding

To analyze entry and exit phases of the social play and grooming interactions, we used the computer software ELAN, v. 5.0 (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). For behavioral coding of phases, we randomly selected 1242 interaction beginnings (421 for bonobos, 821 for chimpanzees), and 1240 interaction endings (420 for bonobos, 820 for chimpanzees). Of these, 27 interaction beginnings (11 for bonobos, 16 for chimpanzees) were out of sight, and 45 interaction endings (23 for bonobos, 22 for chimpanzees) were out of sight. Our final dataset includes only "visible events", of which the beginning or the end of an interaction was fully visible to the human observer. For social play, this included 307 visible beginnings (156 for bonobos, 151 for chimpanzees) and 305 endings (153 for bonobos, 152 for chimpanzees). For social grooming, this included 836 beginnings (227 for bonobos, 609 for chimpanzees – of which 28 stem from a pilot study conducted at Basel Zoo in 2015) and 825 endings (223 for bonobos, 602 for chimpanzees– of which 26 stem from the pilot study). For mixed contexts (i.e., switching between sex, grooming and/or play), this includes and 72 beginnings (27 for bonobos, 45 for chimpanzees) and 65 endings (21 for bonobos, 44 for chimpanzees).

We defined an entry phase as the process by which partners recruit each other, determine the type of the activity and its temporal and spatial properties (Genty et al., submitted; Heesen et al., 2017, Appendix S1). This is achieved through the exchange of mutual gaze (i.e., both partners simultaneously looking at each other's face), and activity and species-specific initiation signals, such as gestures (e.g., Byrne et al., 2017), vocalizations (e.g., Clay, Archbold, & Zuberbühler, 2015a; Crockford et al., 2018; Fedurek, Zuberbühler, & Semple, 2017), and facial expressions (e.g., Palagi, 2008; Parr, Hopkins, & DeWaal, 1998; Parr et al., 2007). For example, great apes have been shown to use gestures (single or combined in sequences) to initiate playful activities (Fröhlich, Wittig, et al., 2016a; Genty et al., 2009), social grooming (Hobaiter & Byrne, 2014) and joint travel between mothers and infants (Fröhlich, Wittig, et al., 2016b; Halina et al., 2013; Rossano, 2013). We coded an entry phase as present if both partners engaged in mutual gaze before initiating and engaging in grooming or play. If no mutual gaze was attained before the encounter, the entry phase was coded as absent. The entry phase could contain communicative signals such as vocalizations, gestures or facial expressions. We coded an exit phase as present if, before or while leaving their partner, partners engaged in either: unilateral gaze, no gaze but using signals only, bilateral gaze, or mutual gaze (including also those cases where one partner gaze and the other uses tactile gestures). If partners did not engage in either type of gaze described above before or while taking leave of their partner, the exit phase was coded as absent. To compute the duration of phases, we coded the start and end time of the phases. The beginning of the entry phase was coded from the instant where mutual gaze is established, and the end was coded with the onset of movements typical of the activity, i.e., first grooming movement with hands or mouth on partner's body for grooming, first body contact for contact play possibly accompanied by laughter and play-faces (Palagi et al., 2015), and first running movement for pursuit of partner for chase play (Pozis-Francois, Zahavi, & Zahavi, 2004).

We defined an exit phase as the process by which partners reach the mutual conviction that they are ready to terminate the activity and take leave of each other (Genty et al., in revision; Heesen et al., 2017). An exit phase includes the cessation of body movements that created the activity, disruption of attention and taking leave of each other through the use of communicative signals (gestures, vocalizations, facial expressions) and gaze. Since signals of great apes in exit phases are much less studied than signals in entry phases (but see Genty et al., 2009; Hobaiter & Byrne, 2014; Luef & Liebal, 2013), we kept a more inclusive definition for exit phases, as compared to entry phases, allowing a categorization of exit types according to their degree of

"sharedness" based on the signals involved. We classified an exit phase as "unilateral exit type" if individuals used unilateral gaze (one partner is looking at the other), as "bilateral exit type" if individuals used bilateral gaze (partners look at each other's face but not simultaneously), and as "signals only exit type" if individuals used no gaze and communicative signals only. We defined an exit as "mutual exit type", if it involved mutual attention by both partners, either through the exchange of mutual gazes, or through the use of unilateral gaze combined with a tactile gesture of the other partner. An exit was considered if the interaction was not followed by a renewed interaction between the two same partners within a time interval of 2 minutes (i.e., except if partners remained in close proximity and showed signs of reengagement attempts, such as gazing at each other, or communicating to each other). The duration of the exit phase was coded from the cessation of movements, characteristic of the activity to the instant where mutual attention is disrupted by both partners. If one partner is gazing back multiple times, then the exit stops with the first gaze back to the partner. We did not code the exit duration any longer if one partner is trying to reengage the other, but the other individual is already engaged in a new activity with another partner and ignores the reengagement attempts. If an activity suddenly ended due to an interruption event, the end of the interaction was classified as an "interruption-end", but an exit could still be considered present after it if the conditions for exit definitions were met. We additionally recorded the IDs of the initiators of entry phases and the IDs of their partners, and the IDs of the individuals who first terminated the interaction in the exit phase and the IDs of their partners. Initiators were those individuals who recruited a partner in the entry phase, or who first initiated leaving or stop of grooming / play movements in exit phases.

Coding reliability

We carried out an inter-observer agreement test for the presence and duration of entry and exit phases using the software EasyDIAg (Holle & Rein, 2015). The test revealed good agreement between RH and EG (n= 38 clips), and AP and EG (*n*= 20 clips) (Cohen's $\kappa = 0.71$ for both), and moderate agreement between LP and EG (*n* = 33 clips) (Cohen's $\kappa = 0.55$).

Statistical analyses

To explore research questions ("RQ") 4-6, we fitted Bayesian generalized mixed models (*b*GLMMs) and Bayesian linear mixed effect models (*b*LMMs) using the Stan computational framework (http://mc-stan.org/), accessed through the brms package (Bürkner, 2015, 2017) in R v. 3.5.0 (R Core Team, 2013). Each model included 4 MCMC chains, with 10,000 iterations per

chain, of which we specified 2,000 iterations as warm-up to ensure sampling calibration, leading to 32,000 posterior samples for each model. The model diagnostics revealed that the posterior distributions reflect the distribution of the original response values appropriately, R-hat statistics were <1.05, the number of effective samples was >100, and the MCMC chains had no divergent transitions (see ESM S2.2-2.3). For all models, we used the default priors of the brms package (Bürkner, 2017), which were weakly informative with a student's t-distribution of 3 degrees of freedom and a scale parameter of 10 (for the mean of specific prior distributions, see ESM S2.2). For inference, we calculated 95% credible intervals from the posterior distributions and checked whether 0 was comprised in this interval. Due to the incapability of determining a clear phase initiator for mutually initiated exit phases, we could not integrate *n*=66 observations of mutually initiated exit phases in any of our mixed models. We included age difference and whether the dyad was uni-sex as control variables. The results of the complete models can be found in ESM, S2.2.

Specificities of model fitting

To test whether the engagement in mutual exit types depends on species and activity type (RQ 4), we fitted a *b*GLMM –"Model 1", ESM S2.2-, including as dependent variable (DV) the use of mutual exit types (fitting a Bernoulli distribution with presence or absence of mutual exit type as binary outcome), as fixed effects age difference in years, sex difference (i.e., same or different sex), species, and interaction context (i.e., play or grooming). Additionally, we fitted crossed random intercepts for initiator and partner ID. We could not include gaze coding events in three cases; a) cases of the pilot study and nine remaining cases, for which no gaze categories had originally been coded for exits (*n*=35), b) where no gaze category could be coded due to the absence of an exit (*n*=120), and c) where gaze categories could not be determined as they were out-of-sight (*n*=38). We identified a mutual exit type as present (1) if partners either engaged in mutual gaze or used unilateral gaze and there was a tactile gesture by the other partner; we determined a mutual exit type as absent (0), and if partners engaged in any other gaze category while terminating the activity (i.e., using signals only, using unilateral gaze only, using bilateral gaze).

To analyze whether entry and exit phases are affected by the partners' social bond (RQ 5) and rank difference (RQ 6), we first fitted *b*GLMMs for the presence of entry phases (yes/no DV, model 2) and exit phases (yes/ no DV, model 4). Next, *b*LMMs for the log duration (DVs measures in seconds) of entry phases (model 3), and exit phases (model 5). For each of these models, we fitted as fixed effects the social bond strength (DSI; as a measure of social distance), and rank

difference (as a measure of power difference), species, age difference in years, sex difference (same or different sex), interaction effects between social bond strength and species, and interaction effects between rank difference and species. Additionally, we fitted cross random effects with initiator and partner ID (Pinheiro & Bates, 2000). We z-transformed both social bond strength and rank difference for all individuals (both species) to mean = 0 and SD = 1 (Schielzeth, 2010).

To acknowledge the possibility of subject-variation with respect to the effects of DSI and rank difference, we fitted a second version of the models using random slopes for initiators' and partners' social bond strength and rank differences, and compared those models to their first version (excluding the random slopes) using leave-one-out cross-validation (Vehtari, Gelman, & Gabry, 2015) (see ESM S2.4 for results). Since the differences in the expected log predictive density (ELPD) between slope and intercept models were much smaller than twice the estimated standard error, the random slope models were not expected to have a better predictive performance than the intercept only models. In addition, low sample sizes on subject level can lead to unreliable subject-level slope estimates (i.e., for entry and exit phases respectively, where 11 and 17 initiators had interactions with maximally four partners). Due to these reasons, we ran our primary analysis without subject-level slopes and present results with subject-level slopes in the ESM, S2.5-S2.6.

Results

*1. Are entry and exit phases present**?***

From all coded events of visible interaction beginnings and interaction endings, entry phases were present in 76% (90% for bonobos, 69% for chimpanzees), and exit phases in 88% (92% for bonobos, 86% for chimpanzees).

2. What are the phases' mean durations?

The mean duration of entry phases in bonobos was 12.7s and in chimpanzees 11.5s. The mean duration of exit phases in bonobos was 16.8s and in chimpanzees 13.6s. See Table 1 for details on the distribution of the duration of entry and exit duration.

Table 1 Descriptive statistics of the distribution of entry and exit duration.

Phase	Species	Mean (s)	S.D. (s)	Median (s)	IQR[1]
Entry	Bonobo	12.7	20.2	5.9	3.4;14.4
Entry	Chimpanzee	11.5	21.0	6.4	3.0;12.0

Phase	Species	Mean (s)	S.D. (s)	Median (s)	IQR[1]
Exit	Bonobo	16.8	22.3	9.2	5.0;18.5
Exit	Chimpanzee	13.6	14.6	9.0	4.8;16.9

[1]IQR= Interquartile range with lower (1st) and upper (3rd) quartile

3. What are the kinds of exit types that bonobos and chimpanzees engage in?

To terminate activities, bonobos engaged in mutual exit types in 51% of all cases, and chimpanzees in 39%. Bonobos engaged in bilateral exit types in 20% of all cases, and chimpanzees in 21%. Bonobos used unilateral exit types in 29% of all cases, and chimpanzees in 39%. While bonobos never engaged in signals only exit types, chimpanzees did so in 1% of all cases.

4. Are mutual exit types affected by the species and the type of activity?

We found substantial relationships between the probability of using mutual exit types and the two predictors of species and activity type (fig. 4A).

The termination of playful interactions involved clearly more mutual exit types compared grooming interactions (fig. 4B; b (SE) = 1.01 (0.18), 95%CrI [0.67, 1.36]). Additionally, the termination of mixed interactions involved more mutual exit types compared to grooming (fig. 4B; b (SE) = 0.77 (0.32), 95%CrI [0.14, 1.41]).

Bonobos and chimpanzees differed in the degree to which they engaged in mutual exit types. Chimpanzees were less likely than bonobos to produce a mutual exit type compared to a non-mutual exit type, i.e., bilateral, unilateral or signals only (fig. 4C; b (SE) = –0.43 (0.19), 95%CrI [-0.81, -0.05]).

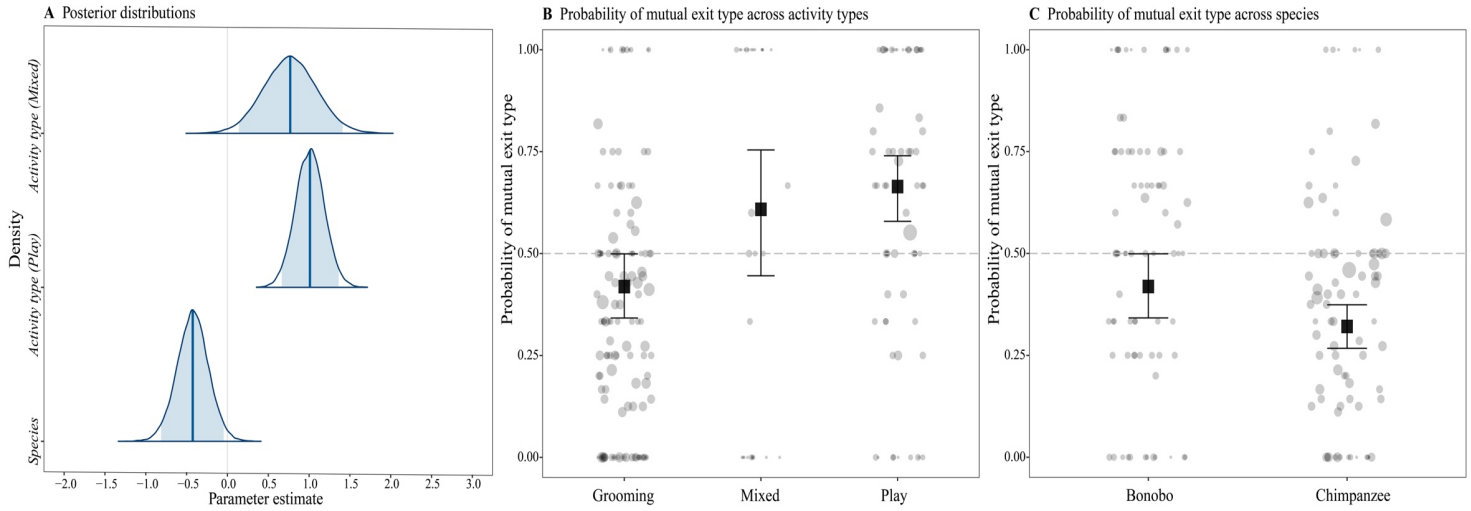


Figure 4 Variation of mutual exit types across activity types ("contexts") and species (model 1). Plot **A)** presents the posterior distribution for the marginal effects of species, activity type (play) and activity type (mixed). The blue area corresponds to the 95% credible interval of the estimated effect. The value of 0 is highlighted by a vertical grey line, and the estimated posterior mean is highlighted by a blue vertical line. Plots **B) and C)** show the predicted probability of mutual exit types for the marginal effects of activity type and species for the Bayesian generalized linear mixed model, and how the model fits the actual data. The upper and lower bars correspond to the upper and lower 95% credible intervals, respectively, and the squares represent the posterior mean. Each circle corresponds to the proportion of mutual exit types present per interaction dyad, out of each dyad's total number of gaze-coded events.

5. Are the phases affected by social bond strength, and are there species differences in this effect?

Entry phase presence

There was no substantial species-difference in how social bond strength (DSI) was related to presence of entries (fig. 5; b (SE) = 0.33 (0.19), 95%CrI [-0.05, 0.70]). However, bonobos seemed more likely than chimpanzees to produce slightly less entry phases with socially close partners compared to distant partners. Furthermore, across the entire range of social bond strength, bonobos were more likely to produce entries than chimpanzees (fig. 5).

Entry phase duration

Bonobos produced shorter entry phases with increasing social bond strength (DSI), while the duration of chimpanzee entries was largely unaffected by social bond strength (fig. 6; b (SE) = 0.17 (0.08), 95%CrI [0.01, 0.33]). In addition, entry phases of chimpanzees and bonobos were on average of similar duration (fig. 6).

Exit phase presence

Bonobos and chimpanzee strongly differed in how social bond strength (DSI) was related to presence of exits (fig. 7; b (SE) = 0.69 (0.25), 95%CrI [0.20, 1.19]). Bonobos were much less likely to produce an exit phase with socially close partners compared to distant partners. In contrast, the probability that chimpanzees produced an exit phase was approximately equal regardless of social bond strength. In addition, bonobos were generally more likely to produce exits compared to chimpanzees (fig. 7).

Exit phase duration

Bonobos and chimpanzees differed in how exit phase duration was related to social bond strength (DSI) (fig 8; b (SE) = 0.27 (0.07), 95%CrI [0.13, 0.42]). Bonobos produced shorter entry phases when the social bond strength increased. Chimpanzees, in contrast, produced longer entry phases when the DSI increased.

6. Are the phases affected by rank difference, and are there species differences in this effect?

Entry phase presence

Bonobos and chimpanzees did not substantially differ in how the presence of an entry phase was related to rank differences between initiators and partners (fig 5; b (SE) = -0.02 (0.30), 95%CrI [-0.61, 0.57]): in both species, there was no pronounced effect of rank difference.

Entry phase duration

In both species, entries became slightly shorter with increasing rank difference (fig 6; b (SE) = 0.13 (0.10), 95%CrI [-0.07, 0.33]). However, this effect was slightly more pronounced in bonobos

Exit phase presence

There was no substantial species-difference with regards to the relationship between exit presence and rank difference (fig 7; b (SE) = -0.44 (0.36), 95%CrI [-1.15, 0.28]). However, chimpanzees seemed to produce slightly less exits with increasing rank difference (fig 7).

Exit phase duration

There was no substantial species-difference in the relationship between rank distance and exit phase duration (fig 8; b (SE) = 0.14 (0.09), 95%CrI [-0.03, 0.31]). However, bonobos seemed to produce slightly shorter exits with increasing rank difference.

Table 2 summarizes the results for the interaction terms between species and social bond strength, and species and rank difference.

Table 2 Summary of model results for interaction terms. Downward-pointing arrows indicate that the relationships between social bond strength (DSI) or rank difference and phase presence or duration were more *negative* for bonobos than for chimpanzees. Upward-pointing arrows indicate that the relationships between social bond strength or rank difference and phase presence or duration were more *positive* for bonobos than for chimpanzees. Horizontal arrows indicate that there was clearly no species-difference. The parentheses indicate that 0 was included in the 95% credible intervals, thus the species-difference was uncertain.

	Interaction	
Model	species :DSI	species :rank difference
Model 1 (Entry presence)	(↘)	→
Model 2 (Entry duration)	↘	(↘)
Model 3 (Exit presence)	↘	(↗)
Model 4 (Exit duration)	↘	(↘)

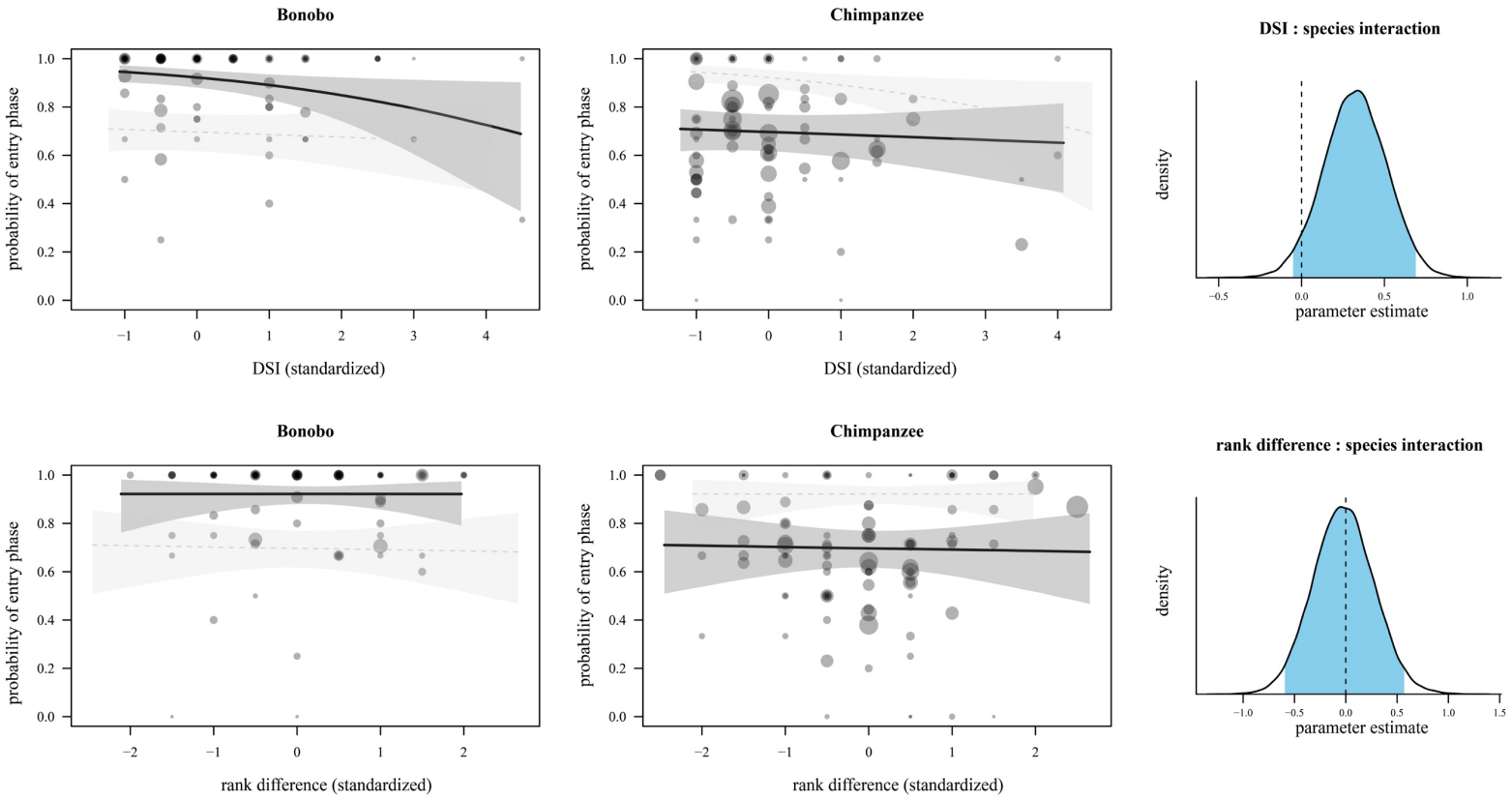


Figure 5 Variation in the presence of entry phases (model 2). The top row shows the relationship between social bond strength (DSI) and the occurrence of entries, separated by species and the posterior distribution of the interaction term between species and DSI. Each circle represents the proportion of events with entries present out of all events. The bottom row shows the relationship between rank difference and the occurrence of entries. Prior to producing this figure, values for DSI and rank difference were binned with a width of 0.5. Circle size is proportional to the total number of events. Circles are colored with transparent filling and hence overlap (multiple individuals with the same value) will appear darker. The lines represent model fits along with shaded 95% credible bands. To aid comparison between species, in each panel we added the model predictions for the other species in fainter color and with dashed lines. In the posterior distributions, the blue area corresponds to the 95% credible interval of the estimated effect with the value of 0 highlighted by a vertical dashed line.

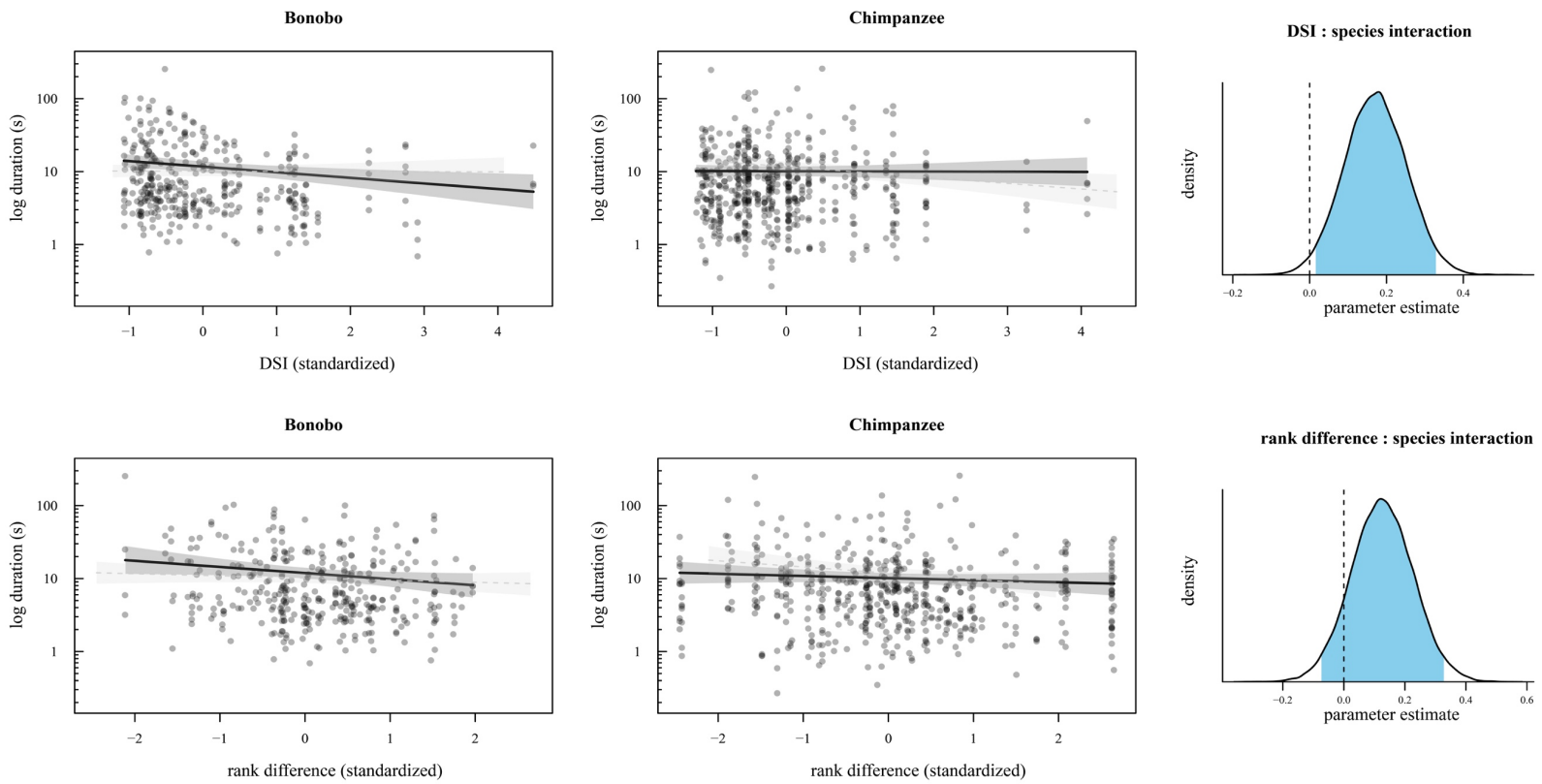


Figure 6 Variation in the log duration of entry phases (model 3). The top row shows the relationship between social bond strength (DSI) and the log duration of entries, separated by species and the posterior distribution of the interaction term between species and DSI. The bottom row shows the relationship between rank difference and the occurrence of entries. Each circle represents the log duration of entry events depending on the partners' DSI or rank difference. Prior to producing this figure, values for DSI and rank difference were binned with a width of 0.5. Circles are colored with transparent filling and hence overlap (multiple individuals with the same value) will appear darker. The lines represent model fits along with shaded 95% credible bands. To aid comparison between species, in each panel we added the model predictions for the other species in fainter color and with dashed lines. In the posterior distributions, the blue area corresponds to the 95% credible interval of the estimated effect with the value of 0 highlighted by a vertical dashed line.

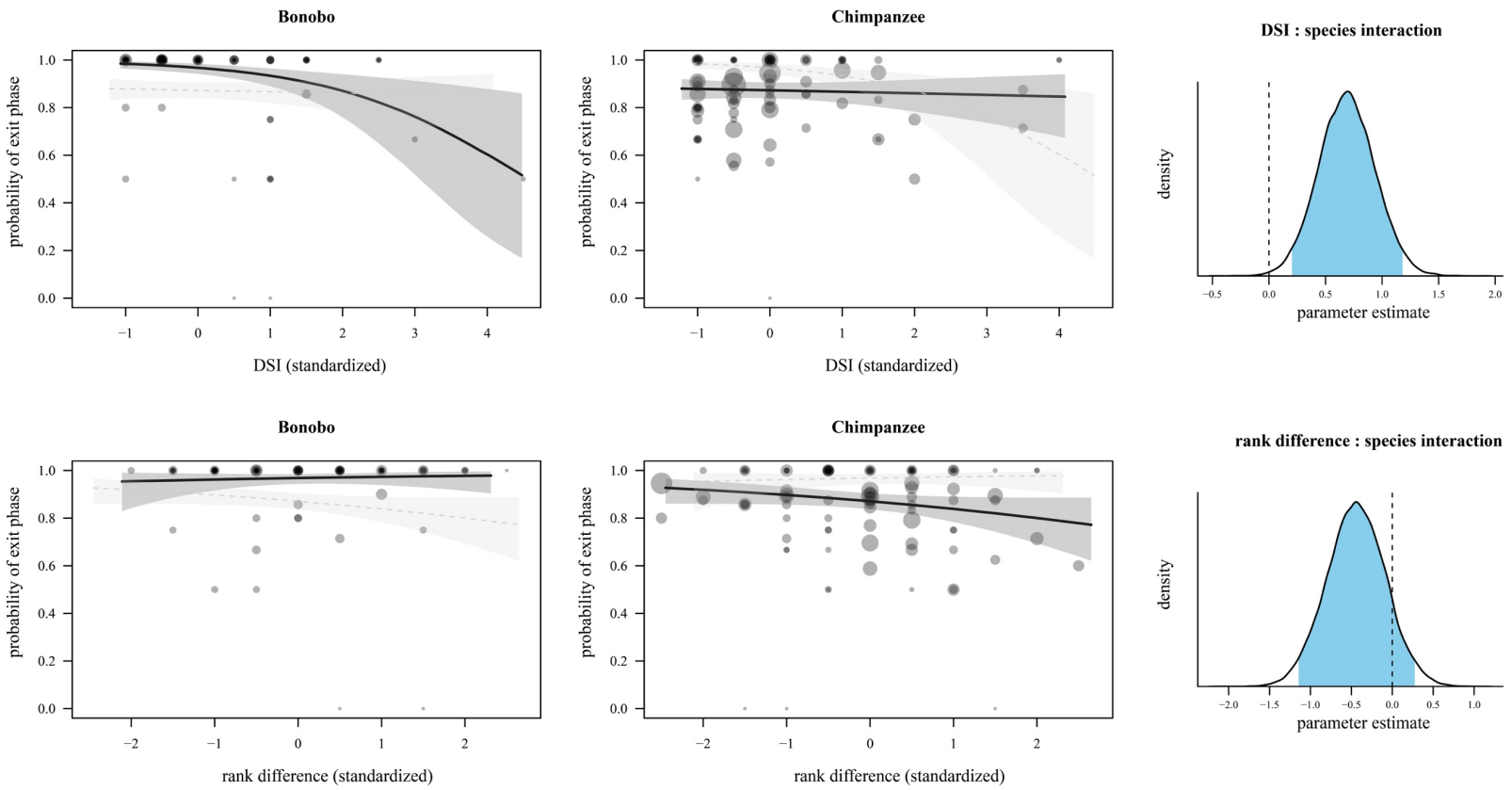


Figure 7 Variation in the presence of exit phases (model 4). The top row shows the relationship between social bond strength (DSI) and the occurrence of exits, separated by species and the posterior distribution of the interaction term between species and DSI. Each circle represents the proportion of events with exits present out of all events. The bottom row shows the relationship between rank difference and the occurrence of exits. Prior to producing this figure, values for DSI and rank difference were binned with a width of 0.5. Circle size is proportional to the total number of events. Circles are colored with transparent filling and hence overlap (multiple individuals with the same value) will appear darker. The lines represent model fits along with shaded 95% credible bands. To aid comparison between species, in each panel we added the model predictions for the other species in fainter color and with dashed lines. In the posterior distributions, the blue area corresponds to the 95% credible interval of the estimated effect with the value of 0 highlighted by a vertical dashed line.

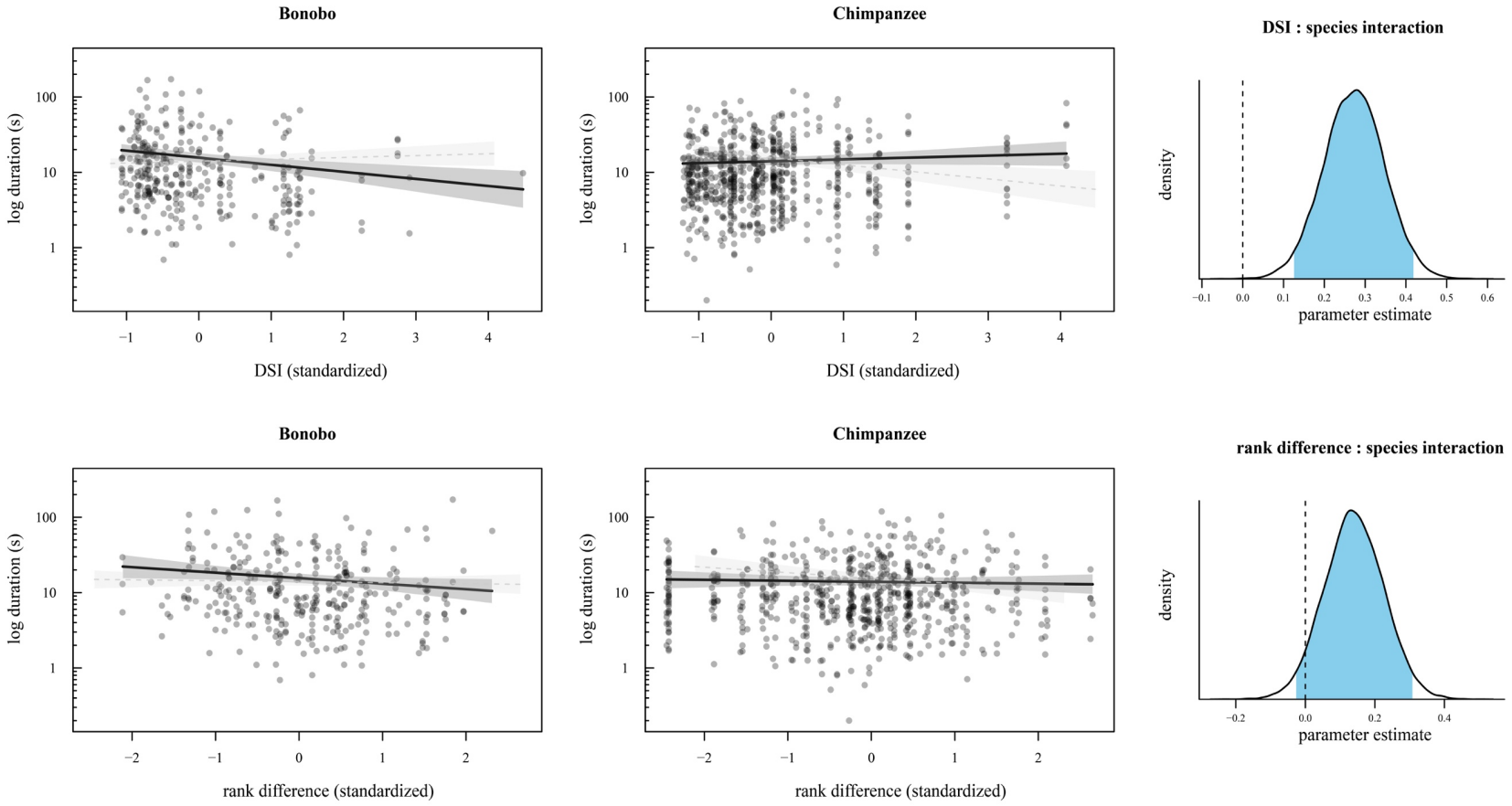


Figure 8 Variation in the log duration of exit phases (model 5). The top row shows the relationship between social bond strength (DSI) and the log duration of exits, separated by species and the posterior distribution of the interaction term between species and DSI. The bottom row shows the relationship between rank difference and the occurrence of exits. Each circle represents the log duration of exit events depending on the partners' DSI or rank difference. Prior to producing this figure, values for DSI and rank difference were binned with a width of 0.5. Circles are colored with transparent filling and hence overlap (multiple individuals with the same value) will appear darker. The lines represent model fits along with shaded 95% credible bands. To aid comparison between species, in each panel we added the model predictions for the other species in fainter color and with dashed lines. In the posterior distributions, the blue area corresponds to the 95% credible interval of the estimated effect with the value of 0 highlighted by a vertical dashed line.

Discussion

We explored how our closest relatives, bonobos and chimpanzees, coordinate joint activities with partners. In particular, we investigated the way two social partners engage and disengage from a joint action through gaze and/or communicative signals. Our results show that chimpanzees and bonobos commonly exchange gaze and other communicative signals to enter into and exit from interactions. This gave rise to a sequential, temporal organization of joint action comparable to that of humans, with potential establishment of participation frameworks in entry phases, and potential leave-taking exchanges in exit phases (e.g., Broth & Mondada, 2013; Clark & French, 1981; Schegloff & Sacks, 1973). The presence of entry and exit phases in bonobos and chimpanzees joint actions suggests the possibility that great apes use communication to solve similar coordination problems found in human joint action. The sequential structure characterizing great ape joint action resembles that of humans, where participants achieve a mutual state of attention in *entry* phases (Kendon, 1976, 1990; Lorenza Mondada, 2009) and mutually condone their leaving in *exit* phases (Broth & Mondada, 2013; Schegloff & Sacks, 1973) through a step-wise exchange of signals and gazes. This suggests a potential mutual awareness from both partners that they are doing something together. Although overt behaviors do not serve as full proof for underlying cognitive processes (Leavens et al., 2017), we argue that communicative expressions used for similar purposes of joint action coordination in both humans and great apes may indicate a shared socio-cognitive capacity for joint action coordination (see also Rossano, 2013), notably for the understanding of joint commitment and possible precursors for face management.

Regarding such potential precursors of face management, our findings indicate that the coordination of joint activities in apes mirrors the patterns predicted by politeness theory in humans (Brown & Levinson, 1987), in that the communicative efforts deployed during entry and exit phases were affected by the social bond strength and rank difference between interacting partners, although differently across species. In humans, politeness theory (Brown & Levinson, 1987) predicts that politeness use (as a function of the weightiness of a FTA) increases with the social distance and power difference between two partners. In accordance with our hypothesis, bonobos – compared to chimpanzees - did produce fewer and shorter entry and exit phases when the social bond strength between partners increased (i.e., when partners were "better friends"). When looking at how rank differences affect the presence and duration of phases, the pattern is less clear, and possibly more complex; regarding the presence of phases, chimpanzees seemed slightly less likely than bonobos to produce exit phases when the initiator of the phase increased

in rank relative to the partner. Regarding the duration of phases, however, bonobos were more affected by rank differences compared to chimpanzees, producing shorter entry and exit phases when the initiator increased in rank relative to the partner. These findings have implications for our understanding of how the two species coordinate joint action and what kinds of social parameters predict tactical decisions of individual communication patterns, in other words how they calibrate their communication in species-specific and partner-specific ways.

On the one hand, one could argue that bonobos may be more likely than chimpanzees to exhibit patterns similar to those predicted by politeness theory, in that they may overall be more affected by rank difference and social bond, compared to chimpanzees. This would suggest bonobos are more sensitive to their relationships with group-members than chimpanzees, and more capable of flexibly adjusting their communicative efforts to these social matrices. On the other hand, the phase duration in bonobos may have been more affected by rank difference than in chimpanzees simply because phase duration can be influenced by certain cofounding variables. Duration could increase as a result of longer negotiation, increased rate of signaling, longer signals being produced (closely bonded individuals interacting frequently with one another can potentially reduce the amplitude of their signals), or longer signal interval and longer response waiting, for example. Although, in this study, we have not looked into the details of the signaling patterns during these phases, it has been demonstrated that chimpanzees and bonobos differ in the speed with which they respond to their communicative partner's signals– with chimpanzees being generally slower than bonobos (Fröhlich, Kuchenbuch, et al., 2016). Since the duration of phases includes gaps of response waiting, after the production of communicative signals, phase duration may not be the most accurate measure to compare joint action coordination between the two species. An exciting avenue for future studies would be to examine in more detail *how* communication patterns are directly affected by social variables. Assuming that phase presence is a more suitable explanatory variable than phase duration when studying politeness theory in apes, then bonobos are clearly more affected by social bond, and chimpanzees slightly more so by rank differences (at least in exit phases). This may perhaps not be surprising, since bonobos are known to behave more prosocial, tolerant and less despotic towards group members compared to chimpanzees (Hare et al., 2007; Tan, Ariely, & Hare, 2017; but see Jaeggi, Stevens, & Van Schaik, 2010), and since egalitarian behavior generally attenuates the effect of "power politeness" (Morand, 1996). The fact that bonobos responded more to the parameter of social bond strength also suggests that bonobos, more so than chimpanzees, may take into account interactional

histories with certain individuals – a possible stage for the development of dyad-specific conventions, partner-appreciation and negotiation skills. Less despotic societies allow for more negotiation between partners, and are therefore considered to be more socially complex, as well as to foster the emergence of more complex communication systems (Dobson, 2012; Freeberg, Dunbar, & Ord, 2012; Maestripieri, 2005). Bonobos also tended to engage in more mutual exit phases (gazing into each other's eyes when terminating an activity, or producing tactile signals while the other partner gazes), supporting the idea that bonobos are more sensitive to the partners gaze in exit phases, monitoring the partner's attention towards them, and that the termination of interactions involve more mutual awareness by partners (possibly to manage future relationships). Nonetheless, previous studies have found that chimpanzees also tailor some of their communicative efforts to the social bond with their partner (Fedurek, Machanda, Schel, & Slocombe, 2013), and chimpanzees also did not seem completely unaffected by social bond in our study (i.e., at least they showed a slight decrease of likelihood of producing an exit phase when social bond strength increases). It could be that chimpanzees may only be affected by social bonds when it is strategically required, for instance to enhance a partners' support in future conflicts, while bonobos constantly take into account their partners' shared history to avoid immediate conflicts.

Overall, our findings suggest that politeness may indeed have an ethological basis, by which similar patterns related to the socio-affiliational imperatives may be found in the interactions (or possibly joint actions) of our closest relatives. This suggests the possibility that the interactional patterns predicted by politeness theory have an evolutionary origin, with humans and great apes sharing a biological basis, or precursor, for the sensitivity to partner's face. The question arises about the functionality of face management in apes more generally, and why these species may have evolved behavioral adaptations as such. In humans, the functions of politeness use in humans seem manifold (Brown, 2015), but for most may be considered quite instinctive and basic: people intend to manage their reputation - i.e., individuals want their own personality and values to be appreciated and approved by others (Brown & Levinson, 1987)-, and to avoid getting into trouble with members of their group (Boyd & Richerson, 1992; Limberg, 2009). Bonobos and chimpanzees experience similar social pressure, yet, it remains to be studied how communicative strategies toward partners of differing rank or social bond strength may facilitate (or threaten) an individual's success within the group, or how the violation of possible "interactional rules" may lead to immediate or delayed punishment by receivers. Although the evidence remains anecdotal,

we indeed observed a possible case of punishment/protest against the violation of producing an exit phase in bonobos, where an adult female did not produce signals or gaze to exit from a grooming encounter with a socially distant and higher ranking adult female. The recipient female showed signs of negative emotional arousal (bared teeth facial expression), followed the other female and forced her into a sexual interaction, mounting her with screams and a bared teethed face. To fully understand whether the patterns we find are comparable to normative politeness formulae in humans, we advocate future research to explore the consequences of violations when individuals fail to coordinate entries or exits according to the rank and social bond with a certain recipient, for instance by studying the immediate consequences (e.g., protest, aggression by recipient) and delayed consequences (e.g., reduction of social support by recipient and interaction rates between partners) after the violation of norm.

Joint action coordination in humans is based on partners' mutual awareness that they are pursuing a shared goal, and on their motivation to achieve that shared goal (joint commitment), by taking complementary roles and helping each other (Tomasello & Moll, 2010). The findings from a multitude of previous experimental studies comparing the capacity for shared cooperative activities between humans and great apes revealed that humans invariably outperform apes with respect to cooperation (Herrmann, Hernández-Lloreda, Call, Hare, & Tomasello, 2010; Warneken et al., 2006, 2007; Warneken & Tomasello, 2006); this has led researchers to conclude that humans, but not other animals, have "shared intentionality", engaging in joint actions by sharing psychological states to achieve common goals (Call, 2009; Tomasello & Carpenter, 2007). Politeness allows for the creation of shared intentions without threatening participants' face. The presence of face management in great apes thus raises the possibility of the presence of shared intentionality in great apes (but see Call, 2009). Although previous experimental research comparing great apes with children has largely revealed negative evidence for shared intentionality in our closest relatives (Herrmann, Hernández-Lloreda, et al., 2010; Warneken et al., 2006, 2007), these conclusions are mostly based on the fact that ape subjects never tried to reengage a reluctant partner during interruptions of social games in artificially laboratory settings (Kingstone et al., 2008). We believe that future research should expand the study of shared intentionality by focusing on behavioural manifestations of shared intentionality in more ecologically relevant situations, i.e. during interruptions of natural interactions with conspecifics, or through signals and gaze patterns used for the purpose of joint action coordination.

To conclude, a joint action framework and politeness theory (Brown & Levinson, 1987) provide an interesting research avenue for studying joint action coordination skills and shared intentionality beyond humans. We showed that the relationship between communication and social variables, as predicted by this theory, are prevalent in captive bonobos and chimpanzees; our data are limited to captive individuals however, and therefore more data from wild individuals is needed to fully understand the relevance of these behavioral patterns in natural conditions. In humans, for instance, one condition that can lead to the reduction of an exit phase is the "continuing state of incipient talk" (Schegloff & Sacks, 1973), by which people remain in the same spatial configuration in which their interactions occur, for instance people sharing the same office. In this scenario, it is often superfluous to coordinate an exit phase, since people have continuing opportunities to re-engage. This continuing state of incipient talk may also apply to the captive groups of bonobos and chimpanzees that represent our data. The fact that we still find a considerable likelihood of individuals' engaging in exit phases while terminating their encounter, suggests that exit phases should also be present in fission-fusion societal conditions in the wild. In the quest for studying the evolution of sophisticated joint action skills in humans, we encourage more studies to test politeness theory using a joint action framework in other primate species, as well as in other non-human animal species that are more distantly related to humans. One suitable activity type for the further investigation of joint action coordination seems to be social play, since our study showed that play involves more mutual exit types compared to other types of activities, suggesting that partners share a more common awareness about their social act; this assumption is also supported by our previous theoretical claim about the cooperative nature of play (Heesen et al., 2017). Cooperative behaviors used to coordinate joint action may thus best be observed during playful activities, in which actions are less routinized and more improvised (Bekoff & Allen, 1998; Spinka, Newberry, & Bekoff, 2001). We also particularly advocate species where the successful coordination of joint action increases fitness benefits, such as in cooperative breeders (e.g., arabian babblers: Carlisle & Zahavi, 1986) or cooperative hunters (e.g., killer whales: Smith, Siniff, Reichle, & Stone, 1981). Using other, more distantly related species to compare politeness patterns in communication with that of great apes and humans could furthermore deliver invaluable insights for our understanding of how politeness links to human language; if politeness patterns resided in the bodily communication of other non-linguistic species, politeness may not solely represent a cultural, linguistic phenomenon that is uniquely human.

Acknowledgements

We thank all the Zoos, including their directors and animal keepers, for letting us conduct our observations at their sites, and for providing invaluable support during the study periods. We are also grateful for the invaluable help of our trained research assistants Aude Pajot, Cindy Maslarova, Laura Perrenoud. Additionally, we thank Martin Götz for discussions and input regarding the Bayesian statistic. The present research was supported by the Swiss National Science Foundation (Grant No. CR31I3_166331 awarded to AB and KZ).

Ethics statement

We received ethical agreement for this study from the Commission d'Ethique de la Recherche of the University of Neuchâtel (agreement number: 01-FS-2017), the internal ethical committee of La Vallée des singes, the UCSD Human Research Protections Program (Project number: 161452S), the internal ethical committee of the *Réserve Africaine* de *Sigean, and the Kantonales Veterinäramt BS at Basel Zoo.*

Chapter 4. Joint commitment understanding of apes with a human partner

Raphaela Heesen[1], Adrian Bangerter[1], Guéry Jean-Pascal[5,] Katia Iglesias[2], Federico Rossano[3], Klaus Zuberbühler[1,4], Emilie Genty[1]

[1]University of Neuchâtel, [2]University of Applied Sciences and Arts of Western Switzerland, [3]University of California San Diego, [4] University of St Andrews, [5] La Vallée Des Singes, France

Abstract

Cooperation is a widespread phenomenon among many animal species. However, humans seem to have evolved a unique motivation to cooperate called "we" or shared intentionality. When engaged in joint activities, partners usually create a joint commitment to achieving a shared goal. One way to assess whether partners understand joint commitment is to test their reaction to unilateral interruptions of the joint activity. When one of the two partners breaks from, one should expect protest and reengagement attempts towards the recalcitrant partner. Previous experimental research on chimpanzees showed that subjects never attempted to reengage a human partner when a social game was interrupted, suggesting that apes lack joint commitment. Here, we revisit this claim by testing chimpanzees and bonobos of different age classes in a less complex but more reciprocal "tug-of-war" game. We played with subjects and suddenly interrupted the game, while manipulating the attentive state of the experimenter. Our results showed that older individuals attempted to reengage more often than younger ones, and older individuals used more game-related behaviors when reengaging their partners compare to younger ones. We discuss potential explanations for these age and species differences in performance and implications in light of the shared intentionality hypothesis.

Keywords: triadic engagement, joint commitment, shared intentionality, great apes

This manuscript is in preparation, and will be submitted in due course after the thesis submission on 30th September 2019.

Introduction

Many social animal species engage in cooperative activities, including collaborative hunting activities in different species of primates, carnivores and fish (Boesch, 2002a; Holekamp et al., 2007; MacNulty et al., 2014; Pitman & Durban, 2012; Smith et al., 1981; Vail et al., 2013), mate guarding (Watts, 1998), and border patrolling (Watts & Mitani, 2001). However, human joint action seems to have reached uniquely sophisticated forms, allowing for the accumulation of cultural artifacts, the emergence of social institutions and the evolution of conventionalized languages (Tomasello, 2016).

Much research has investigated the prerequisites needed to engage in joint action, especially by comparing behaviors between humans and other non-human species in order to understand possible differences in the underlying psychological processes (e.g., Herrmann, Hernández-Lloreda, et al., 2010; Warneken et al., 2006; Warneken & Tomasello, 2006). The current consensus from this work is that human joint actions – as opposed to non-human animal cooperation – require a certain form of joint commitment (Clark, 2006), by which individuals share responsibility for performing a joint action toward a joint goal. Once individuals agree to a joint activity, they are expected to perform their roles and take responsibility for their own action, in order to achieve the joint goal and hence to satisfy the group's shared intention (Tuomela, 2007); thus, joint commitment is considered as the driving force behind people's actions and decisions (Miller, 2006) when engaging in joint actions. A sudden breaking of this commitment by one participant has negative consequences for the group's success in achieving their shared goal, and thus should either be avoided (Gilbert, 1999; Miller, 2002). If interruptions are not communicated appropriately, individuals feel entitled to protest against others' breaking of the obligations entailed by the joint commitment (Gilbert, 1990).

If joint commitments are the consequences of people's shared intentions (or plans) to achieve shared goals, studying whether or not people show signs of joint commitment can reveal insights into their perception of shared goals (hence their ability for shared intentionality more generally). To empirically study joint commitment understanding, on may either observe people's reaction towards perturbations that prevent the group's success in reaching a shared goal (e.g., Mayor & Bangerter, 2015), or by study individuals' reactions after their partner in the joint activity suddenly disengages from the interaction without warning (e.g., Gräfenhain et al., 2013). If partners had formed shared goals, the reaction of the subject towards the sudden leaving of the partner should involve protesting, complaining or actively attempting to reengage the partner in

the activity (Gräfenhain et al., 2013; Ross & Lollis, 1987; Warneken et al., 2006).

In children, the understanding of the obligations entailed by joint commitments emerges with their general awareness of shared intentions and goals (Gräfenhain et al., 2009, 2013) and social norms (Rakoczy, 2007; Rakoczy & Schmidt, 2013) around the age of three years. For example, Gräfenhain and colleagues (2009) tested 2-3 year olds' reactions to an interruption initiated by the experimenter in two conditions: a non-commitment condition (i.e., child and experimenter play their own game independently) and a commitment condition (i.e., partners play the same game together). Three-year-olds - but not 2 year-olds – attempted to reengage their partner significantly more often in the commitment condition than in the non-commitment condition, suggesting they understood that they were engaged in a joint activity. In further tests using the same two conditions, the authors investigated whether three and four-year olds communicate their own leaving if they decided to interrupt the game themselves (Gräfenhain et al., 2009). Only by around 5 years of age do children arrive at a level of distinguishing implicit from explicitly expressed joint commitments with partners, by knowing when they are committed by task interdependence alone (implicit), or by explicit verbal acknowledgement (Kachel & Tomasello, 2019).

The question of whether shared intentionality is a human unique quality or whether it has ancestral primate roots has mainly been addressed by comparative studies examining the understanding of joint commitment in young children and great apes. Warneken et al. (2006) tested the reaction of human children and chimpanzees to the sudden interruption of a cooperative social game. When the experimenters suddenly ceased to take their role during the game, only children - but never chimpanzees – attempted to reengage the reluctant partner to continue the game. The authors concluded that the chimpanzees lacked a critical requisite for shared intentionality, namely the awareness of joint commitment, and more broadly, the interest to perform shared goals (Tomasello & Moll, 2010).

An alternative explanation for this finding may be that the game was simply not adapted for ape species (since they were developed for human children) or too complex (MacLean & Hare, 2013), resulting in disinterest or miscomprehension on behalf of the ape subjects. In addition, the objects used within these games were novel to the subjects, and since chimpanzees are neophobic towards novel objects (Herrmann, Hare, Cissewski, & Tomasello, 2011), their continuing the game alone instead of reengaging the partner may be due to their intention to explore the objects before engaging with them properly (see also MacLean & Hare, 2013). Moreover, the chimpanzees were

rewarded with food for their participation after the trials; since obtaining a reward could be considered a competitive and individualistic –rather than mutual – goal, it remains unclear to which extent subjects considered these games as shared. In addition, the study only tested three infants, with an average age of 3.75 years. Experiments involving less complex social games, with more naturalistic elements (e.g., water or branches), species-typical behaviors deployed by the human partner (e.g., experimenters imitating ape laughter), or other species (bonobos: MacLean & Hare, 2013; Pika & Zuberbühler, 2008, orangutans: Gruber 2014), showed that species were indeed motivation to reengage a reluctant partner after an interruption was initiated by the partner, even from a younger age (e.g., MacLean & Hare, 2013; Pika & Zuberbühler, 2008 Gruber, 2014).

Nonetheless, because human children only develop a more complex understanding of joint commitment at around 5 years (Kachel & Tomasello, 2019), it may be worth asking whether there may be age differences in the joint commitment understanding of great apes. None of the previous studies have investigated a possible developmental trajectory of joint commitment understanding in great apes, even though great apes develop more profound skills in communicating later in ontogeny, possibly through interactional experience (Fröhlich, Müller, Zeiträg, Wittig, & Pika, 2017; Genty, 2019). For example, young chimpanzees learn which gestures function most accurately in reaching one's goal by gesturing in sequences, and then only later become more proficient at using the right gesture as single signal (Hobaiter & Byrne, 2011a), or increasing signal functional specificity (Genty, 2019).

In addition, the social games used to test subjects lacked sufficient reciprocity, spontaneity or were possibly ritualized. MacLean and Hare (2013) had experimenters engage in a ball game with the subjects, handing a ball back and forth. Yet, in order to reach a point where subjects were sufficiently engaged in the game, the experimenter verbally and non-verbally (e.g., reaching out to the ball) encouraged the subject to hand back the ball. If the subject did not return the ball, the ball was retrieved by the experimenter and the game re-started by handing the ball back again to the subject (MacLean & Hare, 2013). In a second game of the same study, the experimenter tickled the subject with a stick, while the ape's role consisted in being tickled. Here, the experimenter took a more active role in relation to the subject, and the subject's motivation to the game remained largely based on the individual goal of being tickled. If both partners had tickled each other, this could have been considered as shared, since both partners would share the same goal ("we tickle each other"). In Gruber (2014), it remains unclear to which extent a shared goal had been established between subject and experimenter, and to which extent the game involved reciprocity,

since the experimenter merely interacted with the subjects using a stick (i.e., no clear definition of the cooperative nature of the "stick" game is provided). Therefore, it cannot be ascertained whether the behaviors identified as reengagement attempts during interruption periods can be interpreted as signs of understanding joint commitment, or whether they instead constituted attempts to reach individual goals. Moreover, touching the experimenter or the object were noted as reengagement attempts (MacLean & Hare, 2013), but such behaviors may alternatively be explained as intention to get into bodily contact with the experimenter or gaining access to the object rather than referring to the previous game. Handing back the game-object would have revealed stronger evidence that subjects attempt to reengage a partner to resume their original goal, namely to play the game together.

Our goal here is to address these shortcomings and provide a new look at joint commitment understanding in bonobos and chimpanzees by involving different age classes of both younger and older individuals (to determine possible developmental differences), using a more reciprocal game, and systematic definitions of what should count as communicative reengagement attempts (i.e., communication should be game-related). We additionally introduced a control of the experimenter's attentive state (attentive vs inattentive), in order to ensure that subjects' reengagement attempts are not merely a consequence of reactions by subjects to the partners' attention in these breaks. In other words, the reengagement attempts by subjects in previous studies could have been a consequence of the fact that subjects felt urged by their partners to react, after the partner had broken from the game. Therefore, the apes may have communicated as they felt addressed, or frustrated due to the passive but staring partner.

More specifically, we aim at revisiting the question of whether great apes are capable of recognizing joint commitment when engaging in social games with a human experimenter, by engaging in a cooperative, voluntary tug-of-war game with the apes that requires two roles to be played reciprocally (i.e., one pulls strongly while the other partner relaxes their pulling strength and vice versa in rapid sequence). If the subjects understand joint commitment, and if they are motivated to engage in cooperative games, they should attempt to reengage their partner by handing back the game object during interruption periods. Furthermore, we assess the existence of a developmental trajectory of joint commitment understanding in adult chimpanzees (Experiment 1), adult bonobos (Experiment 2), and infant bonobos (Experiment 3).

General Methods

The cooperative game consisted in a tug of war between a human experimenter and an ape subject using a plastic garden hose. The game required both partners to pull simultaneously at the opposite ends of the hose, but one partner at a time had to reducing their pulling force (i.e., thereby self-handicapping), leading to a sequence of back-and-forth movements of the hose. Subjects were either habituated to the hose prior to the experiment, or were habituated to it prior to the experiment by leaving some hoses in their enclosure. The experimenter and subject played the game through the cage mesh, with the experimenter standing outside of the cage (fig. 9) for adult chimpanzees and bonobos, and standing inside the cage together with the infant bonobos (fig. 10). The adult subjects were tested on a voluntary basis and were not isolated from the other members of their group. Testing sessions were filmed by a researcher (RH) or a research assistant, who was standing outside the cage 1.5 m away using a Panasonic HC-V770 Camcorder placed on a tripod and with an external Sennheiser unidirectional microphone.

Procedure

Each trial began either by an initiation of the experimenter, holding the hose on one end and presenting the other end to the subject while verbally encouraging the subject, laughing or imitating ape laughter (%), or by a voluntary initiation attempt by the subject to hand a hose already present in their enclosure (%). As soon as the subject grabbed the hose, the experimenter initiated a back-and-forth movement, to encourage the subject to play. If the subject did not pay attention, the experimenter encouraged the subject several times by saying the subject's name, laughing, shaking the hose and handing it conspicuously to the subject. The game started once both partners actively perform the repeated back-and-forth movement; then suddenly the experimenter interrupted the game by dropping the hose. the interruption period lasted 15 s or 30 s, depending on the experiment (see details in each experiment method section).

During these interruptions, two experimental conditions were applied: the still-faced condition, in which experimenters faced the subject while standing still (figs. 9A and 10A), and the back-turned condition, in which experimenters turned their back to the subject and stood still (figs. 9B and 10B). The order of presentation of the conditions was counterbalanced within subjects and across sessions. A play session lasted as long as subjects participated and could include interruptions once a tug-of-war sequence was established at any given time point. We attempted to apply a within-subject design with repeated measurements, by which all subjects were

tested at least once in each of the two conditions. The number of trials varied across subjects and experiments (see each experiment section for details).

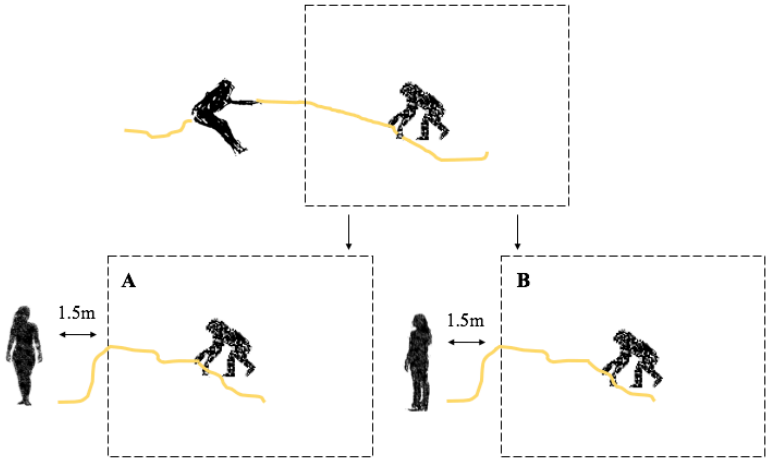


Figure 9 . Experimental design for Experiments 1 and 2 with adult chimpanzees and bonobos. The experimenter is standing outside of the cage and engages in a tug-of-war game with the subject through the cage mesh, using a garden hose (in yellow). In the still-faced condition (A), the experimenter interrupts the game and faces the subject while remaining still for the interruption period; in the back-turned condition (B), the experimenter interrupts the game and turns his back to the subject while remaining still for the interruption period.

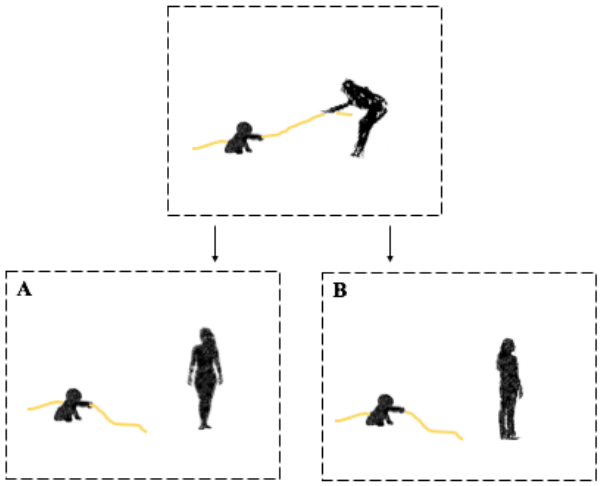


Figure 10 Experimental design of Experiment 3 with infant bonobos. The experimenter remains inside of the cage and engages in a tug-of-war game with the subject in the cage, using a garden hose (in yellow). In the still-faced condition (A), the experimenter interrupts the game and faces the subject while remaining still; in the back-turned condition (B), the experimenter interrupts the game and turns her back to the subject while remaining still.

Video coding

We coded the behaviors and communicative signals deployed during interruption periods using the ELAN package (Version 5.2, 2018, April 04; Nijmegen: Max Planck Institute for

Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/; see Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). We coded the 15- or 30-second segments of interruption periods, starting from the moment the experimenter's hand released the hose, and annotated whether subjects attempted to reengage their partner (or not) and all behaviors and signals produced during the interruption. Reengagement attempts were identified as present if behaviors or signals were produced during the interruption period to either:

a) Attract the attention of the experimenter if inattentive (e.g., audible behaviors or signals, subject moving into the visual field of the experimenter);
b) Behaviors to encourage the experimenter to reinstate the play interaction, defined as "game-related behaviors" containing some information about the game movements or objects that were part of the game (including handing back the hose, rapid prompting movements with the hose, simulating the game action, touching the experimenter with the hose, dropping the hose outside the case, see ESM S3.2);
c) Production of signals to encourage the experimenter to play again, including vocalizations (e.g., laughter), gestures (i.e., see ESM S3.2 for detailed definition of gestures) and facial expressions (e.g., play-face). Gestures used to merely get into bodily contact with the experimenter without being directly related to the tug-of-war game were not counted as reengagement attempts when produced alone (e.g., present body or touch experimenter used without additional game-related behaviors, or gaze alternation between partner and game object, were not counted as reengagement signals).

We also annotated the sensory modality of signals used to reengage a partner in the two conditions, for which we present results in the ESM, S3.3. We classified signals or behaviors as *audible* if their production resulted in a sound, as *tactile* if their production resulted in physical contact with the experimenter, and as *visual* if their production only involved through graphical performance without sound or physical contact (e.g., Pika, Liebal, & Tomasello, 2003, Genty et al. 2009).

Coding reliability

We carried out an overall inter-observer reliability test for the type of signals and behaviors during interruption periods of 50 trials resulting from all three experiments presented in this study (20 trials for adult chimpanzees from experiment 1, 20 trials for adult bonobos from experiment 2, 10 trials for infant bonobos from experiment 3). The test revealed a good agreement between RH and EG (Cohen's $\kappa = 0.67$).

Statistical analysis

To assess the extent to which a certain behavior, signal or reengagement attempt was produced, we first computed proportions of the event (reengagement, signals, behaviour) produced by an individual in each condition. We tested the effect of condition (two within-individual measures) on the proportion of individuals' a) reengagement attempts, b) the use of game-related behaviors, and c) signals, using Wilcoxon signed-rank tests for repeated measurements of individuals in two related samples.

Experiment 1: The Tug-of-War Game in adult chimpanzees

In our first experiment, we investigated the reaction of sub-adult and adult (henceforth subadults and adults are collapsed in one group and referred to as adults) chimpanzees during interruptions of a social game initiated by a human partner. A previous experiment (Warneken et al., 2006) revealed negative results regarding chimpanzees' capacities to engage in shared activities with a human partner; however, the three chimpanzees participating in this experiment were very young (average 3.75 years), and subjected to quite complex, human-like games requiring complementary roles and a spontaneous understanding of the novel task-mechanisms. To address these limitations, we deployed an alternative game that is less complex than the one used in Warneken et al.'s (2006) experiment, and engaging older individuals (above the age of 9 years). We were interested in investigating whether adult chimpanzees would attempt to reengage reluctant partners to a tug-of-war game, and if so, what kinds of behaviors and communicative signals they would use (i.e., specifically, we were interested whether the signals and behaviors used are referring to the game itself). In addition, we manipulate the attentive state of the experimenter interrupting the game (i.e., attentive vs inattentive), to explore whether the subjects adapt the signals and behaviors to the attentive state of their partner when attempting to reengage. Our research questions were therefore: Do adult chimpanzees attempt to reengage reluctant human partners in a cooperative game? And if so, how do they attempt to reengage (what types of behaviors and signals do they use)?

Methods

Study subjects and site

Our experiment included five chimpanzees (*Pan troglodytes verus*; M_{age} = 19.2 years (SD = 5.76 y); range: 9-23 y; 1 female; 4 males; see ESM S3.1), housed at the zoo La Vallée des Singes, France. The chimpanzees were tested between May and October 2018. At the time of study, the

group consisted of seven chimpanzees, of which two did not participate in the study (two females, aged 10 and 11 years). The facility includes an outside enclosure with a large forest area and climbing structures in grassy areas (7500m^2), and an inside enclosure with enrichment and various climbing structures (220m^2). In stable weather conditions, the group is locked outside in their outdoor enclosure. Food is distributed five to six times a day and includes daily ratios of primate pellets, fruits, and vegetables (and a daily portion of rice). The chimpanzees are also occasionally fed with nuts, meat once a week, and eggs twice a week, and can forage for natural vegetation outside (wild berries, herbaceous vegetation). Fresh water is always available from a source at the building and a stream surrounding the island.

Materials and design

See General Methods section for the experimental design and procedure. There was no need to habituate subjects to the garden hose prior to the experiment, since all subjects already had extensive experience with the hose throughout their life in captivity (i.e., the keepers use the hose daily to clean the cages and the hose often was left inside the cage, and was therefore accessible to the chimpanzees as a toy.) The game was played inside the indoor facility, and through the case mesh, at a location wherever subjects spontaneously engaged with us in the game. Five different persons of differing relationship and degree of familiarity with the chimpanzees acted as experimenters and engaged in tug-of-war games with the subjects. These included one female researcher, RH (interacting rarely with the chimpanzees), one female zoo keeper (working approx. once a month with the chimpanzees), and three male zoo keepers (one working approx. five times a month with the chimpanzees, and two working full-time with chimpanzees). Their participation in a testing session was determined based on their availabilities throughout the days. Testing sessions were filmed by a researcher (RH), who was standing outside the cage at a distance of 1.5m using a Panasonic HC-V770 Camcorder with an external Sennheiser unidirectional microphone placed on a tripod.

To allow for comparison with a previous study involving chimpanzees (Warneken et al., 2006), the interruption periods were aimed at lasting 15s ; however, due to differences in reaction times of experimenters to the researcher' cue to restart the game, the interruption period ultimately lasted M=22.06s; S.D.=7.76s. To allow for a more consistent evaluation of all behaviors and communicative signals between trials, we therefore only looked at behaviors/communicative signals occurring during the first 15s of the interruption period.

We conducted 58 trials overall, with 27 trials in the attentive, aka still-faced condition (M=

5.4 ; S.D.=4.16; see ESM S3.1), and 31 trials in the inattentive, aka back-turned condition (M= 6; S.D. = 8; see ESM S3.1). Six additional trials (three in the still-faced condition; three in the back-turned condition) were excluded due to errors in the behavior of the experimenter.

Results

Do adult chimpanzees attempt to reengage reluctant partners to a cooperative game?

Adult chimpanzees attempted to reengage their reluctant partner to the tug-of-war game with a mean proportion of 69% of all trials, with the first reengagement attempt appearing after an average latency time of M=5.29s; S.D.= 4.43s.

Is there an effect of condition on the likelihood of reengagement?

Adult chimpanzees attempted to reengage their partners 69.67% of the time in the still-faced condition and 68.33% of the time in the back-turned condition, , Z=5.5, p=1 (fig. 11).

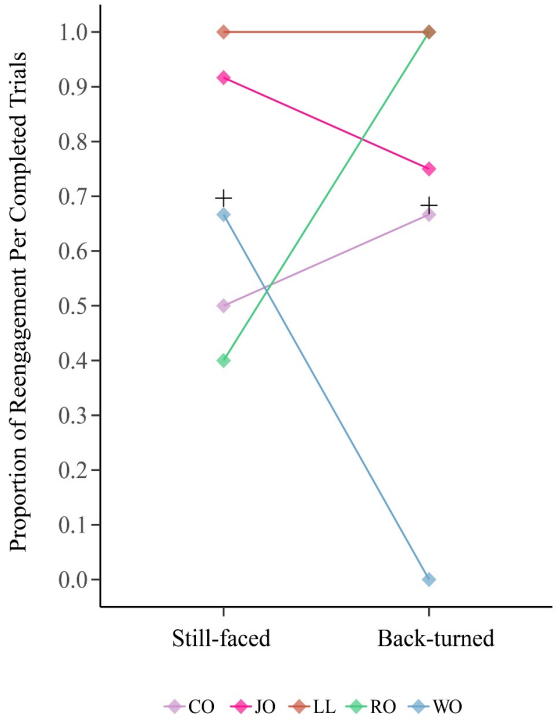


Figure 11 Proportion of reengagement attempts of adult chimpanzee subjects to reengage reluctant partners to the tug-of-war game as per number of completed trials. Each colored data point represents a subject (see ESM S1 for information about subjects). The plus symbol indicates the mean proportion of reengagement attempts across individuals.

What type of behaviors and communicative signals are used to reengage reluctant partners?

Across conditions, chimpanzees reengaged their partners during interruption periods using game-related behaviors (mean proportion 61.4%), gestures (37.2%), a pout face expression (1%), and a *hoo* vocalization (1%), see ESM, S3.2 for definitions. Game-related behaviors consisted in handing back the hose (44.1%), dropping the hose out of cage (15.2%), touching experimenter with hose (12.7%), prompting (17.3%), and simulating game action (10.8%). Gestures used to reengage a partner were shaking the hose (23.4%), stomping (26.4%), touching the hose (31.7%), knocking mesh (6.3%), reaching (3.7%), attempting to touch experimenter (2.4%), head shaking (2.4%), hitting object with object (2.4%), and banging cage (1.2%).

Is there an effect of condition on the likelihood of using game-related behaviors when reengaging a reluctant partner?

There was no statistically significant change in the subjects' proportions of using game-related behaviors between the still-faced and back-turned condition, Z=5.0, p=1; adult chimpanzees used game-related behaviors with a mean proportion of 53.0% in the still-faced condition and 47.0% in the back-turned condition.

Is there an effect of condition on the likelihood of using signals when reengaging a reluctant partner?

There was no statistically significant change in the subjects' proportions of using signals between the still-faced and back-turned condition, Z=5.0, p=0.4; adult chimpanzees used signals with a mean proportion of 59.7% in the still-faced condition and 40.3% in the back-turned condition.

Discussion

In this experiment, adult chimpanzees demonstrated the ability to engage with human partners in a triadic tug-of-war game. During interruption periods initiated by the human partner, the chimpanzees also attempted to reengage their partner in the game in about two-third of all the interruptions. Besides gestural, the chimpanzees produced a variety of game-related behaviors, notably handing back the hose, simulating the game-action of pulling and relaxing, and touching the experimenter with the hose. The likelihood to reengage and to produce game-related behaviors and signals were not affected by the attentive state of the reluctant partner.

Our results contradict previous findings (Warneken et al., 2006; Warneken & Tomasello, 2006), where chimpanzees performed poorly never attempted to reengage their partner (Warneken

et al., 2006). The conclusion drawn from these findings were that chimpanzees, or great apes generally, lack the critical socio-cognitive ability and motivation to engage in cooperative abilities involving joint attention and joint commitment (Call, 2009; Tomasello et al., 2005; Tomasello & Moll, 2010), and therefore do not engage in shared intentionality and do not perform shared goals (Tomasello, 2016).

Nonetheless, our study is not the first to reveal such results concerning the ability to engage in social games; in fact, several other studies provide evidence that at least three ape species - chimpanzees (MacLean & Hare, 2013), bonobos (MacLean & Hare, 2013; Pika & Zuberbühler, 2008), and orangutans (Gruber, 2014) - are motivated to play with a human partner and to reengage partners who suddenly interrupt play. But our study additionally addresses points of concern in these studies; first, the games deployed in these studies often lacked prolonged episodes of reciprocity or were potentially ritualized rather than spontaneous (MacLean & Hare, 2013), so that it is unclear whether a form of commitment could have been reached between the partners. Moreover, some games were rather unilateral with the human partner taking a more active role, for instance when the human partner tickles the ape with a stick (Gruber, 2014; MacLean & Hare, 2013). The nature of the game renders it difficult to gauge whether the intention for reengagement from the subject was personal and selfish (e.g., making the partner come back to satisfy personal intention to being tickled), or shared (i.e., wanting to reestablish the joint activity to tickle each other).

Last, and perhaps most importantly, our subject were all adult chimpanzees (average age 19.2 years) and the chimpanzee subjects in previous studies were fairly young; for instance, the chimpanzees in the study of Warneken et al. (2006) were on average 3.75 years of age, and the chimpanzees in the study of MacLean and Hare (2013) were on average 4.2 years of age. Although younger chimpanzees have shown to attempt to reengage when games are less complex (MacLean & Hare, 2013), social experience in the subjects' ontogeny has never been addressed; we argue, however, that an individuals' social experience with partners (Baarendse, Counotte, O'Donnell, & Vanderschuren, 2013; Fröhlich et al., 2017) could represent a crucial aspect worth of study when testing this joint commitment in great apes. We tested the joint commitment understanding during the tug-of-war game at two different age stages (infants and adults) as well as in another great ape species known to be more cooperative (Hare et al., 2007; Tan et al., 2017; Tan & Hare, 2013) – the bonobo. We represent our findings commencing with joint commitment understanding in adult

bonobos in Experiment 2 and then compare those findings to those gained from infant bonobos in Experiment 3.

Experiment 2: The Tug-of-War Game in adult bonobos

In our second experiment, we investigated the reaction of sub-adult and adult (henceforth subsumed as adult) bonobos during interruptions of the tug-of-war game initiated by the partner. We addressed the same questions as in our first experiment, namely do adult bonobos attempt to reengage reluctant human partners to a cooperative game? And if so, how do they attempt to reengage their partner (type of behaviors and signals)?

Methods

Study subjects and site

Our experiment included thirteen bonobos (*Pan paniscus*; average age M=13.54; S.D.= 8.16 years; age range: 6-30 years; 9 females; 4 males; see ESM S3.1), housed at the Lola Ya Bonobo sanctuary in DRC, where the experiment lasted from October 2018 until November 2018. The bonobos lived in three different social groups in 8-15 ha enclosures, consisting of streams, swamps, lakes, primary rainforest, and grassy open areas. They are provided with food by caregivers in four feedings per day (6-9 types of vegetables per day at 9am, 2-4 types of fruits per day at 11am and again at 4pm, and protein balls at 2pm), but can also range freely within their enclosures to forage for herbaceous vegetation and wild fruits. Water is available from streams, lakes, and sources (provided throughout the day). During the night, the groups are held in 75m^2 dormitories furnished with hammocks.

Materials and design

See General Methods section for the experimental design and procedure. Since it was not certain whether the bonobos had previous experience in touching or seeing a garden hose prior to the experiment, we habituated the bonobos to the hose by distributing 2-3 garden hoses identical to those used in the experiment in all enclosures one week before the start of the experiment. The subjects were tested in their indoor and outdoor facility, wherever the experimenter and the subjects could play the game through the cage mesh. Four different persons of differing relationship and familiarity towards the subjects acted as experimenters and engaged in tug-of-war games with the subjects. These included two female researchers, RH and EG, (interacting rarely with the bonobos), and two male keepers (working full-time with the bonobos). To allow for comparison with a previous study involving bonobos (MacLean & Hare, 2013), the interruption

periods lasted 30 s. We therefore looked at all behaviors communicative signals occurring during the 30 s of the interruption period.

We conducted 54 trials overall, with 32 trials in the still-faced condition (average trials per subject M= 2.9; S.D.=2.51; see ESM S3.1), and 22 trials in the back-turned condition (average trials per subject M=2; S.D.=1.84; see ESM S3.1). Three additional trials (one in the still-faced condition; two in the back-turned condition) were excluded due to errors in the behavior of the experimenter or lack of engagement in the tug-of-war game on behalf of the subject. For the statistical analyses, we had to exclude 4 subjects (reducing the dataset to n=9 bonobos), since they only completed trials in one of the two conditions.

Results

Do adult bonobos attempt to reengage reluctant partners to a cooperative game?

Adult bonobos attempted to reengage their reluctant partner to the tug-of-war game with a mean proportion of 55.29% of all trials, with the first reengagement attempt appearing after an average latency time of M=14.15s; S.D.=22.19s.

Is there an effect of condition on the likelihood of reengagement?

Adult bonobos seemed to attempt to reengage less in the back-turned condition (mean proportion 44.59%) compared to the still-faced condition (65.98%, see fig. 12), however there was no statistically significant change in the subjects' proportions of reengagement attempts between the conditions, Z=22.0, p=0.2.

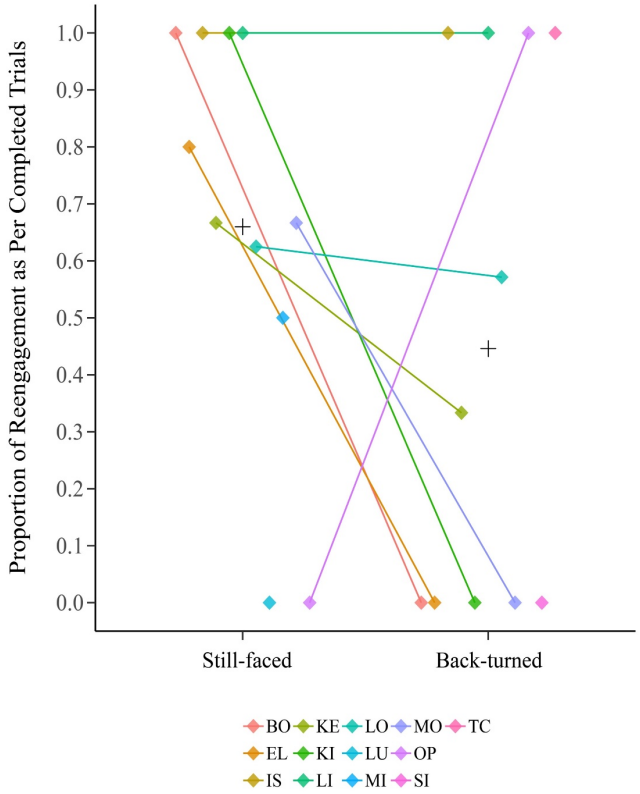


Figure 12 Proportion of reengagement attempts of adult bonobos to reengage reluctant partners to the tug-of-war game as per number of completed trials. Each coloured data point represents a subject (see ESM S3.1 for information about subjects). The plus symbol indicates the mean proportion of reengagement attempts across individuals. Dots without connecting lines indicate cases where individuals had only participated in one of the two conditions.

What type of behaviors and communicative signals are used to reengage reluctant partners?

Across conditions, adult bonobos reengaged their partners during interruption periods using game-related behaviors (mean proportion 70.7%), gestures (26.0%), facial expressions (2.5%), and vocalizations (0.8%). Game-related behaviors constituted hand back hose (75.4%), prompt (9.5%), simulate game action (6.1%), drop hose out of cage (6.5%), and touch experimenter with hose (2.5%). Gestures used to reengage a partner were shake hose (53.1%), shake hand (13.3%), present sex (13.3%), hit object with object (9.6%), touch hose (6.6%), shake head (2.7%), and present body part (1.5%). Vocalizations included a *peep* vocalization and in one instance laughter, and facial expressions included four times the play-face (see ESM, S3.2 for definitions).

Is there an effect of condition on the likelihood of using game-related behaviors when reengaging a reluctant partner?

Adult bonobos seemed to produce similar amounts of game-related behaviors in the back-turned condition (mean proportion 49.7%) compared to the still-faced condition (50.3%), and tests

showed that indeed there was no statistically significant change in the subjects' proportions of using game-related behaviors between the conditions, Z=7, p=0.59.

Is there an effect of condition on the likelihood of using signals when reengaging a reluctant partner?

Adult bonobos produced much less signals in the back-turned condition (mean proportion 16.9%) compared to the still-faced condition (83.1%), yet no statistical test could be run due to the limitations of our sample size (only two signals were used by two subjects in the back-turned condition).

Discussion

In this experiment, adult bonobos demonstrated the ability to engage with human partners in a triadic game, and attempted to reengage their partner back to the game at least in more than half of all the interruptions. Bonobos, like the chimpanzees tested in Experiment 1, most frequently attempted to reengage partners in both conditions by using game-related behaviors as compared to signals. Such game-related behaviors included handing back the hose, simulating the game-action of pulling and relaxing, and touching the experimenter with the hose.

These results further confirm the assumption that joint commitment understanding is probably not uniquely human (see also Pika & Zuberbühler, 2008), and support the assumption that some of the cognitive prerequisites needed for shared intentionality may have been present in our last common ancestor with bonobos and chimpanzees. Our design (i.e. leaving the object accessible for the subjects during interruption periods) allowed for testing whether the apes would use the game object (garden hose) when attempting to reengage a partner rather than gestures or other signals. The apes in our experiment especially produced game-related behaviors when reengaging a partner, constituting a more convincing finding with respect to the apes' intentions when communicating for reengagement. Other studies mainly provided evidence for reengagement through gestures or facial expressions (e.g., Gruber, 2014; MacLean & Hare, 2013; e.g., Pika & Zuberbühler, 2008), but our study is the first to report reengagement attempts directly related to the game action, for instance through the handing back of the hose and simulating the game action. The apes' performance in our experiments therefore does not seem to rely on mere intentions to get back into body contact with the experimenter, but rather quite resembles the performance of human children when engaging in triadic games, who attempt to reengage a passive partner by

offering back the toy or verbally teaching the experimenter how the game was played (e.g., Gräfenhain et al., 2009).

Knowing that adult bonobos and chimpanzees are readily motivated to reengage their partners when dropping from a social game, we now turn to our last Experiment (Experiment 3), where we address a potential effect of age on the ability to recognize broken joint commitments.

Experiment 3: The Tug-of-War Game in infant bonobos

The goal of our third experiment was to investigate the understanding of joint commitment in infant bonobos by testing their reaction to a social partner's sudden breakage from the ongoing tug-of-war game. Our research questions were identical to those of Experiment 1 and 2: do infant bonobos attempt to reengage reluctant partners to a cooperative game? If so, how do infant bonobos attempt to reengage a reluctant partner to a cooperative game (type of behaviors and signals)?

Methods

Study subjects and site

Our experiment, conducted from October 2018 until November 2018, included five infant bonobos (average age M= 3; S.D.=1 years; range= 2.5-5 years; 2 females; 3 males, see ESM S3.1), housed at the Lola Ya Bonobo sanctuary in DRC. The infants were orphans, victims of the bushmeat trade and confiscated by authorities. They had been cared for by humans from the moment of their arrival at the sanctuary. Each infant was guided by a human substitute mother (4 in total) who cared for them for a few years before they were introduced into an existing social group. During the day, infants are free to range in an outdoor enclosure of approximately 500m^2, comprising a forested patch that offers climbing opportunities. In addition, the enclosure is equipped with climbing structures, ropes, a pool and a trampoline. The subjects are free to interact together, or with the human substitute mothers who were present in the enclosure. The substitute mothers provide individual care, social interactions, food and emotional support to the infants. Four substitute mothers were present most of the days (never less than two mothers were present on a single day), and if they were stayed with the infants from 8 a.m. to 5 p.m. The orphans receive a bottle consisting of a mixture of cow milk and cereals with water twice a day, and additionally receive fruits, sugar cane, peanuts and vegetables; each individual receives 5 l of water per day. On rainy days, an indoor enclosure is available (approximately 150 m^2), provided with climbing

structures and ropes and the testing isolation cage (5 m^2). At night they slept in the indoor enclosure (2 individuals per cage) within large cages (2 m^2) furnished with hammocks.

Materials and design

See General Methods section for the experimental design and procedure. Subjects were habituated to the hose prior to the experiment, by placing several plastic hoses in their outdoor enclosure for one week prior to the study. The subjects were tested in the isolation cage of their indoor enclosure (fig. 10). The hose was placed in the cage before the subject and experimenter arrival. The subjects were carried to the testing cage by one of the four human substitute mothers who entered the cage with them. Once the pair entered the testing cage, its door was locked and remained closed throughout the session. Each testing session was stopped after a maximum of 15 minutes even if the trials had not been completed, in order to avoid a decrease of motivation and concentration. The other four subjects remained in their outside enclosure under the supervision of the other substitute mothers. Because of their emotional attachment to and dependence on to these substitute mothers, it was not possible to test the infants in their absence. The substitute mothers therefore acted as experimenters for this study. For each session and each subject, the experimenters were selected depending on their availabilities during the day, thus subjects were assigned to an experimenter based on the most convenient schedule. On a testing day, subjects were called into the testing area and were brought in with one of the substitute mothers, one after the other in random order. The experimenter represented the assigned infant's substitute mother. As in Experiment 2 for adult bonobos and following a previous experiment with bonobos in a similar paradigm (MacLean & Hare, 2013), the interruption periods lasted 30 s. We therefore looked at all behaviors/communicative signals occurring during the 30 s of the interruption period.

We conducted 29 trials overall, with 17 trials in the still-faced condition (average trials per subject M=3.4; S.D.=2.3; see ESM S3.1), and 12 trials in the back-turned condition (average trials per subject M=2; S.D.=1; see ESM S3.1). One additional trial (still-faced condition) was excluded due to the accidental premature ending of the trial before 30s.

Results

Do infant bonobos attempt to reengage reluctant partners to a cooperative game?

Infant bonobos attempted to reengage their reluctant partner to the tug-of-war game with a mean proportion of 27.62% of all trials, with the first reengagement attempt appearing after an average latency time of M=8.57s; S.D.=7.48s.

Is there an effect of condition on the likelihood of reengagement?

There was no statistically significant change in the subjects' proportions of reengagement attempts between the still-faced and back-turned condition, Z=4, p=0.790; infant bonobos attempted to reengage their partners with a mean proportion of 30.24% in the still-faced condition and 25% in the back-turned condition (fig. 13).

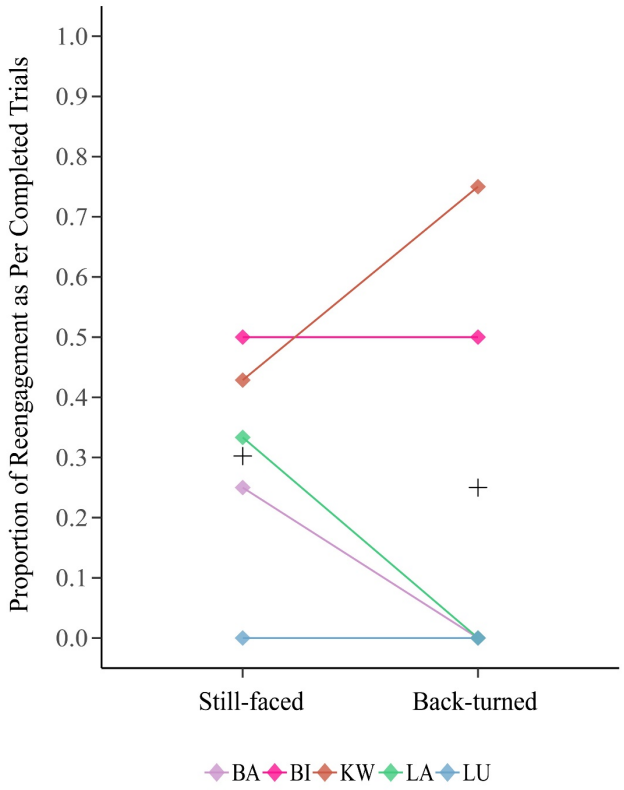


Figure 13 Proportion of attempts of infant bonobo subjects to reengage reluctant partners in the tug-of-war game as per number of completed trials. Each coloured data point represents a subject (see ESM S1 for information about subjects). The plus symbol indicates the mean proportion of reengagement attempts across individuals.

What type of behaviors and communicative signals are used to reengage reluctant partners?

Across conditions, infant bonobos reengaged their partners during interruption periods using gestures (mean proportion 85.5%), play faces (11.8%), game-related behaviors (2.6%), but never vocalizations. The gesture types used to reengage a partner were grab experimenter (27.7%), shake hose (15.4%), dangle (13.8%), shake experimenter (13.8%), reach (7.7%), touch experimenter (7.7%), present body (6.2%), shake head (1.5%), hit object (1.5%), hit object with object (1.5%), jump (1.5%), and move hose (1.5%). Game-related behaviors were only used

twice, produced by one individual ("KW", male, 2.5 years) and constituted once the hand back hose and once the touch experimenter with hose.

Is there an effect of condition on the likelihood of using game-related behaviors when reengaging a reluctant partner?

One of the infant bonobos produced only two game-related signals in the still-faced condition, not in the back-turned condition; due to the insufficient sample size on game-related behaviors in infants, we could not apply statistical tests.

Is there an effect of condition on the likelihood of using signals when reengaging a reluctant partner?

There was no statistically significant change in the subjects' proportions of using signals between the still-faced and back-turned condition, $Z=0$, $p=0.5$; infant bonobos used signals with a mean proportion of 38.8% in the still-faced condition and 61.2% in the back-turned condition.

Discussion

In this study, infant bonobos engaged in a triadic social game with human partners, by pulling at garden hose simultaneously and thereby establishing a form of tug-of-war game. During interruption periods initiated by the experimenter, subjects rarely attempted to reengage their partners (slightly more than 1/5 of the time – 27.62%), which was much less compared to adult bonobos and chimpanzees (average of 62.15 %). When they attempted to reengage their partner they mainly used tactile gestures (but very rarely game-related behaviors). Our results are in line with previous research (MacLean & Hare, 2013; Pika & Zuberbühler, 2008), showing that if conditions are right even very young apes can reengage a recalcitrant partner after having established a shared activity. Although the reengagement rates of infant bonobos were low, our findings are noticeably different from the previous findings by Warneken et al. (2006), where the young chimpanzees never attempted to reengage a partner.

However, when comparing the proportion of reengagement attempts between adults and infants bonobos, adults were more likely to reengage partners and to use behaviors that were relevant to the game. It is difficult to identify whether the infants reengaged the partner to the game after all, since they mainly used gestures that did not involve the hose. The intention behind these gestural reengagement attempts may equally have been to get into body contact with the passive partner, possibly due to insecurity or confusion. Infant bonobos have a very long development period and stay dependent on their mother until approximately 4-5 years of age (Furuichi et al.,

1998). It is possible that, similarly to human infants, the understanding of joint commitment underlying this behavior needs to develop in ontogeny as individuals become more and more independent from their mother, start interacting more and more with different partners and become more aware of how to interact with partners depending on their social status and interaction history (e.g., Fröhlich et al., 2017; Fröhlich, Wittig, & Pika, 2019; Hobaiter & Byrne, 2011a). And indeed, human children actually become skillful engaging in cooperative activities – and recognizing joint commitment – only by the age of around three years (Eckerman & Peterman, 2001; Gräfenhain et al., 2013; Hamann et al., 2012; Kachel & Tomasello, 2019). This may explain why the previous study (Warneken et al., 2006) has found negative results when testing individuals below the age of 4 years; the conclusion drawn from this study, that the cognitive prerequisites for shared intentionality are lacking in great apes and are only present in humans, may not hold if social experience – rather than a special cognitive adaptation - is the critical force behind the development of joint commitment understanding.

General Discussion

In this study, we assessed great apes' understanding of joint commitment by testing whether adult chimpanzees and bonobos as well as infant bonobos would attempt to reengage a reluctant human partner during interruption of a social game. Subjects displayed the motivation to engage in triadic social games with a human partner, sometimes even by initiating the game themselves, and produced behaviors and signals to encourage a reluctant partner to reinstate the game when interrupted; yet, infant bonobos attempted to reengaged partners less often than adult subjects. Additionally, when attempting to reengage their partner, adult subjects used more game-related behaviors compared to infants, who relied mainly on gestural communication, potentially suggesting a better understanding of the concept and the joint nature of the game with each partner having their individual role to fulfill. Previous studies had tested mainly young individuals (MacLean & Hare, 2013; Pika & Zuberbühler, 2008), been largely descriptive (Pika & Zuberbühler, 2008), or involved games lacking aspects of sharedness or reciprocity (Gruber, 2014; MacLean & Hare, 2013). However, in line with some previous studies, findings altogether contradict the results from Warneken et al., (2006), where chimpanzees failed to show any motivation or skill to engage in shared activities, and therefore rebut the conclusion that apes lack the motivation to engage in cooperative activities with others and do not engage in shared intentionality (e.g., Tomasello, 2014). Our findings are in line with other research on chimpanzees and other ape species showing that when using more species-adapted games, apes can engage in

triadic games with a human partner and attempt to reengage their partners when the game is interrupted (Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008).

Although our sample size did not allow for more systematic comparisons and statistical evaluation of the effect of age, we demonstrated higher reengagement rates and greater use of game-related behaviors for reengagement attempts of adults compared to infants, suggesting that a more sophisticated understanding of joint commitment might also emerge later in the ontogeny of apes (though we can only make this assumption for bonobos), along with social experience, which might provide a potential explanation for the negative results of a previous study (Warneken et al., 2006). In human infants, an understanding of the obligations linked to joint commitment does not emerge fully until around three years of age, as several studies have previously shown using experimental paradigms similar to ours (Gräfenhain et al., 2009, 2013; Hamann et al., 2012). The development of joint commitment understanding in human children emerges in concert with their ability to engage in more coordinated and less ritualized joint actions (Eckerman, Davis, & Didow, 1989; Eckerman & Peterman, 2001; Eckerman & Whitehead, 1999), and we argue that similar possible trajectories may as well exist for great ape joint action coordination capacities. Future research with a larger sample size and a longitudinal approach would be necessary to confirm this developmental aspect of joint commitment understanding. An additional exciting future avenue would be to identify the kind of experiences that allow individuals to understand and recognize the consequences of one partner dropping from an ongoing joint action without acknowledgement. Moreover, our results extend previous findings, demonstrating that apes are well capable of producing game-related behaviors when reengaging recalcitrant partner to the game (MacLean & Hare, 2013). For example, subjects produced helpful behaviors for their partners indicating possible game-related intentions, such as simulating the game-action, dropping the hose out of the cage, handing back the hose through the mesh, touching the experimenter with the hose, or prompting the experimenter by rapidly moving the hose when gazing at them. This suggests a motivation to reengage a partner to continue the game itself, rather than to engage in a dyadic activity more generally (see also MacLean & Hare, 2013). Helping partners in a joint action fulfill their individual role in order to successfully reach a common goal is additional behavioral evidence of the establishment of shared intentionality (Tomasello, 2014).

One could still argue that the observation of subjects' reengagement attempts do not qualify as sufficient criteria to identify underlying psychological processes (Leavens et al., 2017), and that the issue shall be addressed with control conditions. One may also argue that the

reengagement attempts of subjects in previous studies (Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008) could have been a consequence of the partner's attentiveness while being reluctant; after the partner had broken from the game, the apes may have communicated as they felt addressed by the gaze, or frustrated due to the passive but staring partner. Our control of the experimenter's attentive state (attentive vs inattentive) showed that the apes attempted to reengage to a similar degree independent of whether the experimenter is looking or not, suggesting that their motivation to reengage is not caused by the gaze of the recalcitrant partner, but possibly by joint commitment.

In future studies, it may also be interesting to manipulate the availability of the partner in other ways. For example by applying conditions in which the experimenter is unable to (but presumably willing) to) pursue the joint goal, or unwilling (e.g., Warneken et al., 2012). If subjects were interested in performing joint activities and recognized their partner as an intentional agent (rather than a mindless social tool), they should attempt to reengage an unwilling partner less than a willing but unable one, presumably understanding the causes of breaking the activity. Additionally, one could introduce an individual control condition, where partners play in parallel rather than together, and then compare these reengagement rates to the ones where both play together (e.g., Gräfenhain et al., 2009, 2013); subjects should attempt to reengage a partner more frequently when they were previously playing together (presumably having established a joint commitment) than when they were playing in parallel.

As for a limitation, one may argue that the presence of the person filming the trials might have provided unintended cues revealing that the trial was not over and that the game might be resumed shortly. Another limitation of our experiments (and previous ones) is that great apes were tested when engaging in joint activities with human partners, rather than peers. It is possible that great apes have all the cognitive capacities to engage in shared intentionality, but do not express or develop this quality in their natural environment. To address whether apes express an appreciation of joint commitment in a natural setting, an exciting avenue for future research would be to study how apes cope with interruptions while being involved in a social interaction (e.g., social grooming or play) with another conspecific, and then compare the reengagement rates in the social situation to one where an individual had been interrupted during a solitary activity (Heesen et al., unpublished data). If individuals then attempted to reengage their original partner more after an interruption in the social setting compared to the solitary setting, this would constitute evidence that apes understand commitments beyond ape-human interactions.

Overall, our findings, suggest that great apes are capable of engaging in cooperative triadic games with a human partner and attempt to reengage the reluctant partner during interruptions, when games are less complex. Nonetheless, the evidence that apes engage in joint action is sparse, suggesting that the crucial difference between humans and great apes may be the underlying *motivation* (not necessarily capacity) to share intentions (see also MacLean & Hare, 2013). Another crucial difference in the performance of joint action between humans and non-human animal species may be type of commitments partners engage in; we hypothesize that apes do not lack the motivation to engage in cooperative activities and the ability to understand joint commitment but that, contrarily to humans, the type of commitments they engage in do not endure beyond the interaction itself and do not allow them to engage in long term projects and to form human-like institutions. The present study aimed at revisiting the debate on the presence of shared intentionality in great apes (Tomasello et al., 2005; Tomasello & Moll, 2010). Our combined results suggest that it is timely to either reconsider the dated origin of shared intentionality (see also van Schaik & Burkart, 2018), or to identify the species-specific sub-categories of this specialization.

Ethical statement

We received ethical agreement for this study from the Commission d'Ethique de la Recherche of the University of Neuchâtel (agreement number: 01-FS-2017), the internal ethical committee of La Vallée des Singes, and from the Ministère de la Recherche Scientifique et Technologie de la République Démocratique du Congo (research permit number: MIN.RST/SG/180/020/2018).

Acknowledgements

We thank La Vallée des Singes and its director Emmanuel Le Grelle as well as Le Conservatoire pour la Conservation des Primates and ABC Congo, particularly Fanny Minesi, Raphaël Belais and Claudine André, for allowing access to the study sites and subjects. We are particularly thankful for the support of the animal keepers at La Vallée des Singes (Carole Michelet, Franck Alexieff, Raphaël Béguet, Steven Thomas, Thomas Urbain) and at Lola Ya Bonobo (Alain Miteti Muyamba, Patrick Kinkani Kinkani, Ivonne Vela Ntona, C'arrive Nsimba Kiasi, Niclette Bonyoka Mboyo, Micheline Nzozi Katiayi), who acted as experimenters and helped us implement the trials, and to Héritier Izansone, for his invaluable assistance during the trials. The present research was supported by the Swiss National Science Foundation (Grant No.

CR31I3_166331 awarded to AB and KZ). RH thanks Kölner Gymnasial- und Stiftungsfonds and EG thanks the Service Egalité des chances of the University of Neuchâtel for supporting their field stay in DRC.

Chapter 5. Joint commitment understanding of apes with conspecifics in natural interactions

Raphaela Heesen[1], Adrian Bangerter[1], Jean-Pascal Guéry[5], Katia Iglesias[2], Federico Rossano[3], Klaus Zuberbühler[1,4], Emilie Genty[1]

[1]University of Neuchâtel, [2]University of Applied Sciences and Arts of Western Switzerland, [3]University of California San Diego, [4] University of St Andrews, [5]Zoo La Vallée des Singes, France

Abstract

The psychological foundation of human cooperation is the propensity to form joint commitments with partners during joint actions, to achieve a common goal. Joint commitments have been studied in laboratory experiments, by looking at individual reactions to sudden interruptions of an ongoing social activity by one of the partners. In great apes, these studies have produced contradictive results for a variety of reasons, possibly also due to the fact that subjects interacted with human experimenters. Whether great apes experience joint commitment in spontaneous joint actions with conspecifics, however, has remained untested. Here, we studied bonobos' and chimpanzees' reactions to a large set of naturally occurring interruptions of joint actions. We found that partners overwhelmingly reengaged each other after such interruptions, by continuing the same actions in the activity that were performed before interruptions, and at the same location even after spatial separation of partners Our findings have implications for the assessment of great ape cooperative capacities, and support the hypothesis that non-human primates are able to experience joint commitments.

Key words – Joint action, joint commitment, great apes, social grooming, social play

This manuscript is currently in preparation, and will be submitted shortly after the thesis submission on 30th September 2019.

Introduction

Joint commitment is the driving force behind human joint action, and cooperation more generally (Clark, 2006; Gilbert, 2014), allowing us to complete conversations, engage in disputes, and perform a large range of joint social activities. Joint commitment emerges when partners experience an obligation to adhere to a joint activity until a joint goal has been reached (Clark, 2006; Gilbert, 1999). Failure to perform one's role jeopardizes joint success in response to which humans tend to sanction partners with disproval, disappointment, reprobation, and can lead to other costly social consequences (Brown & Levinson, 1987; Goffman, 1955). Joint actions, in other words, are built on the mutual expectation that each partner is committed to fulfil their part (Clark, 2006; Tomasello & Moll, 2010), a guiding force of what has been termed the human interaction engine (Levinson, 2006), and an essential part of the unique suite of cognitive and behavioural adaptations allowing humans to engage in sophisticated collaborative activities.

Joint commitment understanding is a cognitive ability that can only be studied in animals by observing behavioural evidence of its presence. It reveals itself in people's reactions towards a sudden suspension of an ongoing joint activity, for instance caused by an interrupting third party or other external perturbations (e.g., Bangerter et al., 2010; Mayor & Bangerter, 2015). Suspension can also be caused by internal perturbations, by which partners themselves are responsible for the sudden cessation of the ongoing joint activity (e.g., Gräfenhain et al., 2013). The latter case has been used as the main experimental paradigm to study the reaction of great apes and human children, with a human partner suddenly disengaging from a triadic social game (Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008; Warneken et al., 2006). It has thus been assumed that reengagement attempts of a reluctant partner after an interruption provided evidence for joint commitment understanding (e.g., Gräfenhain et al., 2009, 2013; Tomasello & Carpenter, 2007; Tomasello & Moll, 2010; Warneken et al., 2006).

For great apes, results have been contradictive with some studies showing disengagement following disruption (Warneken et al., 2006) and others active attempts to reengage the (seemingly) reluctant human partner (Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008). The difference may be due to the fact that the games involved different degrees of complexity and reciprocity or due to age differences in subjects. Indeed, human children only start recognizing joint commitment by the age of three (Gräfenhain et al., 2013). It is possible that great apes' recognition of joint commitment is also subject to developmental trajectories and that the

ape subjects in that study were simply too young. Another major drawback of the current primate literature on joint commitment is that ape subjects interacted with human experimenters and hyman artefacts rather than, as in natural joint actions, with conspecifics, suggesting that conclusions about great ape joint commitment capacities are most likely premature.

In this study, we addressed these issues by focussing on dyadic social interactions between group members engaged in different joint activities, namely social grooming and social play. By focussing on natural joint actions, we were able to obtain a much larger dataset than normally available in experimental studies, which allowed us to test novel hypotheses.

First, we described the frequencies and durations of interruptions depending on the type of joint activity, i.e., social play and grooming, and the species, i.e. chimpanzees and bonobos. For both joint activities, we also catalogued the types of disturbances that led to interruptions. We then looked at the likelihood by which the two species reinstate activity with their previous partners after the disturbances that were caused by external events (i.e., not by the apes themselves). Comparing bonobos and chimpanzees seems relevant to study the effects of social organisation types on the propensity to value joint commitment (Furuichi, 2011; Watts et al., 2006). Indeed, although both chimpanzees and bonobos are both capable of tackling cooperative tasks relying on joint efforts (Hare et al., 2007; Melis et al., 2006a), bonobos have higher tolerance levels than chimpanzees (Hare et al., 2007), share food frequently, act pro-socially towards out-group members (Tan et al., 2017; Tan & Hare, 2013), do not engage in inter-group killing (Gruber & Clay, 2016) and show high levels of empathy towards group members (Clay & de Waal, 2013, 2015). We predicted that bonobos therefore would exhibit a higher likelihood of reinstatement after interruptions than chimpanzees.

However, the mere reengagement rate of partners may be enough to uncover the apes' understanding of their own role and contribution during the activity. Therefore, following the occurrence of an interruption, we evaluated the likelihood of partners to reengage in the activity with their original partner and when did so, whether individuals continued the same activity where they left it before the interruption, i.e. grooming the same body part or resuming the same play type, rather than an alternative body part or play type. Additionally, we looked at whether following interruptions, partners continued their activities at the same location in their enclosure. By assessing this, we wanted to see whether we could distinguish reengagement of a suspended activity from initiations of a new activity, and thus demonstrate natural behavioural evidence for

the understanding of joint commitment. In humans, The awareness of one's own responsibility and a sense of continuation, is considered a crucial aspect of human cooperative action, because it allows people to maintain the integrity of their joint action (Chevalley & Bangerter, 2010; Mayor & Bangerter, 2015). For example, asking for help to collaboratively reconstruct a topic is a common way of reinstating conversation in humans, with common questions asked by listeners, such as "you were saying?", or common phrases to continue the conversational topic by speakers, such as "I resume […], where was I […]" (Bangerter et al., 2010, p. 6).

Lastly, we tested whether the propensity to reengage partners in the suspended activity and the likelihood of continuing the activity (by performing the same action and at the same location than before the interruption) depended on the relationship between interacting partners. We wondered whether partners valued their joint commitment differently depending on the social bond of partners or their rank difference. In human joint action, interruptions pose social threats to people's face (Chevalley & Bangerter, 2010), as their joint commitment to achieve their joint goal is at stake; one way humans deal with this face threat is to deploy politeness (Brown & Levinson, 1987). Indeed, the longer suspensions of interactions last, the more politeness people deploy when reengaging their partners back to the activity (Chevalley & Bangerter, 2010). The use of politeness within speech depends on social distance between interlocutors, their power difference and absolute social ranking within their cultural group (Brown & Levinson, 1987). We predicted that, if the apes were sensitive to the socio-affiliational imperatives of joint commitments with different partners, they should increase their efforts to reinstate, and to resume the same action and location when interacting with partners with whom they are less closely bonded (presumably high social distance), and with partners who are higher-ranked compared to themselves.

Our research questions were the following:

- Does the likelihood of reengagement in the activity depend on perturbation severity, species, activity type and partners' social relationship (social bond and rank difference)
- Is the reinstatement of the activity a continuation of the suspended activity or the initiation of a new activity?

Methods

Study species

We collected observational data of two captive groups of bonobos and three captive groups of chimpanzees (total 50 individuals) at four zoological parks. The bonobo groups (n =25

individuals; range= 4-52 years) were observed at San Diego Zoo and at La Vallée des Singes. Data for chimpanzees were taken from three groups (total 25 individuals, range= 5-49) at La Vallée des Singes, La Réserve Africaine de Sigean, and Basel Zoo. See ESM S2.1 for a more detailed composition of each group.

The subjects at La Vallée des Singes lived in large enclosures comprising an outdoor island enclosure with a large forest and climbing structures in grassy areas (8,000m^2 bonobos, 7,500m^2 chimpanzees), and an indoor enclosure with various enrichment and climbing structures (600m^2 bonobos, 220m^2 chimpanzees). The groups received food five to six times a day, including daily ratios of primate pellets, fruits, and vegetables. Occasionally, both species received rice, nuts, meat, and eggs. Individuals could additionally forage for wild berries and herbaceous vegetation in their outdoor enclosure.

The bonobos at the San Diego zoo lived in an enclosure including both an outdoor (560m^2) and an indoor enclosure (356m^2). The outdoor enclosure consisted of numerous climbing structures, a fresh-water stream, and climbing enrichment. The group received diverse food types up to four times a day (primate pellets and cereals, fruits, vegetables, nuts). Occasionally, the group received additional supplementary foods (honey, peanut butter, popcorn, seeds), often combined with enrichment toys.

The chimpanzees at Basel zoo lived in a building comprising of six lodges in the inside enclosure (total 233.3m^2) and two in the outside enclosure (total 477m^2). The lodges included climbing structures, ropes, puzzle boxes and other enrichments, and the subjects could roam between the lodges (two of the lodges were not visible to visitors). The group received food six times a day, including salad, vegetables, fruits, eggs and primate pellets.

The chimpanzees housed at La Réserve Africaine de Sigean lived in a 200m^2 indoor enclosure and a 10,000m^2 outdoor island area. During the day, they were kept on the outdoor island that consisted of natural vegetation, bushes, trees, sand, a surrounding water pond, climbing structures, bridges and shelters. The group was provided with food six times a day, including vegetables, fruits primate pellets and nuts, and chicken and eggs once a week.

All groups had access to fresh water, and could additionally drink from water sources in the outdoor enclosures, such as ponds and streams (except the chimpanzees from Basel zoo).

Data collection

We collected data using continuous focal sampling (Altmann, 1974), maintaining a balanced record of focal samples across individuals. We video-recorded focal individuals (of 15 or 30 minutes, depending on the site and visibility) using two Panasonic HC-V770 Camcorders with two external Sennheiser unidirectional microphones. Overall, we collected 1298.55h of observation (600.25h for bonobos, 698.30h for chimpanzees), with on average 30.00h per bonobo at San Diego Zoo, on average (±SE) 20.64h (±0.52h) per bonobo at La Vallée des Singes, 37.01h (±0.45h) per chimpanzee at La Vallée des Singes, 30.00h per chimpanzee at Basel Zoo, and 18.08h (±0.42h) per chimpanzee at La Réserve Africaine de Sigean.

During focal samples, we recorded data on proximity, socially directed approaches, and social interactions (grooming and play). We collected proximity data using 5-minute scan samples, determining the proximity of focal individuals to nearest neighbours and the presence of group members. We continuously recorded socially directed approaches, determining if a focal individual approached another group member or was approached by one; we recorded approaches only if their outcome was positive (resulting in a social interaction or contact sitting) or neutral (resulting in proximate sitting). We continuously recorded the duration and directionality (who initiated the social interaction) of social grooming and play between a focal individual and a partner. We also recorded the outcome of agonistic interactions *ad libitum,* noting the IDs of involved individuals and assigned winner and loser, or recorded whether the interaction ended undecided without clear winner and loser. We defined wins as behaviours that involved aggressing an individual physically or by showing aggressive behaviours, taking away resources, displacing or chasing an opponent. We defined losses as behaviours that involved fleeing from the opponent, being displaced, showing submissive behaviours such as bared-teeth displays or whimpers, or giving away a resource. If we could not define a clear winner or loser (e.g., both showing aggressive behaviours, chasing each other, hitting or slapping another, sharing the resource upon negotiation), we defined outcomes as undecided.

Data on proximity, approaches, social interactions and conflicts were collected live and recorded on an iPad 6 using individual customised templates designed specifically for this study using FileMaker Pro 15 (v. 15.0.3.305).

Social bond and rank difference

To compute the social bond between partners, we used an inverse proxy referring to the social affinity or social bond strength between partners, the Dyadic Composite Sociality Index (DSI) (Silk et al., 2006, 2013) using the socialindices package in R (Neumann, 2017, accessed via https://github.com/gobbios/socialindices). The values of the DSI scale can range from $0 \rightarrow \infty$ (see Chapter 3 for a detailed description on how DSI was computed). Since the average DSI equals one, DSI values of dyads below one indicate weaker than average bonds, and DSI values of dyads above one indicate stronger than average bonds (Silk et al., 2013). Our DSI computation included the measures of partners' duration of grooming and play, their proximity rates (i.e., having been in arm reach distance) and their approach rates (i.e., focal having approached the partner or the partner having approached the focal). Because the bonobos of San Diego Zoo were not always let out as one coherent group on the same day, we additionally controlled for observation time and co-residency of dyads for this group.

To compute rank differences, we computed Elo-ratings (Albers & de Vries, 2001; Neumann et al., 2011) for individuals of each site using the "EloRating" package in R (Neumann & Kulik, 2014). Elo-ratings rely on the outcomes of agonistic interactions between partners and depend on the sequence in which interactions take place over time. Undecided outcomes were incorporated as a disadvantage to higher-ranked individuals (Neumann et al., 2011). Since Elo-ratings typically reflect traditional ordinal dominance ranks very closely, for simplicity we refer to Elo-ratings as ranks and Elo-rating differences as rank differences (Neumann et al., 2018).

Coding of interruptions

We coded all behaviors and communicative signals using the computer software ELAN (Version 5.2, 2018, April 04; Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/; see Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

We focused on the interruptions of ongoing social interactions of play and grooming. The interruptions lasted from the moment of suspension of the activity until the reinstatement of the activity. To be considered an interruption, the activity had to be suspended for more than 3sec and the attention of one or both partners (measured as gaze direction) had to be distracted away from the social activity, and drawn to a different partner, an external event or to different social or solitary activity (see fig. 14 for a summary of how we coded interruptions). For grooming,

suspensions of the activity were considered as the cessation of all hand or mouth movements performed to groom the partners' body. For play, suspensions corresponded to the cessation of play-related movements (wrestling, chasing) and play-related signals (play-faces, laughter).

Reinstatements were defined as original partners reengaging in the activity, by resuming the grooming of play specific movements, within 2 minutes after suspension. Partners sometimes used communicative signals before the reinstatement of the activity and were considered at communicative attempts to reengage partners.

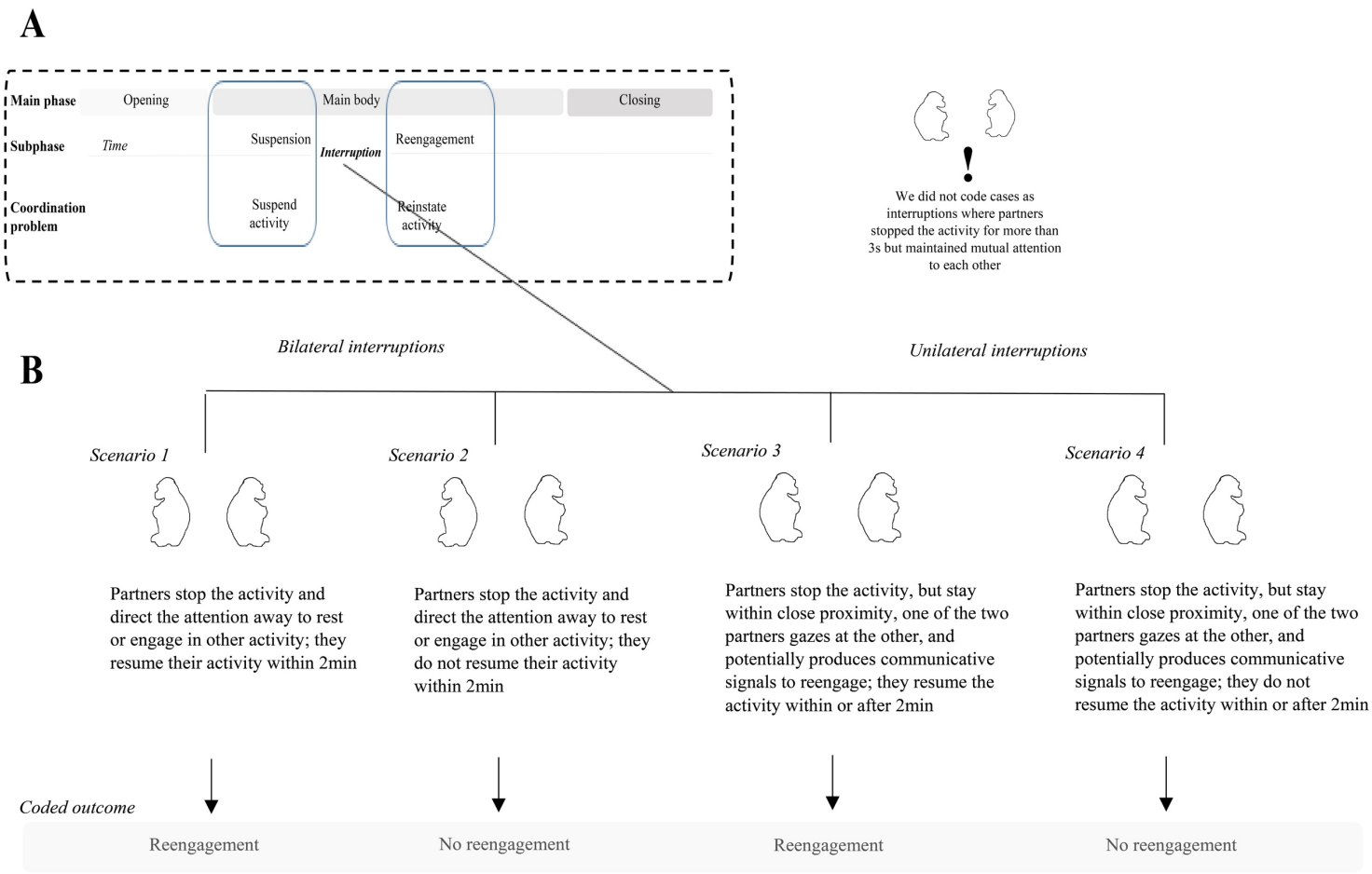


Figure 14 A) The joint action framework after Heesen et al., 2017 (see Chapter 1), showing the temporal placement of interruptions within the overall interaction. B) The coding scheme for interruptions, depicting four scenarios leading to the coding of either reengagement or no reengagement.

We coded our data for the following variables: Partner's IDs, duration of the entire social interaction, duration of the interruption (from onset of suspension of activity movements and attention) to the reinstatement of the activity (from onset of grooming or play typical movements), frequency of interruptions within interactions, types of perturbations stimuli that lead to interruptions, whether reengagement occurred after interruption (yes/no), whether upon reengagement partners continued the same action they were performing at the moment of suspension (yes/no) - for play, this meant resuming the same play-type as before (chase play/contact play), for grooming, this meant continuing to groom the same body part region after the interruption period (see ESM, S4.1 for classifications of body part regions)-, whether one or both partners moved away more than 2m from the original activity location (1,2 or no partner), and when both did, whether the subjects reinstated the activity at the same location in their enclosure than at the moment of suspension (within approximately $2m^2$).

In total, we coded 1,180 randomly selected interactions of social play (329 in total, 169 for bonobos, 160 for chimpanzees), social grooming (810 in total, 239 for bonobos; 571 for chimpanzees), and mixed interactions (i.e., that partners switched between play, grooming, and/or sex; 41 in total, 9 for bonobos and 32 for chimpanzees).

Descriptive statistics for research questions

Frequency and duration of interruptions across activity types and species

To assess the frequency of interruptions we computed the mean rate of interruptions per minute of interaction across play and grooming interactions.

Type and severity of perturbations stimuli

We first described the different type of stimuli that led to the interruption of ongoing interactions and categorized them degree of severity. The degree of severity of perturbations was computed by assessing the proportions by which each type of stimuli caused none or one/both partners to leave the original activity location. Perturbations were categorized as "severe" if their mean proportion of leading to partner(s) leaving exceeded the median proportion of 0.18 (i.e., this median proportion was computed including all proportions where one or both partners had left the original activity location due to the perturbation). We categorized perturbations as "mild" if the mean proportion of perturbations causing partner(s) to leave the initial interaction location was below the median proportion of 0.18.

Likelihood of reengagement

To describe the extent by which subjects reinstated activities, we computed the individual proportions by which dyads reengaged in the original activity after interruptions across activity types and species. To do so, we divided the dyad's frequency of reengagement by the dyad's total number of interruptions.

Likelihood of goal permanence: continuity of activity

To describe the extent by which subjects continued the same activity, we computed the proportions by which dyads continued the same action after interruptions across activity types and species, by dividing the number of times dyads either groomed the same body part or played the same play, by the dyad's total number of reengagement after interruptions.

Likelihood of goal permanence: continuity at same location

To describe the proportions of continuity at the same location, we computed the proportions by which dyads continued their interactions after interruptions at the same location ($2m^2$) across activity types and species, by dividing the number of times dyads continued at the same location by the dyad's total number of reengagement after interruptions.

Statistical analyses

To test our research questions, we fit Bayesian generalized mixed models (*models 1-3*) using the Stan computational framework (http://mc-stan.org/), accessed through the brms package (Bürkner, 2015, 2017) in R v. 3.5.0 (R Core Team, 2013). The model included four MCMC chains, with 10,000 iterations per chain, of which we specified 2,000 iterations as warm-up to ensure sampling calibration, leading to 40.000 posterior samples. The model diagnostics revealed that the posterior distributions reflect the distribution of the original response values appropriately, R-hat statistics were <1.05, the number of effective samples were >100, and the MCMC chains had no divergent transitions (ESM S4.2-4.3). For all models, we used the default priors of the brms package (Bürkner, 2017), which were weakly informative with a student's t-distribution of 3 degrees of freedom and a scale parameter of 10 (ESM S4.2). For inference, we calculated 95% credible intervals from the posterior distributions and checked whether 0 was comprised in this interval.

Due to the large age spectrum of individuals within the sample (i.e., range= 4-52 years, mean =20 years, S.D.=13 years), and the potential effect of sex (i.e., possibly more male-male

bonding in chimpanzees and more female-female bonding in bonobos, see Gruber & Clay, 2016), we investigated whether the model fit improved by including age and sex differences as control variables. We ran LOOIC comparisons (Vehtari et al., 2015) between the full model (including variables age or sex difference) and the reduced model (excluding variables age or sex difference) and found no substantial differences in the expected log predictive density (see ESM S4.4). This indicated that the full models did not improve accuracy, leading us to exclude the variables age and sex difference from the final model.

Finally, we tested the assumptions of the generalized mixed models. Concerning the postulate of the linear relation between the log of the probability of the presence of the event (1) divided by the probability of the absence of the event (0) and the quantitative independent variables (i.e., DSI and rank difference), we transformed the independent variable using a cube root when the assumption was not met. As the results remained similar, we present the results drawn from the original independent variable.

To test for the likelihood of reengagement in the activity depending on perturbation severity, species, activity type and partners' social relationship (social bond and rank difference), we fit a model (*model 1*), with the dependent variable being "reengagement" or "no reengagement" (fitting a Bernoulli distribution with presence or absence of reengagement as binary outcome). The fixed effects were the severity of the perturbation (mild vs. severe), an interaction between the activity type (grooming vs. play) and species (bonobos vs. chimpanzee), and interaction terms between species and DSI (i.e., as a measure of social bond), and species and rank difference (i.e., as a measure of power difference). We z-transformed both DSI and rank difference for all individuals (both species) to mean = 0 and SD = 1 (Schielzeth, 2010). Additionally, we fit cross random effects with unique ID's for activity and interaction dyads (Pinheiro & Bates, 2000). For this dataset, we only considered interruptions occurring during grooming or play, that were caused by external perturbations (total n=1013) – i.e., we excluded internal perturbations that were initiated by subjects themselves ("breaks") from the model. Due to the limited sample size ofinterruptions within mixed activities of bonobos (i.e., ≤10 interruptions), we also excluded "mixed" interactions (i.e., where individuals switched between grooming, play, or sex) (n=40).

To test whether the reinstatement of the activity a continuation of the suspended activity or the initiation of a new activity, we fit two models (*model 2 - 3*). The first one (*model 2*) served to test for the continuity of individuals' actions. We used the same fixed and random effects as for

model 1, but the dependent variable was "continuation of same action" or "no continuation of same action" (fitting a Bernoulli distribution with presence or absence of continuation as binary outcome). For this dataset, we only considered interruptions occurring during grooming or play, which were resumed and caused by external perturbations (total n=729). Due to the limited sample size of interruptions within mixed activities in bonobos (i.e., ≤10 interruptions), we excluded the mixed activity types from the model (n=30).

The second one (*model 3*) served to test for the continuity of the activity at the same location. We used the same fixed and random effects as for model 1, but the dependent variable was "continuation at same location" or "no continuation at same location" of the same action (fitting a Bernoulli distribution with presence or absence of continuation as binary outcome). For this dataset, we only considered interruptions occurring during grooming or play, which were resumed, caused by external perturbations and involved cases where one or both partners left the activity location (n=109). However, we found that there was a small variation between dyads and a small sample of repeated measurements per dyad. The model with the crossed random effects of main body and dyad ID had bad convergence with transient MCMC chains, a very bad pareto-k diagnostic (38% of pareto-k values were >1) (Vehtari et al., 2015), and did not improve accuracy when compared to the reduced model without random effects (see ESM, S4.4). We therefore did not fit random effects for this model.

Results

Frequency and duration of interruptions

Table 3 summarizes the data of interruptions across interactions of different activity types and per species, and fig. 15 provides a summary statistic of the duration of interruptions and interactions across activity types and species. Of the 1,180 interactions, 845 interactions (558 grooming interactions, 250 play interactions and 37 mixed interactions) included 2,704 interruptions (1,681 in grooming, 886 in play, and 137 in mixed interactions), and 335 included no interruptions. Overall, n=156 distinct dyads (80 for bonobos, 76 for chimpanzees) experienced interruptions while interacting, with on average 17.3 interruptions per dyad overall (S.D. = 15.1, range=1-74), 14.1 interruptions on average for bonobos (S.D. = 9.7, range=1-48) and 20.3 interruptions on average for chimpanzees (S.D. = 18.8, range=1-74).

We recorded 1,681 interruptions during grooming interactions (bonobos: 596 interruptions in 190 interactions; chimpanzees: 1,085 interruptions in 368 interactions). The average duration of grooming interactions was 12 min for bonobos (median = 7 min; S.D. = 12 min), and 4 min for chimpanzees (median = 2 min; S.D. = 6min), and the average duration of interruptions of 25 s for bonobos (median = 9 s; S.D. = 41 s) and 20 s for chimpanzees (median = 9 s; S.D. = 29 s). We recorded 886 interruptions of 250 play interactions (bonobos: 549 interruptions in 151 interactions; chimpanzees: 337 interruptions in 99 interactions). The average duration of play interactions was 5 min in bonobos (median = 3 min; S.D. = 5 min), and 2 min for chimpanzees (median = 1 min; S.D. = 4 min), with an average duration of interruptions of 26s for bonobos (median = 14s; S.D. = 32s), and 27s for chimpanzees (median = 15s; S.D. = 32s). Additionally, we recorded 137 interruptions during 37 mixed interactions (bonobos: 14 interruptions in 6 mixed interactions; chimpanzees: 123 interruptions in 31 mixed interactions). The average durations of mixed interactions was 5min for bonobos (median = 4min; S.D. = 4min), and 4min for chimpanzees (median = 3min; S.D. = 4min), and the average duration of interruptions was 26s for bonobos (median = 13s; S.D. = 28s), and 24s for chimpanzees (median = 16s; S.D. = 26s).

Table 3 Summary statistics of interruptions in the interactions (in which interruptions occurred, N=845) of bonobos and chimpanzees.

	Bonobos (*n*= 80 dyads)		**Chimpanzees (*n*= 76 dyads)**			
Interaction type	*N Interactions (mean duration, median, standard deviation)*	*N Interruptions (mean duration, median, standard deviation)*	*N Interactions (mean duration, median, standard deviation)*	*N Interruptions (mean duration, median, standard deviation)*	*Total Interactions*	*Total Interruptions*
Grooming	190 (12 min; median = 7 min; S.D. = 12 min)	596 (25s; median = 9 s; S.D. = 41 s)	368 (4 min; median = 2 min; S.D. = 6min)	1085 (20s, median = 9 s; S.D. = 29 s)	558	1681
Play	151 (5 min, median = 3 min; S.D. = 5 min)	549 (26s, median = 14s; S.D. = 32s)	99 (2 min, median = 1 min; S.D. = 4 min)	337 (27s, median = 15s; S.D. = 32s)	250	886
Mixed	6 (5min, median = 4min; S.D. = 4min)	14 (26s, median = 13s; S.D. = 28s),	31 (4min, median = 3min; S.D. = 4min),	123 (24s, median = 16s; S.D. = 26s)	37	137
Total	**347**	**1159**	**498**	**1545**	**845**	**2704**

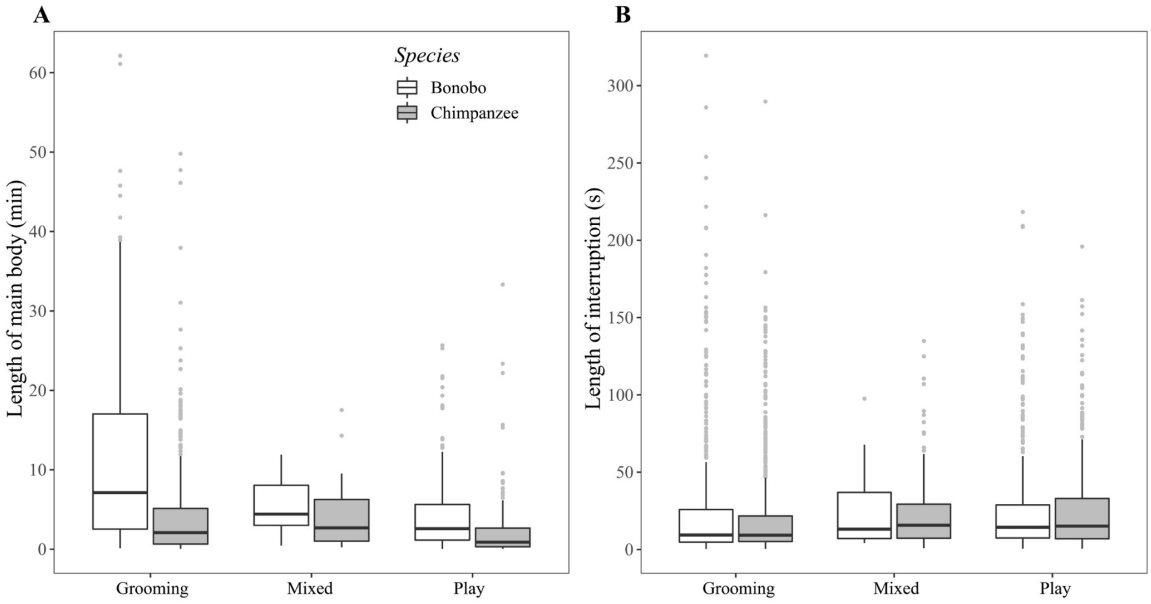


Figure 15 Summary statistics for duration of interactions, aka main bodies (A) and interruptions (B) across activity types and species. The lower and upper hinges correspond to the 25th and 75th percentiles. The upper whiskers extend from the hinge to the largest value no further than 1.5 * the distance between the 25th and 75th percentiles from the hinge (after Wickham, 2016). The lower whisker extends from the hinge to the smallest value at most 1.5 * the distance between the 25th and 75th percentiles from the hinge. Outlier points beyond the end of the whiskers are plotted individually.

Rate of interruptions across activity types and species

The overall mean rate of interruptions across species and activity types was 0.74 interruptions per minute (Table 4). Across species, mixed interactions had the highest interruption rates (0.11), followed by social play (1.04), and grooming the one with the lowest rate of interruptions per minute (0.60). In bonobos, there were on average 0.70 interruptions per minute, including 0.45 during grooming, 0.49 during mixed, and 1.06 during play activities. Chimpanzees exhibited on average 0.76 interruptions per minute, including 0.66 during grooming, 1.28 during mixed, and 1.02 during play activities.

Table 4 **Mean rate**, median rate and inter-quartile range [1st quartile;3rd quartile] of interruption frequency across all interactions (N=1180) per minute of interaction length, grouped by the type of activity and species.

Species	**Type of activity**			
	Grooming (*N*=810)	Mixed (*N*=41)	Play (*N*=329)	*Total mean*
Bonobo;N_{dyads} =80	**0.45**;0.20;[0.04;0.51]	**0.49;**0.23;[0.00;0.50]	**1.06;**0.79;[0.44;1.29]	**0.70;**0.40;[0.10;0.87]
Chimpanzee;N_{dyads} =76	**0.66**;0.36;[0.00;0.86]	**1.28**;1.18;[0.78;1.65]	**1.02**;0.67;[0.00;1.37]	**0.76;**0.46;[0.00;0.99]
Total mean	**0.60;**0.30;[0.00;0.75]	**1.11;**0.99;[0.52;1.50]	**1.04;**0.75;[0.22;1.35]	**0.74;**0.43;[0.00;0.97]

Type and severity of perturbations stimuli

We identified 11 types of perturbations that led to interruptions of activities (Table 5). The most common perturbation was internal (i.e., a "break", initiated by one or both partners, *n* =1648). The second most common perturbation was external (i.e., a conspecific interrupting the activity, n= 388), followed by unknown (i.e., indeterminable for the human observer) perturbations (n=366). Other common external perturbations were food-related events (*n* =95), environmental stimuli (*n* =71), and group vocalizations (*n* =53). Relatively uncommon perturbations were group conflicts (*n* =24), visitors (*n* =21), the presence of keepers (*n* =18), holding-related procedures (*n* =13), and group-travel events (*n* =4).

Following our severity categorization method (Table 5, see Methods section), we identified six perturbations (including one internal and five external) as "mild" (i.e., partners left the activity location in less than 18% of cases): *breaks, conspecific interrupting, unknown events, environmental stimuli, visitors, and the presence of keepers*. We identified five perturbations (all external) as "severe" (i.e., partners left the activity location in more than 18% of cases): *food-related events, group vocalizations, group conflicts, holding-related procedures and group travel*.

Table 5 Perturbations causing interruptions of natural grooming and play activities in bonobos and chimpanzees.

Perturbation	Nature of perturbation	Frequency of event	Proportion of partners' leaving	Severity category[1]
Break	Internal	1648	0.09	Mild
Conspecific interrupting	External	388	0.15	Mild
Unknown	External	366	0.12	Mild
Food-related	External	95	0.37	Severe
Environmental stimulus	External	71	0.10	Mild
Group vocalizations	External	53	0.23	Severe
Group conflict	External	24	0.20	Severe
Visitors	External	21	0.06	Mild
Presence keepers	External	18	0.00	Mild
Holding-related procedure	External	13	0.50	Severe
Group travel	External	4	1.00	Severe
Out of sight	Unknown	3	Unknown	Unknown

[1] Perturbations were categorized as "severe" if the proportion of perturbations causing partner(s) to leave the initial interaction location exceeded the median proportion of 0.18.

Likelihood of reengagement depending on perturbation severity, species and activity type

After external perturbation events, bonobos and chimpanzees reengaged in the social activity with their original interaction-partners by a mean percentage proportion of 70% (Table 6). The likelihood of reengagement was much lower if partners were subjected to a severe perturbation event, compared to a mild perturbation event (fig. 16A; b (SE) = -1.65 (0.25), 95%CrI [-2.15, -1.19]). Being subjected to a severe perturbation event compared to being subjected to a mild perturbation event decreases the chance of reengagement of 81% (Odds Ratio = 0.19).

There was no difference in the likelihood of reengagement between grooming vs. playful activity types (fig. 16A; b (SE) = -0.03 (0.29), 95%CrI [-0.59, 0.54]). However, we found a pronounced relationship between the likelihood of reengagement and species (fig. 16A; b (SE) = -0.85 (0.28), 95%CrI [-1.41, -0.30]). Being a chimpanzee compared to being a bonobo decreases the chance of reengagement, but this decrease was slightly reduced for play (fig. 16B; b (SE) = 0.31 (0.48), 95%CrI [-0.66, 1.26]).

Table 6 **Mean proportion,** median proportion and inter-quartile range [1st quartile;3rd quartile] of reengagement attempts across dyads after interruptions caused by visible external perturbation events (N=1053), grouped by the type of activity and species.

Species	**Type of activity**			
	Grooming ($N_{interruptions}$=640)	Mixed ($N_{interruptions}$ =40)	Play ($N_{interruptions}$=373)	*Total mean*
Bonobo;N_{dyads} = 78	**0.73**;0.80;[0.52;0.90]	**0.92**;1.00;[0.92;1.00]	**0.76**;0.83;[0.68;1.00]	**0.75**;0.80;[0.60;1.00]
Chimpanzee;N_{dyads} =67	**0.67**;0.50;[0.71;1.00]	**0.72**;1.00;[0.50;1.00]	**0.58**;0.50;[0.50;0.84]	**0.65**;0.67;[0.50;1.00]
Total mean	**0.70**;0.78;[0.50;1.00]	**0.78**;1.00;[0.67;1.00]	**0.69**;0.79;[0.50;1.00]	**0.70**;0.79;[0.50;1.00]

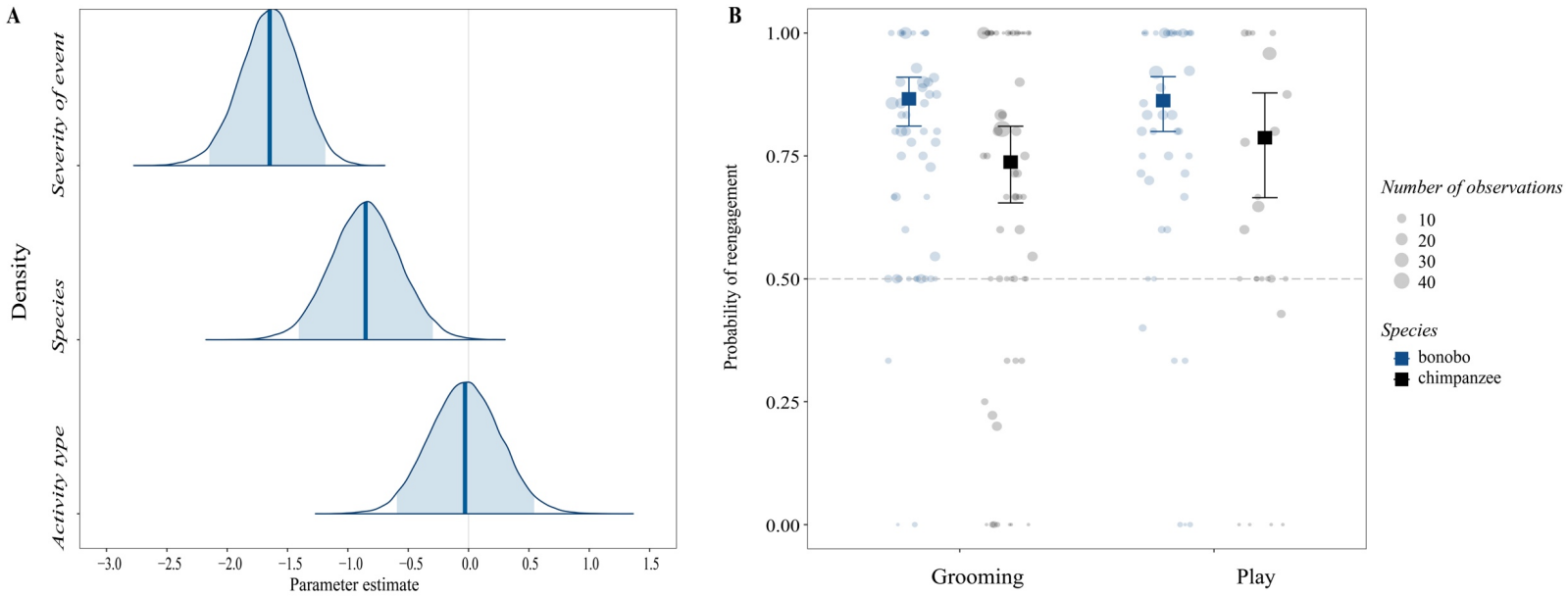


Figure 16 Variation of reinstatement attempts after interruptions across interactions type, species and perturbation event severity (model 1). Plot **A)** presents the posterior distribution for the marginal effects of activity type, species and the severity of the perturbation causing the interruption. The blue area corresponds to the 95% credible interval of the estimated effect. The value of 0 is highlighted by a vertical grey line, and the estimated posterior mean is highlighted by a blue vertical line. Plot **B)** shows how the predicted probability of reengagement for the marginal effects of the interaction term between activity type and species for the Bayesian generalized linear mixed model. It shows how the model fits the actual data. The upper and lower bars correspond to the upper and lower 95% credible intervals, respectively, and the squares represent the posterior mean. Each circle corresponds to the proportion of reengagement per interaction dyad, with the size of the circles corresponding to the number of observations per dyad.

Goal permanence: likelihood of continuity of activity depending on perturbation severity, species and activity type

After external perturbations, bonobos and chimpanzees continued the same action that they had at the time of suspension, with original interaction-partners, by a mean percentage proportion of 72% (Table 7). Partners subjected to a severe perturbation event, compared to a mild perturbation event, were about 65% less likely to continue the same action (fig. 17A; Odds Ratio = 0.35; b (SE) = -1.04 (0.35), 95%CrI [-1.73, -0.37]).

Both species were more likely to continue the same play type compared to grooming the same body part (fig. 17A; b (SE) = 0.97 (0.40), 95%CrI [0.22, 1.78]). There was also a pronounced relationship between the likelihood of continuing the same action and species (fig. 17A; b (SE) = -1.44 (0.38), 95%CrI [-2.19, -0.70]). Being a chimpanzee, compared to being a bonobo, decreases the chance of continuing the same action within grooming activities (fig. 17B), but this decrease slightly reduces for play (fig. 17B; b (SE) = 1.28 (0.73), 95%CrI [-0.11, 2.78]).

Table 7 **Mean proportion,** median proportion and inter-quartile range [1st quartile;3rd quartile] of continuing the same action for resumed interactions, after interruptions caused by visible external perturbations (N=759), grouped by activity type and species.

Species	**Type of activity**			
	Grooming ($N_{interruptions}$=444)	Mixed ($N_{interruptions}$ =30)	Play ($N_{interruptions}$=285)	*Total mean*
Bonobo;N_{dyads} = 74	**0.75**;0.78;[0.67;1.00]	**0.33**;0.17;[0.00;0.50]	**0.83**;1.00;[0.80;1.00]	**0.77**;0.88;[0.67;1.00]
Chimpanzee;N_{dyads} =59	**0.60**;0.60;[0.33;1.00]	**0.45**;0.50;[0.00;0.70]	**0.93**;1.00;[0.93;1.00]	**0.66**;0.86;[0.38;1.00]
Total mean	**0.68**;0.75;[0.50;1.00]	**0.41**;0.42;[0.00;0.70]	**0.86**;1.00;[0.83;1.00]	**0.72**;0.88;[0.50;1.00]

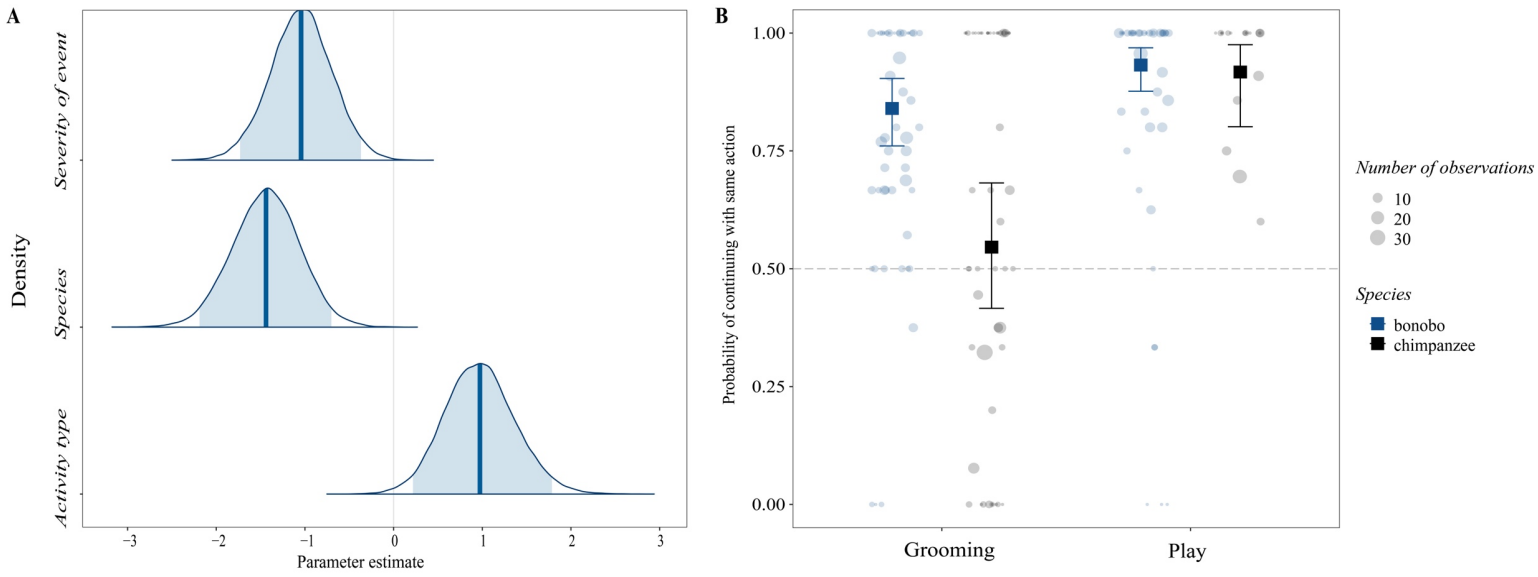


Figure 17 Variation of proportion of continuation of the same action after interruptions of social grooming or play across species and for external perturbations (model 2). Plot **A)** presents the posterior distribution for the marginal effects of activity type, species and the severity of the perturbation causing the interruption. The blue area corresponds to the 95% credible interval of the estimated effect. The value of 0 is highlighted by a vertical grey line, and the estimated posterior mean is highlighted by a blue vertical line. Plot **B)** shows how the predicted probability of continuing the same action for the marginal effects of the interaction term between activity type and species for the Bayesian generalized linear mixed model. It shows how the model fits the actual data. The upper and lower bars correspond to the upper and lower 95% credible intervals, respectively, and the squares represent the posterior mean. Each circle corresponds to the proportion of reengagement per interaction dyad, with the size of the circles corresponding to the number of observations obtained per dyad.

Goal permanence: Likelihood of continuity at same location depending on perturbation severity, species and activity type

After external perturbation events, bonobos and chimpanzees resumed the activity at the same location as the one where they were at the time of suspension (within $2m^2$), with original interaction-partners, by a mean percentage proportion of 63% (Table 8). Partners subjected to a severe perturbation event, compared to a mild perturbation event, were about 71% less likely to continue at the same location (fig. 18A; Odds Ratio = 0.27; b (SE) = -1.32 (0.64), 95%CrI [-2.60, -0.10]).

There was no clear relationship between continuation at the same location and the activity type (fig. 18A; b (SE) = -0.72 (0.62), 95%CrI [-1.96, 0.45]), the species (fig. 18A; b (SE) = -0.31 (0.99), 95%CrI [-2.21, 1.71]), and the interaction term between activity type and species (fig. 18B; b (SE) = -0.79 (1.25), 95%CrI [-3.30, 1.59]).

Table 8 **Mean proportion,** median proportion and inter-quartile range [1st quartile;3rd quartile] of continuing at the same location for resumed of interactions through which partners were separated spatially, after interruptions caused by visible external perturbations (N=115), grouped by activity type and species. Cases "out of sight" were excluded.

Species	**Type of activity**			
	Grooming ($N_{interruptions}$=44)	Mixed ($N_{interruptions}$ =6)	Play ($N_{interruptions}$=65)	*Total mean*
Bonobo;N_{dyads} =41	**0.66**;1.00;[0.25;1.00]	**0.33**;0.00;[0.00;0.50]	**0.65**;1.00;[0.50;1.00]	**0.63**;1.00;[0.00;1.00]
Chimpanzee;N_{dyads} =20	**0.81**;1.00;[1.00;1.00]	**0.00**;0.00;[0.00;0.00]	**0.59**;1.00;[1.00;1.00]	**0.63**;1.00;[0.05;1.00]
Total mean	**0.71**;1.00;[0.46;1.00]	**0.20**;0.00;[0.00;0.00]	**0.64**;1.00;[0.15;1.00]	**0.63**;1.00;[0.00;1.00]

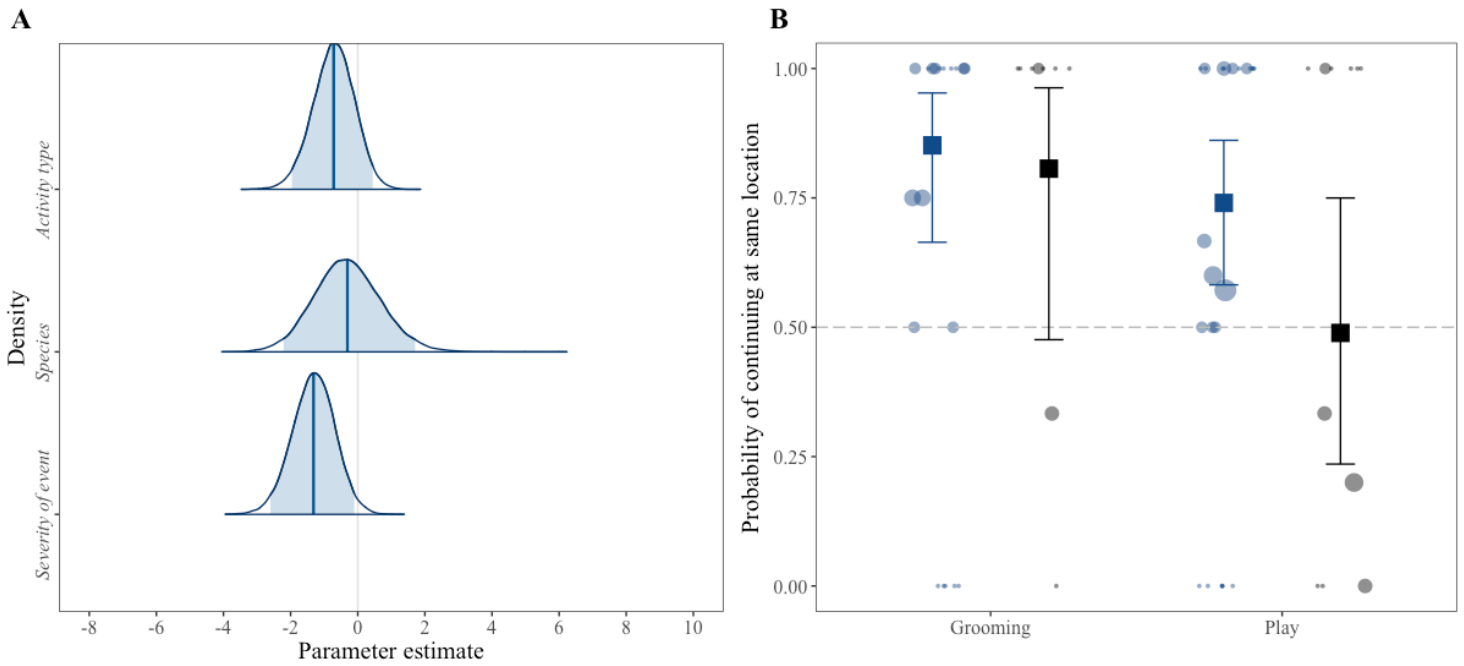


Figure 18 Variation of continuing at the same location after interruptions of social grooming or play across species and for external perturbations (model 3). Plot **A)** presents the posterior distribution for the marginal effects of activity type, species and the severity of the perturbation causing the interruption. The blue area corresponds to the 95% credible interval of the estimated effect. The value of 0 is highlighted by a vertical grey line, and the estimated posterior mean is highlighted by a blue vertical line. Plot **B)** shows how the predicted probability of continuing at the same location for the marginal effects of the interaction term between activity type and species for the Bayesian generalized linear mixed model. It shows how the model fits the actual data. The upper and lower bars correspond to the upper and lower 95% credible intervals, respectively, and the squares represent the posterior mean. Each circle corresponds to the proportion of reengagement per interaction dyad, with the size of the circles corresponding to the number of observations obtained per dyad.

Likelihood of reengagement and goal permanence depending on partners' social bond, rank difference

We found no clear relationships between the likelihood of any of the dependent variables (i.e., reengagement, continuation of same action, continuation at same location) and social bond, and neither between the likelihood of these variables and rank difference, see for a summary, ESM S4.2.

Discussion

We found that interruptions in the natural interactions of chimpanzees and bonobos are extraordinarily common (about one occurs each minute of an interaction bout), and like in human interactions (Bangerter et al., 2010), they may be caused by the individuals themselves (i.e., self-interruption) or due to external perturbations, such as environmental stimuli or other conspecifics. Play was the activity type that included interruptions most often, compared to grooming and mixed activity types (i.e., where partners switch between grooming, play, sex). When these interruptions occurred, both species reinstated the social activities very frequently, and they were more likely to do so if they were interrupted by a mild perturbation (e.g., conspecific interrupting, environmental stimulus) compared to a severe perturbation (e.g., food, conflict). Bonobos were also considerably more likely than chimpanzees to reinstate social activities, across activity contexts. We investigated the likelihood by which subjects resumed previous actions (e.g., for grooming, continuing to groom the same body part; for play, continuing to either play contact play or chase play after the interruption) and the likelihood by which partners resume their activity at the same location if they had been separated. Indeed, we found that both species continue the activity with the same action and at the same location in the majority of cases. Goal permanence, for both the continuation of the same action, as well as the continuation of activities at the same location, was also very high, and increased especially for play (as compared to grooming); yet again, bonobos were more likely compared to chimpanzees to continue the same action (though only for grooming contexts).

These results suggest that partners seem to value the commitment they have established with their partner, to complete a shared activity until their apparent common goal is reached. This effect seems to be stronger for bonobos, than for chimpanzees, which is in line with our previous predictions stating that bonobos – as they are more socially tolerant, less aggressive, and more pro-

social (Demuru et al., 2015; Hare et al., 2007; Herrmann, Hare, et al., 2010; Tan et al., 2017) – should exhibit higher motivational levels of reengagement than chimpanzees. This suggests that the appreciation of one's partner may be may be more expressed in species such as bonobos, which live in less competitive environments, and in more egalitarian societies with lose dominance styles compared to chimpanzees (see Gruber & Clay, 2016 for review).

Previous research has considered reengagement of partners after interruption as evidence for the recognition of joint commitment (Warneken et al., 2006), and many have claimed that apes are incapable of it and that it is a human specific capacity that results from their unmatched cooperative nature and motivation to share experiences and goals with other, otherwise labelled as shared intentionality (Tomasello & Moll, 2010). The findings of this study, along with other previous findings (Clay, Pika, Gruber, & Zuberbühler, 2011; Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008) challenge this claim, however, that only humans are capable of recognizing joint commitments (Warneken et al., 2006), and thus to engage in cooperative activities involving joint commitment (Tomasello & Moll, 2010; Warneken et al., 2006). Our findings also extend previous ones by showing that joint commitment understanding extends to spontaneous social interactions between conspecifics. This is important for our understanding of whether apes exhibit joint commitment with peers, and not only with human partners, as evinced by previous experiments (MacLean & Hare, 2013; Pika & Zuberbühler, 2008). This suggests that the behavior expressed in the artificial social games was not a consequence of a learning by the apes in their captive environment with humans, suggesting that the understanding of joint commitment has a biological basis, and is exhibited in interactions between conspecifics. Our study, however, cannot ascertain to which extend this ability is also expressed in wild apes, yet this constitutes a crucial layer that needs to be addressed in future studies, when one is to draw conclusions about evolutionary trajectories of this ability.

Moreover, the games used in the previous experiments (MacLean & Hare, 2013; Pika & Zuberbühler, 2008; Warneken et al., 2006), mostly were triadic games, with an object as a shared focus of attention between two participants. This is different in our study, as play and grooming are dyadic activities carried out between two partners, yet often without an object being used. MacLean and Hare (2013) showed that apes exhibit joint commitment in both scenarios, in dyadic and triadic games, suggesting that joint commitment is not necessarily related to the triadic nature of the game. This also means that natural activities, just as triadic activities, equally involve a

feeling of being jointly committed – and that joint commitment has probably first emerged in dyadic, more basic activities where joint attention was not necessary (i.e., I can groom your body without you being attentive, and yet we are both committed to grooming, as you hold your body still for me to groom).

The pattern of reengagement did however follow the patterns predicted by politeness theory, which are that politeness efforts (in humans) increase between more weakly bonded partners and partners of differing rank (Brown & Levinson, 1987). The difference therefore may not be quantitative but rather qualitative; in other words, future studies may learn more from looking at the calibration of *signals* produced in apes while suspending or reengaging an activity with different partners, than the overall reengagement rate (since, also, politeness in humans is predominantly expressed via speechacts, Brown, 2015). Nonetheless, and importantly, one could argue that the fact that our apes reengaged equally much, regardless of the relationship to their partner, shows how this commitment is a *priority*, and not just an empathic result of the subjects' fear of repercussion, when interacting with a more dominant or less familiar partner.

Nonetheless, the evidence that partners reengage after interruptions does not allow insights into individuals' understanding of their responsibility during the interaction. The awareness of one's own contribution and a sense of continuation, however, is considered a crucial aspect of human joint action, because it allows people to maintain the integrity of their joint action (Chevalley & Bangerter, 2010; Mayor & Bangerter, 2015). The fact that our apes continued same action and at same location after the interruptions suggests two things. First, it suggests that individuals might be aware about what their contributions were during the interaction, as well as about where their interaction had initially started. Reconstructing a topic is a common way of reinstating conversation in humans, with common questions asked by listeners, such as "you were saying?", or common phrases to continue the conversational topic by speakers, such as "I resume […], where was I […]" (Bangerter et al., 2010, p. 6). Future studies could investigate to which extent individuals are also aware about their roles during the interaction, for instance by studying whether individuals communicate or reengage more if they carried greater responsibilities. Second, the fact that the apes exhibited high levels of goal permanence when resuming interactions suggests that the reinstatement may truly counts as a *continuity* of the suspended activity and not the initiation of a new one. This means that participants have some sense of continuation, and are probably truly motivated by the intention to re-establish a previous commitment with their

previous partners, rather than going back to some previous task to satisfy their personal intention (Trafton, Altmann, Brock, & Mintz, 2003).

The most frequent causes of interruptions were the chimpanzees and bonobos themselves, similar to a previous study on perturbations of collaborative activities in humans (Mayor & Bangerter, 2015). External severe perturbations such as food distribution within the zoo facility, holding-related procedures such as shifting groups between out- and indoor areas, as well as group vocalizations and conflicts often caused partners to leave the activity location. Severe perturbations reduced the likelihood of reengagement and the continuation of previous actions. This may be because the initially "cooperative" attitude during social grooming or play then shifted to a more "competitive" attitude during feeding or conflicts, leading to a decay of memory for initial goals (Altmann & Trafton, 2002) and decline of motivation to cooperate. For example, a study with humans showed that the nature of the interaction context (whether it was competitive or cooperative) influences partners' motivation to cooperate, evinced by the findings that competitive conversations tend to involve shorter pauses between participants' turns and more interruptions than cooperative conversations (Trimboli & Walker, 1984).

The likelihood of reengagement, along with the likelihood of continuing to groom the same body part after interruptions, also differed between species. Chimpanzees were overall less likely to reengage and to continue grooming the same body part of partners compared to bonobos. The finding is supported by a range of studies demonstrating higher tolerance and pro-sociality levels towards conspecifics in bonobos compared to chimpanzees (Hare et al., 2007; Tan et al., 2017; Tan & Hare, 2013; but see Jaeggi, Stevens, & Van Schaik, 2010). Taken together, our findings show that bonobos, more so than chimpanzees, seem to exhibit more profound abilities to navigate joint action with partners. The difference between the species in this study became especially apparent in grooming activities, where bonobos were more likely than chimpanzees to reengage and continue to groom the same body part after interruptions. Social play, in both species, involved the highest levels of reengagement and continuation, which further confirms our hypothesis that social play applies as the interaction context with the highest levels of cooperation among participants (Heesen et al., 2017).

Altogether, bonobos and chimpanzees are readily motivated to reengage conspecifics to a social activity after interruptions, and both species – although more so bonobos –reconstruct their previous action after the interruption. These findings have implications for our understanding on

how “joint” interactions in these species may be, suggesting that individuals are aware about their personal contributions to the ongoing interaction. The motivation to resume unfinished social activities and the understanding of one’s previous interactional contribution help to maintain social integrity and are considered part of joint commitment in humans (Bangerter et al., 2010; Tomasello & Carpenter, 2007; Tomasello & Moll, 2010; Warneken et al., 2006).Thus, our findings likely provide evidence that great apes have some appreciation of joint commitment. The evidence could however further be strengthened in three ways; first, one could establish a more detailed account on how individuals’ calibrate their communication when suspending or reinstating activities according to the social relationships with partners, to inform on the abilities of apes to coordinate suspension, and not only reengagement, as suspension coordination is another predictor for that two partners are dealing with a joint commitment (Chevalley & Bangerter, 2010). Second, a more detailed assessment about individuals’ understanding of their own interaction roles would be useful. One could do this by studying whether individuals invest more efforts when they themselves were responsible for the suspension of the interaction. Third, one could look at the rate by which individuals reinstate solitary activities (e.g., self-grooming) and compare it to the rate by which individuals reinstate social activities. This would constitute a powerful comparison, permitting insights whether individuals’ attempts to reengage are really consequences of joint commitment, or whether they were results of a dissatisfied individual intention to achieve a personal goal (Searle, 1979).

To sum up, our findings suggest that apes have an understanding of joint commitment when engaging in social interactions with conspecifics, and that joint commitment therefore may not be a uniquely human characteristic (Tomasello & Carpenter, 2007). We hope that our research motivates future studies to investigate such patterns in groups of wild apes, in order to extend this knowledge further. Human uniqueness, as it seems, may have been shaped by ancestral primate roots, and much is to be studied and learnt from the natural interactions of apes to strengthen this understanding.

Chapter 6. Joint commitment understanding in bonobos: Relationships, responsibilities and roles

Raphaela Heesen[1], Klaus Zuberbühler[1,4], Katia Iglesias[2], Federico Rossano[3], Adrian Bangerter[1], Jean-Pascal Guéry[5], Emilie Genty[1]

[1]University of Neuchâtel, [2]University of Applied Sciences and Arts of Western Switzerland, [3]University of California San Diego, [4] University of St Andrews, [5]Zoo La Vallée des Singes, France

Abstract

Humans routinely engage in cooperative activities with partners committing to common goals by sharing their intentions. This ability to experience joint intentions has been identified as uniquely human, based on species differences in how humans and great apes reengage reluctant partners after interruptions of joint activities. However, conclusions have remained controversial for both theoretical and methodological reasons, such as apes being asked to interact with human experimenters and their artefacts. Here, we revisit the human uniqueness claim and show that bonobos interacting with conspecifics in natural cooperative activities understand joint commitment. In our experiment, we brought one of the partners to suspend ongoing cooperative grooming, with one of two disruptions: one specifically directed at the partner and another non-specifically directed at the entire group. We found strong propensities to reengage distracted partners to resume the previous cooperative activity. Communicative efforts to both suspend and reengage grooming interactions depended on the partners' bonding and power differences and their individual roles in the previous interaction. We compared these findings with human understanding of individual responsibility in joint commitments and the shared intentionality hypothesis more generally.

Key words – Joint action, joint commitment, face management, bonobos

This manuscript is in preparation, and will be submitted in due course after the submission of the thesis on 30th September 2019.

Introduction

Most social species cooperate. But what appears to distinguish humans from other animals is that only humans experience intentions towards common goals as shared with their partners. According to Bratman (1992, p. 328), such shared intentionality manifests itself by specific behavioral patterns, visible during cooperative activities, that is, (1) 'mutual responsiveness', (2) 'commitment to the joint activity', and (3) 'commitment to mutual support'. First, 'mutual responsiveness' requires behavioral evidence for partners' attending and adjusting to each other's actions. Second, 'commitment to the joint activity' requires evidence that both partners are actively trying to achieve the common goal (Tomasello & Carpenter, 2007; Tuomela, 2007), (Clark, 2006, p. 126). Third, 'commitment to mutual support', finally, requires behavioral evidence for partners helping each other fulfilling their roles.

So far, there is no clear evidence as to whether animals, including the great apes, experience cognitive states akin to shared intentionality (Call, 2009; Tomasello & Moll, 2010). This is based on laboratory experiments in which adult human experimenters engaged subjects in triadic social games in order to then unexpectedly interrupt them. In one study, human infants from the age of 3 years typically exhibited behavioral and communicative attempts to reengage their reluctant partner, chimpanzees of about 3 years of age did not (Warneken et al., 2006). In another study, a human experimenter engaged chimpanzees older than 6 (and with an average age of 19.2 years considering the sample) years in a highly reciprocal triadic social game (Heesen et al., unpublished data), which revealed consistent reengagement attempts (see also MacLean & Hare, 2013; Pika & Zuberbühler, 2008).

Inconsistency of results between studies may be due to species or age effects, while interactions with human partners and their artefacts may provide additional motivational barriers. Indeed, when naturally interrupted during an ongoing social activity with a conspecific, chimpanzees and bonobos attempted to reengage their initial partner most of the time (Heesen et al., unpublished data, see Chapter 5), and often resumed the actions performed right before the interruption, and at the same location, indicating some sort of continuity of the activity rather than the initiation of a new interaction. Although these findings provide a first glimpse into the apes' perception of shared activities with conspecifics, and extent previous findings from experimental studies on joint commitment of great apes (Gruber, 2014; Heesen et al., unpublished data; MacLean & Hare, 2013b; Pika & Zuberbühler, 2008; Warneken et al., 2006), the evidence yet

remains to be strengthened due to several limitations. First, the interruption stimuli in our previous study occurred spontaneously and randomly and were thus of different nature and degree of severity. For example, one may consider the delivery of food as a severe interruption stimulus, possibly changing the mental state from a relaxed, cooperative state to a more competitive one, probably reducing the likelihood of reengagement. On the contrary, one may consider a group member interrupting the activity between two partners as a mild interruption stimulus, being shorter in length and causing less stress for competition. As a consequence, it was difficult to disentangle between the different factors potentially affecting the apes' motivation to reinstate activities with partners (e.g., emotional state, severity of interruption, identity of partner).

Second, there was thus far no control condition to relativize the reengagement rates of partners during social activities in comparison to other sorts of activity, such as solitary play or grooming. Apes may just generally resume any activity or task because they memorize initial task goals (Trafton et al., 2003). Third, the previous results do not allow insights into the apes' understanding of the consequences of breaking joint commitments. Dealing with interruptions of joint activities in humans is "face" threatening (Bangerter et al., 2010) – with face being form of publicly manifest self-esteem (Goffman, 1955), especially when one partner is kept waiting because of their partner engaging in another interacting with a third party, such as an interrupted conversation (Bangerter et al., 2010). When interrupted, partners need to coordinate on suspending the activity, dealing with the interruption, and then reengaging in the activity (Chevalley & Bangerter, 2010). Politeness is a form of face management that serves to mitigate face threats inherent to the suspension and reinstatement of interrupted activities, and may involve apologies and justifications on why the interaction had to be suspended (Bangerter et al., 2010), and reconstructing topics when resuming the activity with a partner (Bangerter et al., 2010).

Studies involving human children, for example, have shown that children understand the consequences of breaking commitments as they explicitly acknowledge their own leaving communicatively to their partner (e.g., Gräfenhain et al., 2013), or resist offered rewards to maintain explicitly agreed upon joint commitments from the age of 3 (Kachel & Tomasello, 2019). To date, it remains unstudied whether apes produce intentional communicative signals, such as gestures, vocalizations and facial expressions, to coordinate the suspension of a social activity with peers. Bonobos especially are sensitive to partners' social attributes, calibrating their communicative efforts in the opening and closing phases of an interaction according to the social

distance and (to some extent) status difference with their partners (Heesen et al. unpublished work, Chapter 3). There are thus ample reasons to study whether bonobos may be concerned with some form or precursors of face management when dealing with interruptions of natural social and cooperative activities with conspecifics.

In the present study, we attempted to address the shortcomings listed above and to strengthen the evidence for joint commitment understanding in apes. First, we investigated whether the social nature of the activity (in comparison to a solitary activity) promotes partners' motivation to resume the activity after being interrupted, thus revealing awareness to joint commitment. To do so, we investigated individuals' likelihood of resuming an activity after the interruption of a solitary activity (self-grooming or solitary play) compared to a social activity (social grooming).

Second, as predicted by the politeness theory (Brown & Levinson, 1987) for human interactions, we assessed the possibility that bonobos calibrate their behavior and signals flexibly to reduce putative "face" threats according to the identity of their partner. To do so, we interrupted dyads of differing social distance (hereafter "social bond") and power difference (hereafter "rank difference"), in two conditions using two different interruption stimuli. In the first condition, the interrupting stimulus was specifically directed at a ***focal*** individual ("Targeted interruptions"), i.e., a keeper calling out the target individual's name in the distance. This type of stimulus usually provokes the approach of the individual that is being called. With this condition, we attempted to mirror previous experimental paradigms used for human children, (Gräfenhain et al., 2013; Kachel & Tomasello, 2019), in which one of the partners (the child) was interrupted by being offered a more attractive activity or an individual reward by a third party, thereby threatening the ongoing interaction with the current interaction partner. To coordinate the suspension of their play interaction, children verbally acknowledged their leave-taking, showing awareness of joint commitment. In the second condition, the interrupting stimulus was undirected ("Untargeted interruptions"), i.e., an opening sliding (holding) door noise was produced inside the holding area, suggesting the presence of a keeper ready to scatter food inside the building. This type of stimulus usually interrupts the entire group's ongoing activity (thus both the ***focal*** individual and partner) and the individuals respond by approaching the holding area to check inside. We predicted that the target interruption types would presumably increase the pressure for patterns similar to face management in humans (e.g., Bangerter et al., 2010; Gräfenhain et al., 2013) as only one individual

is leaving and is thus responsible for suspending the interaction. We expect that the likelihood of reengaging partners after interruptions and to communicate about suspension and reengagement will be higher in the Targeted interruption condition compared to the Untargeted interruption condition. Similarly, we expect this likelihood to be higher when the social distance and power difference with partners is larger (Brown & Levinson, 1987). These two conditions were then compared to the solitary activity by observing the reaction of individuals engaged in solitary activities, i.e. solitary play and self-grooming, before and after the interruption stimuli were produced.

Lastly, we explored the extent to which bonobos understand their own role in the joint activity. Indeed, role-understanding plays a key role in the management of social integrity during shared cooperative activities in humans (Bratman, 1992). To this end, we assessed whether individuals resume the role that they had been taken before the interruption, and whether they deploy more behavioral and communicative efforts when they were responsible for suspending the activity, when they were the ones that were responsible for initiating the interaction, and when they were the receiver of grooming at the time of suspension (see Bangerter et al., 2010, where listeners were more polite than speakers). If bonobos are aware of joint commitment and the socio-affiliative imperatives linked to it (Enfield, 2008), we would expect the partner who is responsible for suspending the activity, or for having initiated it, or for being the active partner in the ongoing interaction (groomer), to communicate more, as they would presumably be to "blame" for the threat of integrity, and for demeaning other's efforts and time. Partner's efforts to maintain social integrity may be shown also by their interest to resume previous roles (e.g., an individual taking up its role as the groomer after the interruption to continue where they left off), which is yet another aspect we inspected in this study.

We attempted to address the following research questions:

1. Are individuals more likely to resume grooming vs. solitary activities?
2. Does partner's resumption of social activities depend on the social bond or rank difference (controlling for experimental condition)?
3. Does the likelihood of individuals' reengagement attempts depend on their interaction role (responsible for suspension, initiator, groomee)?
4. When individuals suspend or reengage, does their likelihood to communicate depend on social bond or rank difference (controlling for experimental condition)?

5. Does the likelihood of individuals' communication about suspension / reengagement depend on their interaction role (responsible for suspension, initiator, groomee)?
6. Do suspenders resume the same role after the interruption than before?

Methods

Study subjects and site

The experiment was conducted at the zoological Park of La Vallée des Singes, France from May to August 2018. The bonobo group consisted of 17 individuals, of which 15 participated in our experiment (mean age= 17 years; S.D.=12 years; range age=4-51 years; 10 females, 5 males, see ESM S5.1 for detailed group composition). We focused on 30 dyads with either a strong (*n*=16) or weak (*n*=14) bond (see ESM S5.2). The strength of social bond was determined using the DSI method (see specific methods section Chapter 3 and 5 for details). We classified strong bonds as those with a DSI index >1 (i.e., the average DSI value in this sample), and weak bonds as those with a DSI index <1 (Silk et al., 2006, 2013). The focal individuals were selected based on whether they had partners with social bonds above or below the average DSI index 1; the DSI data was collected in the study of the previous year (2017, see Study 1). We selected 8 focal individuals (see ESM, S5.2), each of which had at least one strongly and one weakly bonded partner.

The group lives in large enclosures comprised of an outdoor island enclosure with a large forest and climbing structures in grassy areas (8000$m^{2)}$ and an indoor enclosure with various enrichment and climbing structures (600m^2). The group receives food five to six times a day, including daily ratios of primate pellets, fruits, and vegetables. Occasionally, the bonobos receive rice, nuts, meat, and eggs. Individuals can additionally forage for wild berries and herbaceous vegetation in their outdoor enclosure. In stable weather conditions (>13°C), the group is locked in their outdoor enclosure. Water is always available *ad libitum* from a fresh water source at the building and a stream surrounding the island.

Materials and design

Social activity type

We selected social grooming as the best joint activity candidate for this experiment as bonobos engage in grooming bouts very frequently on a daily basis (Sugiyama, 1988), often over long periods of time and in a reciprocal manner (Newton-Fisher & Lee, 2011) Furthermore, grooming serves – apart from its hygienic function - myriad of social functions and is beneficial

for managing tolerance (Port et al., 2009), social relationships (Schino et al., 2007), group cohesion (Dunbar, 1991), and social bonding (Fedurek & Dunbar, 2009). This allowed us to interrupt various different partner combinations and repeatedly. We filmed all behaviors during interruption periods by using two Panasonic HC-V770 Camcorders with externally attached directional microphones on tripods located around the enclosure (see fig. 19). We filmed grooming bouts from beginning (one or both partners start grooming the body of the other) until the end (partners stopped grooming for at least 2min).

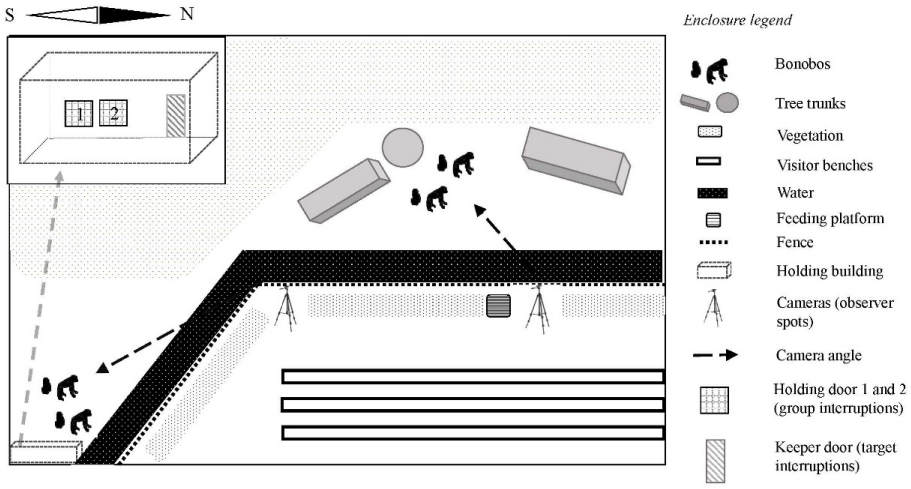


Figure 19 Experimental setup from a birds-eye perspective. The scale of the map does not correspond to the realistic scale and only serves as simplified representation of the setup.

Solitary activity types

As a control to the social activity, we investigated subjects' motivation to resume a solitary activity. We selected two equivalents of social behaviors that subjects could carry out as solitary behaviors - self-grooming and solitary play. We filmed and considered any individual who engaged in either of these two solitary behaviors in the instant of an Untargeted or Targeted interruption stimulus and then checked whether individuals resumed the same solitary behavior within 2min after the interruption stimulus. We filmed solitary behaviors before and during interruption periods by using two Panasonic HC-V770 Camcorders with externally attached directional microphones on tripods located around the enclosure (see fig. 19). We filmed behaviors from beginning (start of self-grooming or solitary play) until the end (partners stopped the behavior for at least 2min). We found 26 solitary behaviors at the right instance of the interruption stimulus

(10 solitary plays and 16 self-groomings), comprising on average 2.4 ± 1.8 solitary behaviors per individual across conditions (*N*=11 individuals; Untargeted interruption: 1.8 ± 1.5; Targeted interruption: 1.6 ± 1.3, see ESM S5.2). Of the 11 individuals that produced solitary activities during the experiment, 4 were also focal individuals in other trials of the experiment.

Procedure

The experimental trials were conducted opportunistically in the late afternoon when the bonobos were in their outdoor enclosure, before the group was let in for the night (fig. 19). We waited for the subjects to be located near the building or nearby tree trunks in the open-space of their outdoor enclosure, as both spots were popular grooming location and in full view. RH was standing in the public area and filmed the dyads of interest as well as the whole group, using two cameras placed at the two different locations (fig. 19). RH timed the grooming interaction from its onset to the determined interruption time. RH called the keeper of the day, standing inside the holding area, via Walky Talky (Motorola GP340) five minutes before the determined interruption time, and again 30 seconds before that time point. Upon cue, the keeper produced the interruption stimulus from within the building.

Using a repeated measures within-subjects design, we interrupted focal individuals interacting with partners in dyads of differing social bond (weak vs. strong) repeatedly during social grooming activities in two conditions: the Untargeted interruption and the Targeted interruption conditions, see fig. 20.

In the first condition, the Targeted interruption condition, only the ***focal*** individual was targeted by the interruption stimulus. Upon receiving the cue, the keeper called the focal subject by its name through the holding door of the indoor enclosure (Table 9). Upon the focal subject's arrival at the holding door, the keeper rewarded him with a desirable food item, such as carrots, juice or raisins (i.e., keepers chose food items randomly each time to avoid habituation). If the focal individual did not approach after several attempts of calling, the keeper closed the holding door and the trial was cancelled. This type of stimulus is regularly used to attract one individual at a time for medical treatments or general health checks, to which individuals willingly respond by approaching. We ran 39 Targeted interruption trials (mean = 4.6 trials per focal individual, S.D. =1.4), of which we discarded two trials as the stimulus did not lead to an interruption of the dyads' grooming bouts. From the 37 remaining Targeted interruptions, 8 were conducted with weakly bonded dyads, and 29 with strongly bonded dyads (see ESM S5.2; mean interruptions per dyad=

1.9, S.D. =0.9).

In the second condition, the Untargeted interruption condition, we broadcasted a stimulus that was not directed at any individual in particular but that potentially affected the entire group (thus both partners of the dyad). Upon receiving the cue, the keeper produced the stimulus by producing noise of a rapidly moving holding door of the indoor facility, simulating a scenario by which the group is usually let inside the indoor facility to receive the evening meal and rest for the night (Table 9). The keeper rapidly opened and closed either the first or the second holding door (randomized across trials; see for location of holding doors fig. 19). This created a typical noise that is being associated with a scenario in which the bonobos would enter the building for the night (Table 9). This type of noise usually interrupts the entire group's ongoing activity and provokes the approach of all group members to the holding doors to check inside. In this condition, no bonobo received any food rewards and the holding doors remain closed after the stimulus. We ran 49 Untargeted interruption trials (mean = 6.1 trials per focal individual, S.D. =2.3), of which we discarded one trial as the stimulus did not lead to an interruption of the dyads' grooming bouts. From the 48 remaining Targeted interruptions, 17 were conducted with weakly bonded dyads, and 31 with strongly bonded dyads (see ESM S5.2; mean interruptions per dyad= 1.6, S.D. =1.2).

These two conditions allowed for comparison between a situation where only one partner is responsible for interrupting the interaction to receive an individual reward while the partner is kept waiting (Targeted interruption), and a situation where both partners are equally responsible for interrupting the activity (with the expectation of obtaining food inside the building).. We expect that individuals will produce less communicative and coordinative efforts in the Untargeted interruptions compared to Targeted interruptions. We ran trials opportunistically, whenever focal dyads engaged in grooming, but counterbalancing the number of Untargeted and Targeted interruptions across the study. To avoid habituation and extinction of behaviors, we kept daily interruptions at a minimum; the maximum was two interruptions a day (either Untargeted and Targeted, Targeted and Targeted, or Untargeted and Untargeted). We randomized the time of Untargeted and Targeted interruptions in the early time slot (4-5pm) and late time slot (5-6pm), maintaining a balanced number of Untargeted and Targeted interruptions across time slots.

Interruption time

We hypothesized that in the process of social grooming, partners may feel more committed to the grooming activity at the beginning of the bout than towards the end. They may feel less

responsibility towards a potential joint commitment if interrupted at a later stage of their interaction compared to an earlier stage. For these reasons, we controlled for the time of interruption of grooming for each dyad. We determine the best interruption time for each dyad based on preliminary observations during a 5-months study conducted at La Vallée des Singes form May to September 2017 (total 330.3h; mean= 20.6h per individual; S.D.=0.5h). The distributions of dyad's grooming durations were estimated from at least three complete interactions per dyad (range=3-6 grooming interactions). We then computed the average duration of grooming bouts for each dyad based on these sampled interactions. We decided to determine the best period of interruption as the period between the onset of the grooming bout and the median of the dyad's grooming bout duration (see ESM S5.2), hereafter referred to as "interruption time range" (ITR). The beginning of the ITR corresponds to the mid-time-point between the start of grooming and the median duration of previously measured dyads' grooming bout lengths. The end point of the range corresponds to the median duration of the dyad's grooming bouts. We interrupted the dyads as accurately with both Targeted and Untargeted interruption stimuli as possible within this time range. For *n*=10 dyads (1/3 of the sample), there were either no or not enough interactions to be analyzed from the previous data collection and we could therefore not pre-determine an ITR for these dyads. As a compromise, we interrupted these dyads within a standard time window of 3 to 18 minutes (based on the ITR of the other dyads).

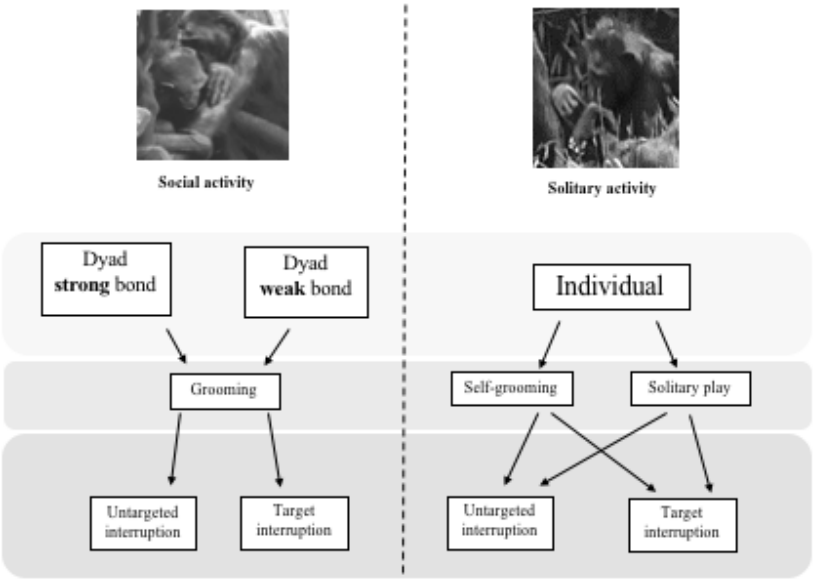


Figure 20 Experimental design. In the social activity, we interrupt weakly and strongly bonded dyads during social grooming interactions in two experimental conditions: Untargeted and Targeted interruptions. In the solitary activity, we recorded individuals engaged in self-grooming and/or solitary play at the time of Untargeted and Targeted interruptions.

Table 9 Description of experimental conditions, interruption stimuli and methods.

Condition	**Interruption stimulus**	**Interruption directed to**	**Implementation**	**Time for stimulus onset (for social grooming)**
Untargeted	Sound of holding door opening and closing	Whole group	Keeper moving holding door from inside the building. Simulation of the sound when holding doors open to shift group from outdoor to indoor enclosure for feeding.	When the dyad arrived at their "interruption time range"
Targeted	Keepers calling focal individual's name	Individual	Keeper calling the name of an individual from the holding door to offer food item.	Same as above

Social bond and rank difference

Data were collected simultaneously to the study period (May-August 2018). This included 1h per day with 15 minutes of continuous focal follows, 5-minute scan samples and ad libitum conflict data on selected focal individuals during the time of the experiment (see ESM, S5.2); we

collected 62.3h of observations in total for social data (7.7h per focal individual, S.D.=0.28).We carried out analyses with the statistical computer software R (v. 3.5.0). We computed social bonds using dyadic sociality indexes (DSI) (Silk, Alberts, & Altmann, 2006; Silk, Cheney, & Seyfarth, 2013) using the socialindices package in R (Neumann, 2016, accessed via https://github.com/gobbios/socialindices). The DSI thus serves as inverse proxy for social distance (Levinson & Brown, 1987). These involved the count of individual's proximity to others (in arm-reach distance), count of approaches and duration of grooming (and play during last year's 5-months observation period). We collected durations for social activities and counts of approaches during continuous 15-minute focal samples, and proximity data via 5-minute scan samples. To compute the rank distance between grooming partners, we produced a rank hierarchy with dominance scores for each individual, based on recordings of conflict data collected during the previous study. We computed ranks using EloRatings (Albers & de Vries, 2001; Neumann et al., 2011), computed with the "EloRating" package in R (Neumann & Kulik, 2014). We then computed rank differences by subtracting individuals' Eloratings; rank differences served as proxy for power differences (Levinson & Brown, 1987). Higher values indicate stronger bonds, lower values indicate weaker social bonds (Silk et al., 2006, 2013). We manipulated the variable social bond (i.e., interrupting weakly vs. strongly bonded dyads), while controlling also for rank differences between partners. As our mean DSI value was 1, values >1 indicate strong bonds, and values <1 indicate weak bonds.

Video coding

We coded all behaviors and communicative signals using the computer software ELAN (Version 5.2, 2018, April 04; Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/; see Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Specifically, we coded the following variables: individual IDs of the focal individual and its partner; the type of activity they engaged in before the interruption (social grooming/ self-grooming/ solitary play); whether the activity performed before the interruption was resumed or not within 2min after the interruption stimulus (resumption/no resumption); whether communicative signals or mutual gazes were exchanged during the interruption period to suspend (3sec before and 3sec after interruption) or to reengage partners (3 sec before resumption of grooming (for both: communication/no communication, see ESM S5.3 for definitions of signals); the role of each partner in instigating the interaction (initiator/receiver); the role of each

partner in the interaction at the time of the interruption stimulus and after reengagement (groomer/groommee); and the role of each partner in the interruption (responsible/not responsible for suspension; the individual responsible for suspension is the one who stopped the activity first).

Statistical analysis

In the Targeted interruption condition, we expected the focal individual (the one called by name) to be the one responsible for interrupting the activity (the one who stopped the activity first and who left the partner standing while collecting its reward), however that was not always the case and sometimes both partners stopped simultaneously, or it was the other partner who was responsible for the interruption. As a consequence, and to ensure comparability between conditions, we either focused the analysis on the focal individual (and compared its behavior across conditions) or on the individual responsible for interrupting or reinstating the activity (not necessarily the focal individual in the Targeted condition), depending on the research question

In order to test the understanding of joint commitment in our subjects, we compared the likelihood of resuming the activity after interruption of a social versus a solitary activity. Additionally, we explored the likelihood of resuming the grooming activity with the same partner after interruption, and the likelihood of communicating with their partner at the times of suspension and resumption, depending on: the experimental condition (Targeted/Untargeted interruptions), the nature of the relationship between partners (strong/weak bond, higher/lower ranking) and the role of partners in the interaction (initiator/receiver, responsible/non responsible for suspension and resumption, active/passive).

To test our research questions, we fit Bayesian generalized mixed models using the Stan computational framework (http://mc-stan.org/), accessed through the brms package in R v. 3.5.0 (R Core Team, 2013). We fit each model with random effects of IDs of individuals and their partners or dyads, where relevant (see ESM S5.4 for selected models' specificities); for all models, we compared full models including respective random effects with reduced models excluding random effects, to verify whether mixed modeling was justified. We ran LOOIC comparisons and chose the model with the best expected log predictive density (ESM S5.6) and sufficient variance of random effect intercepts (>=0). We consistently presented the results of the more parsimonious models if random effect modeling was not justified, meaning we conducted GLMs instead. To ensure transparency, we nonetheless present non-selected models' results in ESM S5.7. The models included four MCMC chains, with 10,000 iterations per chain, of which we specified 2,000

iterations as warm-up to ensure sampling calibration, leading to 40.000 posterior samples. The model diagnostics revealed that the posterior distributions reflect the distribution of the original response values appropriately, R-hat statistics were <1.05, the number of effective samples were >100, and the MCMC chains had no divergent transitions (ESM S5.5). For all models, we used the default priors of the brms package, which were weakly informative with a student's t-distribution of 3 degrees of freedom and a scale parameter of 10. For inference, we calculated 95% credible intervals from the posterior distributions and checked whether 0 was comprised in this interval.

To test whether the nature of the activity affects the likelihood of resumption, we fit a model ("model 1") with the dependent variable "resumption" or "no resumption" (fitting a Bernoulli distribution with presence or absence of resumption as binary outcome). The fixed effect was the nature of the activity (social vs. solitary). This dataset included 30 dyads and 11 individuals, with 111 interruptions (85 social and 26 solitary activities).

To test whether the social bond and rank difference, as well as experimental condition, affect the likelihood of resumption, we fit a model ("model 2") with the dependent variable "resumption" or "no resumption" (fitting a Bernoulli distribution with presence or absence of resumption as binary outcome). The fixed effects were the DSI (i.e., a measure of social bond serving as inverse proxy for social distance), and absolute rank difference between partners (i.e., as a measure of power difference), and the experimental condition (Untargeted vs Targeted interruptions). We z-transformed both DSI and rank difference for all individuals to mean = 0 and SD = 1. This dataset included 30 dyads and 85 interruptions.

To test whether the interaction roles affect individuals' likelihood to resume partners, we fit a model ("model 3") with the dependent variable the "resumption" or "no resumption" (fitting a Bernoulli distribution with presence or absence of resumption as binary outcome). The fixed effects were the role in the suspension (responsible vs not responsible), the role in the opening of the activity (initiator vs.receiver), and the role in the activity itself (active groomer vs. passive groomee). We excluded observations where roles could not be determined, such as in the case of "mutually suspended" interruptions (n=12 cases). The dataset thus included 10 individuals and 158 interruptions (i.e., 158 data points as we looked at each individual's behavior not the dyad's).

To test whether the social bond and rank difference, as well as the experimental condition, affect the likelihood of communication at the time of suspension and resumption, we fit two models

("model 4 and 5") with the dependent variable the "communication" or "no communication" (fitting a Bernoulli distribution with presence or absence of communication as binary outcome). The datasets is comprised of data points of individuals *responsible* for suspension or resumption of the activity. The fixed effects were the DSI, the rank difference between the individual suspending/ reengaging and its partner, and as a control variable the experimental condition (Untargeted vs. Targeted interruptions). We z-transformed both DSI and rank difference as before. The dataset for model 4 included 10 individuals responsible for suspending (with 12 different possible partners) and 79 interruptions (i.e., of n=85 interruptions, n=6 interruptions were excluded as they were mutually suspended by partners, meaning no clear individual responsible for suspension was identified) . The dataset for model 5 included 9 individuals responsible for reengaging and 66 interruptions (i.e., of n=85 interruptions, n= 2 were excluded as we could not identify one individual clearly responsible for resumption, and n= 17 interruptions were not resumed, hence communication was no option).

To test whether the interaction roles affect *any* individuals' tendency (i.e., not just the one responsible for suspension or resumption) to communicate at the time of suspension or resumption, we fit two models ("model 6 and 7") with the dependent variable "communication" or "no communication" (fitting a Bernoulli distribution with presence or absence of resumption as binary outcome). The fixed effects were the role in the suspension (responsible vs not responsible), the role in the opening of the activity (initiator vs. receiver), and the role in the activity itself (active groomer vs. passive groomee). We constructed dataset where each row indicates whether a partner of a dyad communicated at the time of suspension / resumption, taking each row as independent data point of one individual acting within a dyad. We excluded observations where roles could not be determined, such as in the case of "mutually suspended" interruptions (n=12). The dataset used to test the communication at suspension (model 6) thus included 12 individuals and 158 interruptions. Besides data points where individuals responsible for suspension could not be identified, we excluded n=32 data points representing non-resumed interactions, reducing the dataset used to test the communication at resumption (model 7) to 126 interruptions.

Results

We conducted 88 interruptions of social grooming activities in total (n=39 in the Targeted interruptions condition and n=49 in the Untargeted interruptions condition). In three of those interruptions individuals did not interrupt their activity, meaning we were not able to include these

cases in our analyses (*n*=2 in the Targeted interruption condition; *n*=1 in the Untargeted interruption condition). Resumption of the social activity with the initial partner was observed after 73.0% of Targeted interruptions and 85.4% of Untargeted interruptions. In contrast, interruptions of solitary activities resulted in lower resumption rates in both conditions (25% vs 50% in Targeted and Untargeted interruptions respectively*; see* ESM S5.2 for raw data).

1. **Are individuals more likely to resume grooming vs. solitary activities?**

Model 1 suggested a substantial relationship between activity resumption and the nature of the activity; chances of resumptions were substantially lower following interruptions of a solitary activity compared to a social one (fig. 21A; ESM S5.8A; b (SE) = -1.73 (0.49), 95%CrI [-2.71, -0.77]).

2. **Does partner's resumption of social activities depend on the social bond or rank difference (controlling for experimental condition)?**

The result of model 2 revealed no substantial relationships between the likelihood of resuming an activity and the experimental condition (b (SE) = -0.85 (0.58), 95%CrI [-2.02; 0.26]), social bond (b (SE) = 0.15 (0.31), 95%CrI [-0.42, 0.78]), or absolute rank difference between partners (b (SE) = -0.01 (0.29), 95%CrI [-0.59, 0.57]), see fig. 21B-D; ESM S5.8B.

3. **Does the likelihood of individuals' reengagement attempts depend on their interaction role (responsible for suspension, initiator, groomee)?**

The results of model 3 revealed a relationship between resumption and individual roles in the grooming action and initiation of the activity (ESM S5.8C). Individuals were less likely to reengage initial partners when they were the passive partner (i.e., "groomees"), than when they were the active partner (i.e., "groomer") (fig. 21E; b (SE) = -0.98 (0.45), 95%CrI [-1.91, -0.13]). Individuals were more likely to reengage their initial partners to the interrupted grooming interaction when they were the ones who initiated ("initiator") the grooming bout compared to being the receivers ("receiver") (fig. 21G; b (SE) = 1.16 (0.38), 95%CrI [0.43, 1.92]). However, there was no relationship between resumption and the individual role in the interaction at the moment of suspension (groomer/groomee) (fig. 21F; b (SE) = 0.30 (0.42), 95%CrI [-0.52, 1.11]).

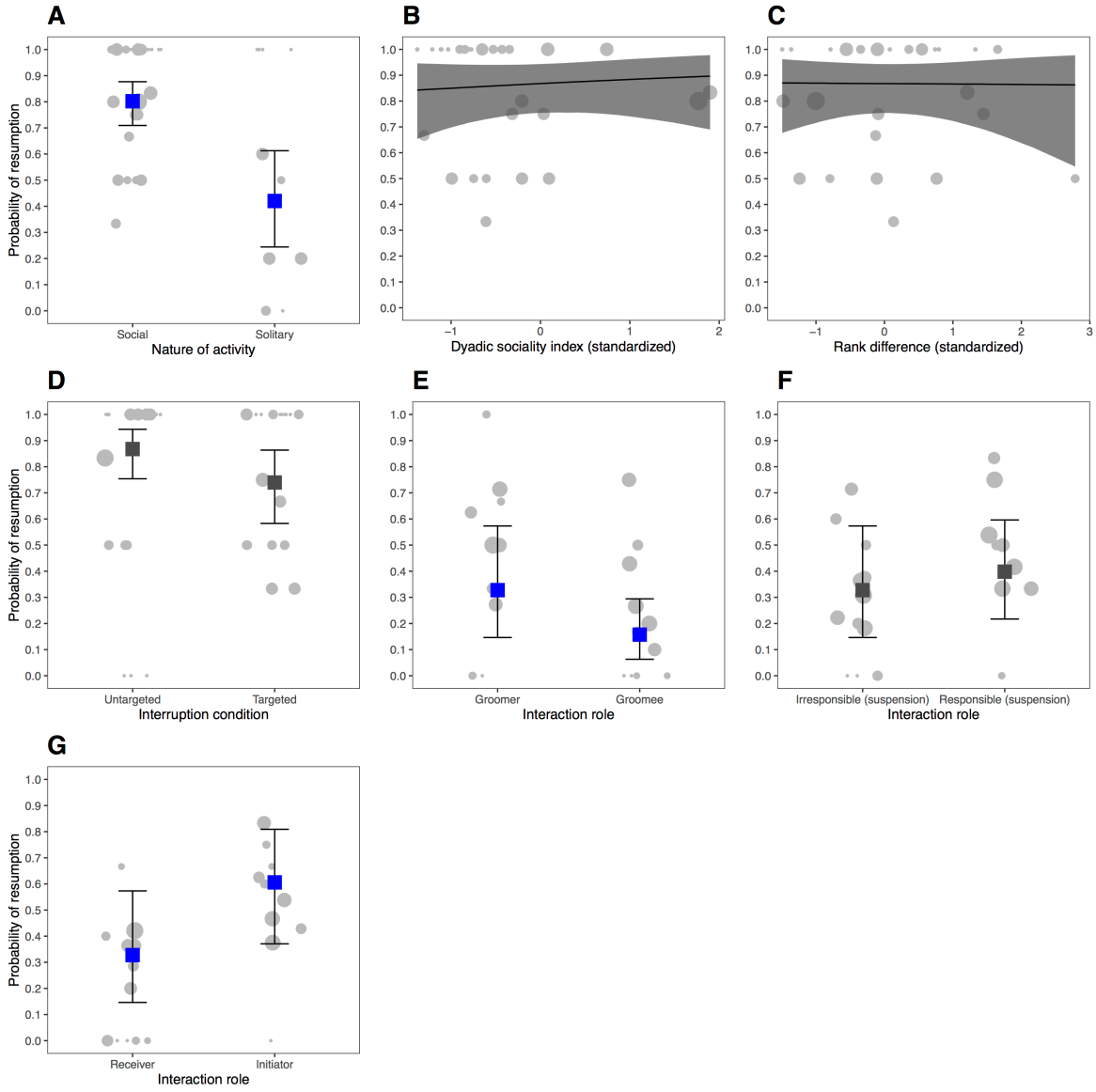


Figure 21 Variation of resumption probability according to activity nature (A), DSI (B), rank difference (C), interruption condition (D), and interactional roles (E-G). Plots depict the predicted probability of resumption for the marginal effects of the complete Bayesian models (A: model 1; B-D: model 2; E-G: model 3) and shows how model predictions match the actual data. The upper and lower bars or upper and lower edges of the ribbons depict the upper and lower 95% credible intervals, respectively, and the mid-square or mid-ribbon-line represents the predicted posterior mean. Squares or ribbons colored in blue represent substantial effects as estimated by the model. Each grey circle corresponds to the proportion of resumption per dyad [a(social); b-d] or individual [a(solitary); e-g]. Size of circles corresponds to number of observations; larger size indicates more frequent observation.

4. When individuals suspend or reengage, does their likelihood to communicate depend on social bond or rank difference (controlling for experimental condition)?

Model 4 revealed substantial relationships between communication at the time of suspension and both social bond and rank difference, as well as the interruption condition (ESM S5.9A).

individuals were less likely to communicate at the time of suspension when they were more closely bonded with their partner (fig. 22A; b (SE) = -0.94 (0.46), 95%CrI [-1.93, -0.13]), and when they were higher-ranking compared to their partner (fig. 22B; b (SE) = -1.24 (0.72), 95%CrI [-2.82, -0.03]). In addition, model 4 revealed a relationship between communication at the time of suspension and experimental condition, with individuals being less likely to communicate in the Targeted interruption condition compared to the Untargeted interruption condition (fig. 22C; b (SE) = -1.46 (0.72), 95%CrI [-2.99; -0.11]).

Model 5 revealed no effect of social bond on communication at the time of resumption (fig. 23A; b (SE) = 0.21 (0.31), 95%CrI [-0.39, 0.82]). However, we found a substantial relationship between the propensity to communicate at the time of resumption and both experimental condition and rank difference (ESM S5.10A). Individuals were more likely to communicate to reengage previous partners when they were higher-ranking relative to their partner (fig. 23B; b (SE) = 1.00 (0.33), 95%CrI [0.38, 1.68]), and in the Targeted interruption condition compared to the Untargeted interruption condition (fig. 23C; b (SE) = 1.93 (0.73), 95%CrI [0.59, 3.47]).

5. Does the likelihood of individuals' communication about suspension / reengagement depend on their interaction role (responsible for suspension, initiator, groomee)?

Model 6 revealed no substantial effects on communication at the time of suspension of any of the tested interactional roles, including responsible for suspension (b (SE) = 0.45 (0.52, 95%CrI [-0.57, 1.48), initiator (b (SE) = 0.05 (0.44), 95%CrI [-0.83, 0.92]), and groomee (b (SE) = -0.83 (0.53), 95%CrI [-1.90, 0.17]), see fig. 22D-F; ESM S5.9B.

Likewise, model 7 revealed no substantial effect on communication at the time of resumption for the interactional role groomee (fig. 23D; b (SE) = -0.12 (0.55), 95%CrI [-1.21, 0.94]). However, we found that individuals were substantially more likely to communicate at the time of resumption when they had been responsible for suspension (fig. 23E; b (SE) = 1.06 (0.49), 95%CrI [0.11, 2.05]), and for initiating the interaction (fig. 23F; b (SE) = 0.97 (0.44), 95%CrI [0.12, 1.83]), see also ESM S5.10B.

6. Do suspenders resume the same role after the interruption than before?

In the majority of the trials (83%), the individuals responsible for suspending the activity resumed the same role in the grooming activity (active groomer/passive groomee) that they occupied before the interruption.

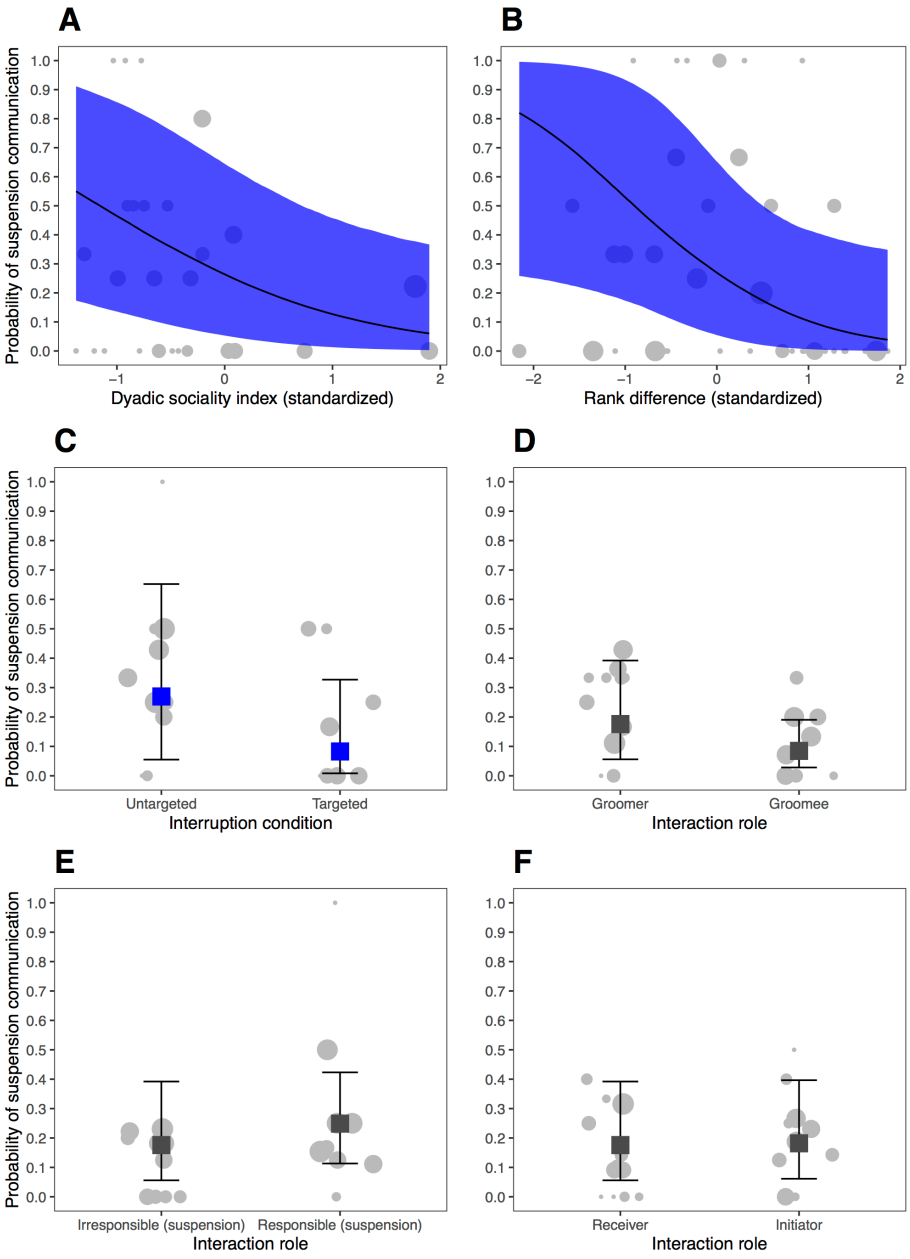


Figure 22 Variation of suspension communication probability according to DSI (A), rank difference (B), interruption condition (C), and interactional roles (D-F). Plots depict the predicted probability of suspension communication for the marginal effects of the complete Bayesian models (A-C: model 4; D-F: model 6) and shows how well the models predict the actual data. The upper and lower bars or upper and lower edges of the ribbons depict the upper and lower 95% credible intervals, respectively, and the mid-square or mid-ribbon-line represents the predicted posterior mean. Squares or ribbons colored in blue represent substantial effects as estimated by the model. Each grey circle corresponds to the proportion of suspension communication per individual responsible for suspension (A-C) or any individual regardless of whether or not they were responsible for suspension (D-F). Size of circles corresponds to number of observations; larger size indicates more frequent observation.

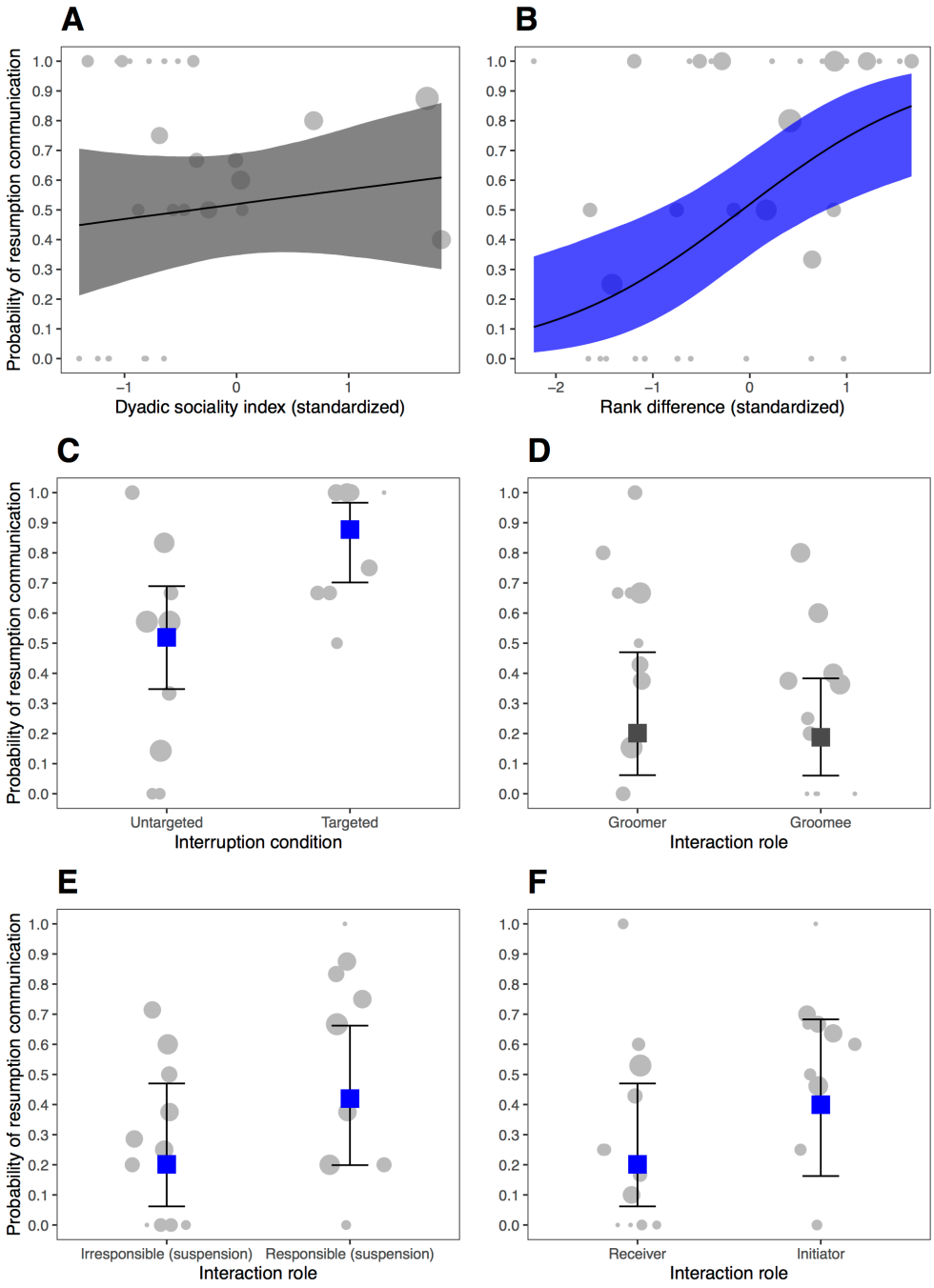


Figure 23 Variation of resumption communication probability according to DSI (A), rank difference (B), interruption condition (C), and interactional roles (D-F). Plots depict the predicted probability of resumption communication for the marginal effects of the complete Bayesian models (A-C: model 5; D-F: model 7) and shows how the model predictions match the actual data. The upper and lower bars or upper and lower edges of the ribbons depict the upper and lower 95% credible intervals, respectively, and the mid-square or mid-ribbon-line represents the predicted posterior mean. Squares or ribbons colored in blue represent substantial effects as estimated by the model. Each grey circle corresponds to the proportion of resumption communication per individual responsible for resumption (A-C) or any individual regardless of whether or not they were responsible for resumption (D-F). Size of circles corresponds to number of observations; larger size indicates more frequent observation.

Discussion

In the present study, we aimed at investigating whether bonobos experience a sense of joint commitment when engaging in naturally-occurring joint actions with conspecifics. To address this, we studied whether partners of an interrupted social grooming interaction would subsequently resume the activity and whether resumptions were driven by a simple individual motivation to complete an unfinished task or, as in human joint action, by an underlying sense of responsibility towards the partner, a joint commitment to complete a shared goal. We also assessed whether this sense of responsibility towards the partner would vary depending on the social relationship between partners and their respective roles in the social activity.

In line with our predictions, we found that bonobos were less likely to resume a solitary than a social activity, suggesting that their motivation goes beyond a mere desire to complete an unfinished task but entails a sense of commitment towards the partner or the joint action itself. It could also be argued that the higher rate of resumption of social compared to solitary activities might simply be explained by the fact that it is more pleasant to be groomed than to groom oneself. However this explanation was ruled out by the finding that subjects' were in fact more likely to reengage when they had been in the role as active "groomer", compared to when they had been in the role as passive "groomees". This suggests that resumption in the social setting was not merely driven by the more pleasant feeling of being groomed by another partner.

When grooming interactions were interrupted, subjects consistently resumed the interrupted social activity with their initial partners, regardless of experimental condition or social relationships. Moreover, individuals who initiated the social activity were also more active in attempting to reengage, as well as to communicate to reengage, their partners after interruptions, while individuals who were responsible for suspending the joint action reliably resumed their abandoned roles (active/passive) following resumption, suggesting that participating in joint actions comes with a sense of responsibility towards these joint actions, that is, a joint commitment, and that bonobos possibly possess some awareness of the social consequences resulting from breaking it. (e.g., Gräfenhain et al. 2013).

Following our prediction, when grooming interactions were interrupted, the communication efforts of bonobos were higher at the time of resumption following Targeted rather than Untargeted interruptions, suggesting that apes do not just operate towards completion of memorized task goals (Trafton et al. 2003), but with an understanding of the social implications

of their actions. This interpretation is also in line with the fact that individuals resumed abandoned joint actions with their original partners and resumed their role prior to interruptions.

However contrary to our prediction, we found that at the time of suspension, individuals communicated less following a Targeted compared to Untargeted interruption. This finding might be explained by the fact that Untargeted interruptions could have been more arousing for subjects, therefore leading to a much intense emotional response involving communication, compared to Targeted interruptions. Although we could not assess this in the current study, we hope this will be tested in future studies by measuring arousal levels via non-invasive psycho-physiological techniques like thermal imaging (Dezecache et al. 2017), or by selecting stimuli that are pre-tested to ensure comparable arousal levels at the time of interruption.

The finding that bonobos communicated more at the time of resumption than at the time of suspension is contrary to what is found in humans (Bangerter et al. 2010), where people generally employ more communicative efforts suspending an activity than when later resuming it. This suggests that it might be difficult to tactically communicate and coordinate when a bonobo experiences high emotional arousal (like at the time of interruption), while in humans, containing one's own arousal state for the purpose of coordinating a joint task might represent a much simpler task. Similar to humans (Brown and Levinson, 1987), however, bonobos increased communication efforts when suspending joint actions as social distance and rank difference increased. More specifically, subjects were more likely to communicate at the time of suspension when they were socially distant and subordinate in rank, and more likely to communicate at the time of resumption after Targeted interruptions, but surprisingly more so when they were higher-ranking. Additionally, and again similar to humans, bonobos were more likely to communicate for resumption when they themselves were responsible for having suspended the activity compared to when they were irresponsible. This finding is reflected in the human literature, where toddlers as young as three years of age show greater likelihood to communicate when it was their responsibility that the joint activity was suspended (Gräfenhain et al. 2009; Gräfenhain 2013). These results suggest that bonobos possess some awareness of their own responsibility in the joint action, and the social consequences linked to breaking the joint commitment, and that they adjust their communication efforts accordingly, probably to deal with coordination problems somewhat similar to human face management (Brown and Levinson 1987; Bangerter 2010; Chevalley and Bangerter 2010).

Inferences from behaviors to mental states, especially complex ones, are always debatable, so it is important to rigidly apply the principle of parsimony (Morgan 1981), also with human data. In humans, it is usually accepted that joint actions are the product of an underlying joint commitment which, at least in adults, is tied to higher-order mental processes that involve concerns about how others perceive them, and concerns about avoiding *face threatening* acts (Brown 1987). In humans, *face management* (Brown 1987) increases as a function of social risk, and face-threatening acts become more likely with increasing social distance and power difference between two partners, necessitating more active use of politeness (Brown 1987). Also, since suspending joint actions is perceived as a bigger threat to a partner's face than resuming them, more active use of politeness is required for the former (Bangerter 2010). Furthermore, longer interruptions of joint actions require more politeness than shorter ones when resuming (Chevalley 2010). Finally, passive partners of a joint action are expected to deploy more politeness than active ones prior to suspending (Bangerter 2010), demonstrating that social roles during joint actions matter, yet this was not the case in our current study (i.e., bonobos were not more likely to communicate when in the role of passive "groomee"). Although our result indicate that bonobos' joint action seem to be governed by similar socio-affiliational imperatives than that of humans and follow some patterns predicted by face management (Brown 1987) (especially when communicating about suspension), there is no clear available evidence that bonobos are concerned about how others perceive them as much as humans do (Enfield 2008; Stivers et al. 2011).

Already from the age of three human children acknowledge their leave-taking when dropping from social games with a partner, but not when they were playing on their own (Gräfenhain 2009, 2013), suggesting that they understand their own involvement and responsibility in a joint action. Interestingly, three-year-old children, but not bonobos, resisted breaking from joint commitments when being offered personal reward by refusing to leave their partners in about 70-80% of N=96 trials (Kachel and Tomasello 2019) whereas in bonobos this was only observed in 0.04% of N= 89 trials, demonstrating that human infants will forgo opportunities for personal gain in order to honor ongoing joint commitments. Although not directly comparable, it is possible that humans and apes differ in how they prioritize social commitments relative to material personal benefits.

Taken together, bonobos seem to possess some awareness of their responsibility toward their partner in a joint action, and of the social consequences linked to breaking a joint action.

Furthermore they appear to follow some basic patterns predicted by face-management while, at the same time, prioritizing personal incentives over joint commitments, especially in situations involving highly arousing causes of interruptions like food rewards. Whether these differences in behavior are due to population-specific peculiarities (e.g., wild bonobos may be more socially committed), due to underlying psychological differences or due to differences in the importance of social norms cannot be decided with our data. Also important is that, compared to humans, non-human animals are thought to be less capable of planning the future and of forming long-term goals (Suddendorf and Corballis 2007), possibly due to differences in relative brain size and corresponding computational capacities (but see Rosati et al. 2007; Vlamings et al. 2010). Our findings nonetheless show that bonobos are more likely to adapt communication based on their interaction roles at the time of resumption (not suspension), suggesting the ability to bear in mind previous circumstances over some time, as well as to reflect on one's own responsibility in past interactions with certain partners. In conclusion, and leaving aside these complexities, our findings provide evidence that bonobos are endowed with social cognition that enables them to understand joint commitment, although possibly in a less rich way than humans. Humans try to avoid affronting others' self-esteems, to the extent that they are willing to sacrifice personal for joint interest. Nevertheless, contrary to previous claims (Call 2009; Tomasello and Moll 2010), bonobos show understanding of others' psychological states (Tomasello and Carpenter 2007) in the sense of joint commitment, a key prerequisite for human-unique "we" or shared intentionality (Tuomela 2007; Tomasello 2010), and part of a layered continuum of primate social cognition (Genty et al.submitted; Pika and Zuberbühler 2008; MacLean and Hare 2013; Gruber 2014).

Chapter 7. General discussion

The overall goal of my thesis strived to answer the one larger question: *Are apes capable of shared intentionality as interactional achievement?* To address this question, I implemented a joint action framework and applied it to study the natural interactions of bonobos and chimpanzees. I described the sequential, orderly process by which apes coordinate to get into, maintain, and get out of joint activities, as well as how they manage interpersonal relationships. Using a bottom-up approach while studying shared intentionality, I was particularly interested in the lower-level alignment processes (i.e., behaviors, gaze and communicative signal exchanges) in the joint action phases of apes. Solving coordination problems in entry, suspension, reengagement and exit phases would be an indicator of joint commitment understanding in humans, and hence the presence of such joint action phases in apes can serve as a behavioral assessment for joint commitment understanding in these non-human primate species. Additionally, I investigated whether the communicative efforts in these phases were adjusted according to the social relationships between two participants, notably to social bond, rank differences, and weightiness of impositions – variables which in humans constitute predictors of politeness use as face management (Brown & Levinson, 1987). Joint commitment and face management are critical for the successful coordination of joint action, and such states may become visible and emerge through the joint action coordination at the behavioral level. Both these abilities are representative of shared intentionality as interactional achievement (Heesen et al., 2017).

In this general discussion, I will first briefly revisit the rationale for implementing a joint action framework for the study of shared intentionality (Theoretical goal). Next, I will summarize my predictions and findings, and briefly touch on the individual implication of each empirical goal (Empirical goal 1 and 2, see fig. 25). After that, I will place my findings into the broader picture, and discuss how my findings might alter our perception that shared intentionality is uniquely human; in other words, I will discuss my findings' broader implications, limitations and potential future avenues resulting from my research. I will finish with a general conclusion.

Theoretical goal: The joint action framework

The evidence addressing the question of whether apes have shared intentionality according to Tomasello's definition (Tomasello, 2016) is controversial. Tomasello's top-down approach to studying shared intentionality implies that humans have the ability to share psychological states

with one another, and that this involves a suite of high-level cognitive abilities, such as joint attention, cooperative communication, perspective-taking, mutual help, complementarity of roles, and so on (see Chapter 1). He claims that, although apes may share some of these abilities, such as helping (Greenberg et al., 2010; Warneken et al., 2007), and role-reversing in collaborative tasks (Fletcher et al., 2012), they still fail at engaging in joint action via joint attention, joint commitment, and cooperative communication (Crockford et al., 2012; Genty & Zuberbühler, 2015; MacLean & Hare, 2013; Pika & Zuberbühler, 2008; Warneken et al., 2006). This led some scholars to assume that they must lack shared intentionality (Tomasello, 2016; Tomasello et al., 2005). In what follows, I will re-emphasize why it is has been still so difficult to show that apes have shared intentionality, to then remind the reader of why shared intentionality should be studied differently.

First, to the theoretical problems inherent to the classical definition of shared intentionality (Tomasello et al., 2005). Assuming that all these high-level cognitive abilities are needed for shared intentionality really is not helpful for studying joint action in any other being besides human adults. To illustrate this point. Children engage in joint action (Eckerman & Whitehead, 1999; Gräfenhain et al., 2009), even before they pass false belief tests by the age of four (Wellman, Cross, & Watson, 2001), and chimpanzees engage in group hunting, with complementarity of roles and sharing of spoils (Boesch, 2005), even though they apparently lack joint attention, cooperative communication or anything of the like (Tomasello & Moll, 2010). How may such sophisticated forms of joint action be possible in these agents, if shared intentionality is missing?

I argued in my introduction that one may study shared intentionality instead via a bottom-up approach, by figuring out the most *minimal* conditions that may characterize a state of togetherness in joint action. In this account, I acknowledged that joint commitment understanding may occur as non-conceptual representation (Rosas & Bermúdez, 2018), in that the participants may not actually be aware about the cognitive state that governs their joint commitment. In that way, joint commitment may not necessarily rely on mind-reading and mutual knowledge, but on participants' perceptions and public agreements. This would explain how joint action may occur in agents supposedly not governed by shared intentionality, such as apes. However, Tomasello's top-down definition of shared intentionality fails to explain exactly this (Rosas & Bermúdez, 2018) and may thus presume an unwarranted level of cognitive sophistication in joint action, although simpler processes may be at work (Vesper et al., 2010). Additionally, it sets the bar for what counts

as shared intentionality very high, which makes it impossible for any researcher to show that apes have shared intentionality or cognitive states alike (Leavens et al., 2017; Townsend et al., 2017). And, frankly, that seems unfair, given that apes really do interact in quite similar ways than we do (e.g., Fröhlich, Kuchenbuch, et al., 2016; Rossano, 2013), with sophisticated cogntive abilities such as recognition of false beliefs (Buttelmann et al., 2017; Krupenye et al., 2016).

Second, to the practical problems of the classical definition. I wish to draw attention to the fact that the assumption that apes lack shared intentionality predominantly rests on laboratory experiments, in which often encultured and over-trained apes were tested in artificial conditions, and often with human partners (Fletcher et al., 2012; Warneken et al., 2006; Warneken & Tomasello, 2006). Why would any chimpanzee coordinate joint action with a partner that shares very little of her communicative repertoire, as well as one that controls and restricts her dietary provision (i.e., assuming that chimpanzees may see their partner as competitor rather than cooperative partner)? And why would one assume that three infant chimpanzees can understand joint commitment, considering the fact that joint commitment understanding only appears in children around the age of 3 (Warneken et al., 2006)? How can the theory of shared intentionality be warranted, if the findings on which it builds mainly rely on chimpanzees, who are in fact the least cooperative of all apes (Hare et al., 2007; Herrmann, Hare, et al., 2010; Melis et al., 2006b) ? My point is that we should worry about whether the current definition and presumptions about shared intentionality is warranted, theoretically sound, and species-fair.

Further problems concern the reconstruction of the evolutionary scenario in which shared intentionality would have evolved in human evolution. Van Schaik and Burkart (2018) argue that shared intentionality may likely have evolved far earlier than 400ky, as previously suggested (Tomasello, 2016), due to another event about 3 Mya that most likely preceded obligate collaborative foraging: collective predator defense (Ferraro et al., 2013). This vital activity must have involved *joint commitment*, and thus most likely shared intentionality (Boehm, 2018).

Moreover, cooperative hunting in chimpanzees (Boesch, 2002) and in lions (Stander, 1992), and probably many other cooperative hunters, actually fulfill Tomasello's criteria for sharing intentions; hunters flexibly interchange individual hunting roles, and the spoils are commonly shared, with respect to roles, age and rank (Boesch, 1994, 2005). It is surprising why these activities are not considered as governed by shared intentionality (Tomasello, 2016).

Lastly, as outlined in detail in Chapter 1 - the current definition of shared intentionality does not take into account the participants' needs to manage socio-affiliative imperatives inherent to joint action coordination, i.e., by using politeness (Brown & Levinson, 1987; Enfield, 2008; Goffman, 1955; Stivers et al., 2011). Joint actions thus not only feature joint commitment, but also face management, and since Tomasello's definition does not account for that, a revision of the theory is even more warranted. The theory of shared intentionality as it stands is just not helpful at all for any future research interested to study the natural interactions of any species.

To bring new light to the study of shared intentionality, I presented a joint action approach in Chapter 1, which permits the study of joint action coordination in apes, thereby giving new opportunities to research for studying shared intentionality in natural interactions of non-human animal species. In particular, it allows studying shared intentionality as publicly visible achievement, avoids assumptions about recursive mind-reading, offers to study behavioral correlates of face management and joint commitment, can be applied to all age classes, any species, natural interactions, as well as interactions between conspecifics. The central questions of the joint action framework are - *how do participants collaboratively create, maintain and dissolve a potential state of joint commitment in joint action, and while doing so, how to they manage socio-affiliational imperatives*? In this thesis, I focused especially on those joint action phases most prominent for joint commitment understanding in humans – i.e., where joint commitment is brought into existence and publicly negotiated (**entry phases**), suspended and re-established (**suspension and reengagement phases**) and dissolved (**exit phases**). Suggestive evidence for joint commitment in apes would be one to show that apes exchange gaze and/or communicate signals in interaction beginnings (evidence for entry phases), when suspending interactions after interruptions (evidence for suspension phases), when reinstating interactions after having dealt with interruptions (evidence for reengagement phases), and when terminating interactions (evidence for exit phases).

As politeness –as a form of management in humans - serves to regulate the socio-affiliational imperatives inherent to joint action coordination (Arundale, 1999; Brown & Levinson, 1987; Enfield, 2008), I argued that it should be integrated in the joint action framework. It plays a crucial role when participants engage in joint action, as it allows for the smooth coordination of joint action without threatening partners' relationships. As face management has a universal foundation in human joint action with a possible ethological basis, its study is also warranted in

apes. Suggestive evidence for precursors of face management in apes, as I argued in Chapter 1, would be that apes increase their signaling efforts in the phases of entry, suspension, reinstatement, and exit when they are less familiar with partners, or when their partners occupy higher dominance ranks compared to themselves.

To summarize, the joint action framework can be used to study shared intentionality in apes on an ethological, natural basis. In what follows, I will review the main findings of each Chapter, and provide a narrow, focused discussion of the findings. Afterwards, I will attempt to place my findings into the broader picture.

Summary and interpretation of findings

Empirical goal 1

My first empirical goal was to study how bonobos and chimpanzees coordinate naturally occurring social interactions (notably social play and social grooming) with conspecifics (fig. 24). Specifically, I aimed at investigating whether the two species exchange gaze and communicate at the beginning and in the end of their encounters, presuming that, if apes were able of engaging in joint action, they would exhibit entry and exit phases comparable to those in human joint action. Additionally, I aimed at exploring whether the structure of these phases (i.e., duration and likelihood of phase presence), depends on social bond strength and rank difference between partners, presuming that, if apes had precursors to face management, phases should have been shorter and fewer when initiated with strongly bonded / subordinate partners. In addition to these aims, I investigated potential differences between species and activity types, as well as compared the differences in the use of mutual gaze in interaction endings between species and activity types. This study served to address the larger question of whether apes may be capable of collaboratively achieving a state of shared intentionality.

In accord with my predictions, I found that both species frequently exchange gaze and communicative signals when initiating and terminating social activities with peers. This suggests that the apes may solve similar coordination problems than humans when interacting socially, leading to a recognizable sequence of entry and exit phases comparable to the sequence of joint action in humans (Clark, 1996; Heesen et al., 2017). Yet, bonobos exhibited considerably more entry phases (with mutual gaze), and slightly more exit phases (including all possible gaze patterns, but notably more mutual gaze) than chimpanzees. As predicted, phase presence and duration were

affected by the partners' social relationships, although this differed between species. Notably, bonobos were more likely than chimpanzees to produce shorter and fewer phases when interacting with strongly bonded partners, and bonobos produced slightly shorter entry and exit phases when they were more dominant than partners. Chimpanzees produced slightly (yet without statistical effect) fewer exit phases when they were more dominant than partners. Additionally, I found differences in the gaze patterns partners in exit phases, depending on the species and activity type. So-called "mutual exit types" (i.e., exit types that comprised either mutual gaze, or a tactile gesture by one partner and gaze by the other) were more common in bonobos (compared to chimpanzees) and in social play (compared to grooming).

The fact that bonobos and chimpanzees communicate and frequently exchange mutual gaze when initiating interactions is in line with previous studies, and suggests that apes might establish so-called participation frameworks (Rossano, 2013). This entails a shared focus of attention by subjects, with signals, body orientation and behaviors being directed to "ratified participants" (Goffman, 1963). Such social arrangements were previously described for humans (Goffman, 1963), and serve the purpose to demarcate physical boundaries in interactions, by which participants are ratified (or not) a participation status, allowing for the effective coordination of joint action. Moreover, the fact that bonobos and chimpanzees communicate and frequently exchange mutual gaze in exit phases (although less than in entries), possibly suggests that they collaboratively engage in a form of leave-taking - a typical coordination process that aids closing encounters in humans (Albert & Kessler, 1978; Broth & Mondada, 2013; Clark & French, 1981; Schegloff & Sacks, 1973). In humans, leave-taking may involve exchanges of well-wishing expressions such as good-byes (Clark & French, 1981).

Communicating in exit phases also serves to manage interpersonal relationships beyond the encounter, and in humans, reflects the degree of affiliation between participants – it is a way of "reaffirming acquaintances" (Clark & French, 1981, p. 1). For example, in exits of telephone calls, people are likely to exchange goodbyes if they had exchanged helpful or personal information during the call. This means, the more service participants exchange, the more they "owe" to each other in the end, and good bye is one way to picture such a state (Clark & French, 1981). The evidence that exit phases are negatively affected by social bond (and to some extent rank difference) in bonobos suggests that individuals perceive different risks when ending encounters with certain partners – a pattern that reflects the predictions by politeness theory

(Brown & Levinson, 1987), and one that should be studied further in future research. One way to study this in more detail would be to investigate the *type* of signals that bonobos (or chimpanzees) use in joint action phases, and whether these may depend on social relationships (e.g., by signal length, abundancy, or amplitude) in ways predicted by politeness theory.

Lastly, since bonobos were overall more affected by social bond than chimpanzees, and more likely to exhibit entry phases, as well as mutual exit phase types (mutual gaze, or unilateral gaze by one partner and tactile gesture by the other), the coordination of joint activities in bonobos seems to involve more mutual attention by partners. This raises questions and warrants speculation on whether bonobos' interactions involve higher levels of inter-subjectivity than do chimpanzees' interactions. These findings have implications for our understanding of how the two species coordinate interactions and for identifying the kinds of social parameters that may predict how individuals calibrate their communication efforts in species-specific and partner-specific ways. The two species, although more so bonobos than chimpanzees, seem to collaborate to get into and out of interactions with peers, which suggests that they may be capable of achieving a state of shared intentionality through the process of joint action coordination – and thus exhibit a form of joint commitment understanding, as well as possible precursors to face management. Politeness theory (Brown & Levinson, 1987) shall be considered for the study of ape interaction, as politeness has a universal, ethological basis in humans and serves the management of socio-affilational imperatives inherent to joint action (Brown, 2015; Enfield, 2008).

Empirical goal 2

My second empirical goal comprised three empirical studies (fig. 24), and concerned the study of how bonobos and chimpanzees generally cope with interruptions of ongoing social activities, i.e., how they may suspend and reinstate interrupted interactions. Presuming that apes would be capable of engaging in joint action, I expected them to exchange gaze and communicative signals to suspend and re-establish interrupted social activities with social partners.

In my first study (Chapter 4), I looked at whether apes were motivated to engage in a spontaneous triadic tug-of-war game with a human (pulling a garden hose back and forth), and to reengage a partner after he/she suddenly dropped from the game. The paradigm was inspired from a previous experiment, in which an experimenter suddenly transitioned from an active to a passive state while playing a social game with an ape / child (Gruber, 2014; MacLean & Hare, 2013; Pika & Zuberbühler, 2008; Warneken et al., 2006). Yet, in this study, I incorporated a less complex, but

more reciprocal game, and a comparison of age and between species. Although previous studies, focusing on infant chimpanzees, revealed no motivation for reengagement by subjects (Warneken et al., 2006), studies involving subadult bonobos and chimpanzees did evince the subjects' motivation for reengagement (MacLean & Hare, 2013; Pika & Zuberbühler, 2008). I predicted that, if the apes understood joint commitment with a human partner, and if older apes are generally better at it than infant apes, they should readily reengage the partner via communication and game-related behaviors, and older apes should perform better than infant apes. The findings of this study confirmed these predictions; older apes reengaged humans at higher levels than infants, and adults' behaviors / communication used to reengage involved a high degree of game-relatedness.

In my second study (Chapter 5), I looked at whether the behavioral pattern to reinstate interactions may extend to natural interactions with conspecifics. I looked at the likelihood by which partners reinstated previous social activities (play and grooming), the kinds of perturbations causing an interruption (e.g., food-related, environment-related, etc.), species differences, and the motivation to resume previous actions (i.e., continuing to groom the same body part in grooming, and playing the same play type in play) and returning to previous locations. I predicted that, if apes valued the joint commitment with conspecifics, then they would readily reinstate interactions with previous partners, and show some sense of "continuation" by resuming the same actions and at the same location. The findings confirmed my predictions; apes readily reengaged with previous partners (though less so after exposure to severe perturbations, such as food), and were motivated to resume previous actions (i.e., they frequently groomed the same body part, and continued the same play type, after the interruption) at previous locations. However, bonobos were more likely than chimpanzees to reinstate activities, and to continue with the same grooming action.

In my last study (Chapter 6), I ran an experiment to explore the potential joint commitment understanding of bonobos in more controlled conditions. First, I used two interruption stimuli – one "Targeted" and one "Untargeted" interruption stimulus; the Targeted interruption stimulus served to only interrupt one partner (personal responsibility) of the dyad by calling its name, and the Untargeted interruption stimulus served to interrupt both partners (shared responsibility) by the rattling of a nearby door of the housing building. The two stimuli served to study whether individuals would communicate more when the activity interruption was their responsibility (Target) as opposed to shared responsibility (Untargeted). I predicted that if apes engaged in face management, they would increase their coordination efforts to suspend / reengage when being

personally responsible, presumably to mitigate threats inherent to the breaking of the commitment. Second, I interrupted individuals while interacting with strongly vs. weakly bonded partners, to compare the coordination efforts between dyads of strong vs. weak bonds. If apes had precursors to face, I expected bonobos to follow politeness patterns (Brown & Levinson, 1987), thus to communicate more during suspension and reengagement phases when interacting with weakly (vs. strongly) bonded and more dominant (vs. subordinate) partners. Moreover, I compared individuals' likelihood to reinstate a previous activity if the activity was social vs. solitary. I expected that, if bonobos were truly motivated by joint commitment, they should resume social grooming much more than solitary activities (self-grooming, solitary play), presuming that the breaking of social interactions features possible "face threats" inherent to joint commitments. Lastly, I investigated the bonobos' understanding of their own responsibility within the interaction, by looking at whether coordination efforts may increase with an individual's responsibility. I predicted that, if bonobos managed face, then they would communicate more during suspension and reengagement phases – and potentially reengage more overall – when they were the ones responsible for suspending the interaction (vs. not responsible), for initiating the interaction (vs. receiving invitation), and receiving grooming (vs. giving grooming).

My findings confirmed several of these predictions. First, individuals communicated more (for reengagement), when they were interrupted by a Targeted interruption stimulus, thus when it was their responsibility to leave and to let the partner wait, than when interrupted by an Untargeted interruption stimulus, when it was a shared responsibility. Second, individuals communicated more (for suspension) when they suspended an activity with a weakly bonded, or a more dominant partner, potentially reflecting the patterns predicted by politeness theory (Brown & Levinson, 1987). Individuals were also much more likely to reinstate social grooming (so carried out with a partner) than solitary activities (carried out alone, e.g., self-grooming, solitary play), suggesting that social interactions feature certain risks (possibly linked to joint commitment), that may drive individuals to return to their previous social partner. I did not find evidence that individuals communicated more during suspension phases when they were responsible for the suspension, activity initiation or when they received grooming. Nonetheless, I found evidence that they reengaged partners more (i.e., by returning back to the partner to resume grooming with or without communication) when they had been the initial recruiters of the interaction, compared to recruitees, and when they had been the active partner ("groomer"); moreover, I found that individuals

communicated more for reengagement when they had been responsible for suspension and for initiating the interaction.

All these findings from Chapter 4-6, together, suggest that apes likely have an understanding of joint commitment, and that bonobos might have precursors of face management. Joint commitment has been long thought to be human unique (Bratman, 1992; Tomasello & Carpenter, 2007), and was previously assessed within the top-bottom approach of shared intentionality – as one of many critical cognitive abilities that allow humans to engage in joint actions with shared goals (Tomasello, 2016). I showed that apes exhibit reactions indicative of joint commitment understanding, both when interacting with a human partner in laboratory environments (Chapter 4), as well as when interacting with conspecifics in natural settings (Chapter 5-6). This refutes the idea that only humans have shared intentionality (as cognitive ability **and** as interactional achievement), and, seen from a joint action perspective – bonobos and chimpanzees have the critical prerequisites to coordinate joint actions via shared intentionality. This implies that shared intentionality can manifest itself in the joint actions of apes, and validates the joint action framework as a method to studying this capacity in non-human animal species.

Behavioral patterns, possibly suggestive of forms of face management beyond humans, have previously not been studied in any non-human animal species, although they may serve as yardstick to assess behavioral evidence for politeness efforts beyond humans. Politeness is a form of face management in humans, by which individuals regulate each other's public identity or self-esteem (Brown, 2015, p. 327), and is a universal phenomenon that occurs across cultures. This warrants its assessment in animal behavior, notably that of our closest relatives. My data suggests that bonobos may follow the patterns predicted by politeness theory (i.e., partners communicate more when socially distant or different in power), and increase their efforts when they are to blame for the threat of integrity. However, it is reasonable to pose the question of why bonobos may at all engage in a form of face management, i.e., what functions may such behavior serve? It is admittedly speculative to assume that bonobos have negative face wants (want not to offend others) or positive face wants (want to be approved), but the data suggests that they may tackle similar problems inherent to face management. At the very least, bonobos seem to have a sort of "appeasement management", by which partners increase their coordination efforts to avoid getting in trouble and reduce risks while being jointly committed. This comes to use especially while interacting with higher-ranking partners who could potentially sanction oneself (via aggression

and restriction of resources), and while interacting with socially less familiar individuals whose reactions may be less predictable. Politeness theory, as an ethological phenomenon, may thus lie at the heart of what it means to deal with the risks of life in a social group where participants have normative expectations about how their equals should be treated (and not). Nonetheless, even though I could show that apes follow politeness theory patterns, one critical condition of face management which I did not address in my current thesis is the public aspect of reputation management. I intentionally used the expression "precursors of face management" when speaking about face in apes, because managing face includes – next to the dyadic component – a necessary public component: an audience who is present and who could potentially judge one's reputation. Since the latter aspect has not been assessed, it may be prematurely to account for face management in apes. However, there is indeed evidence that apes do take into account the audience composition when interacting with a social partner; for instance, apes indeed have been observed to regulate gestural and vocal communication according to who is present around them (Cartmill & Byrne, 2007; Slocombe & Zuberbühler, 2007). Therefore, studying whether joint action coordination is affected by an audience (i.e., and crucially, whether coordination efforts increase along with the importance of the audience members), would represent a fruitful opportunity for future research to address whether apes manage face indeed.

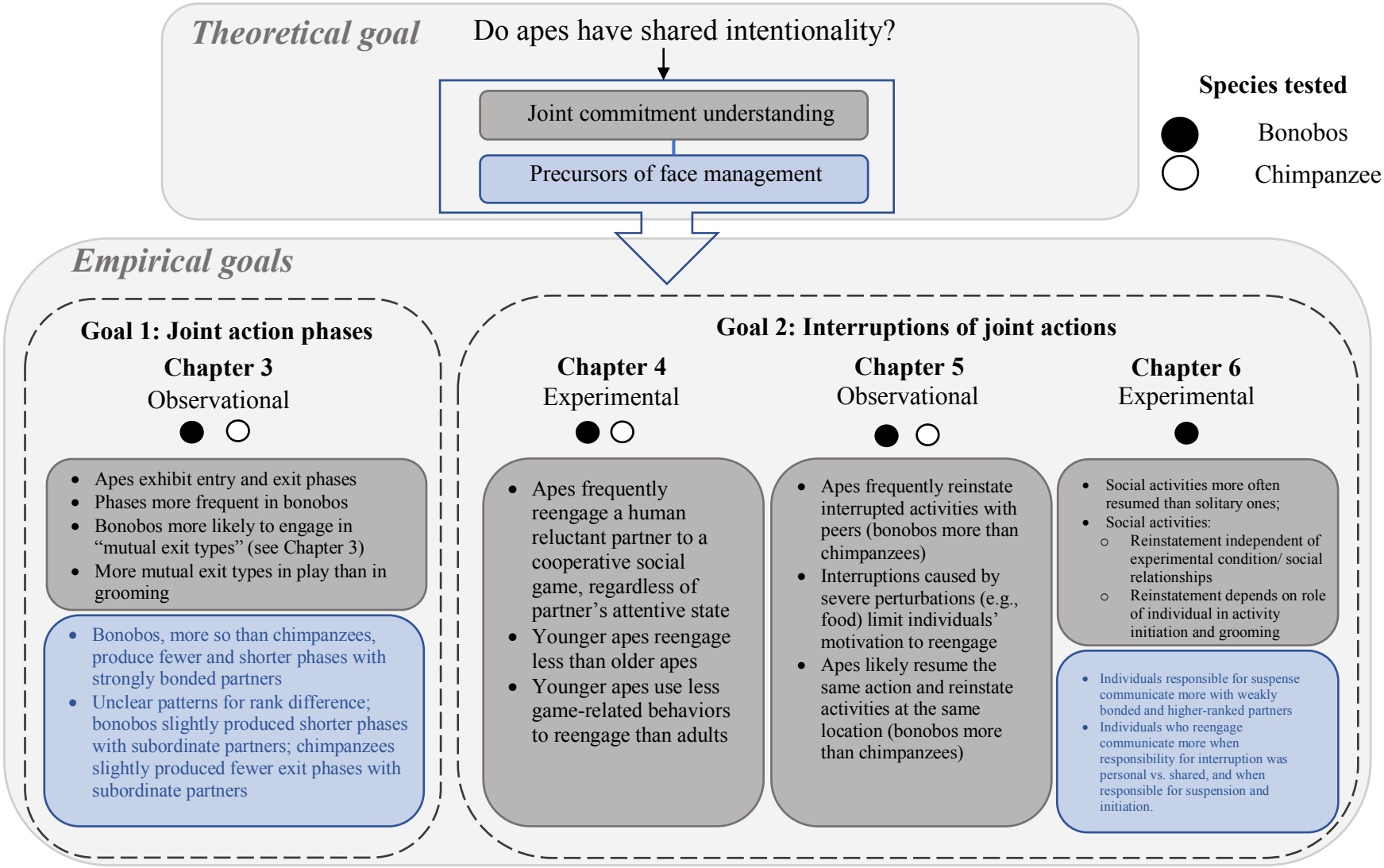


Figure 24 A synthesis of thesis findings, summarizing the overarching question, the goals and most important research findings of each Chapter. The grey boxes indicate suggestive evidence for joint commitment understanding, and the blue boxed indicate suggestive evidence for patterns related to face management.

The broader picture: Implications, limitations and future avenues

In this section, I will discuss the broader implications, limitations, and ideas resulting from my findings. I will place my work with respect to the classical, philosophical definition of shared intentionality, and discuss how we should revise the evolutionary story of shared intentionality. To do so, I address five major points of discussion, starting with what the findings tell us about the two species' joint action capacities, and ending with some points to be considered for a possible revision of the theory of shared intentionality (with future questions).

Great ape joint action capacities

Apes are able to read intentions (Call et al., 2004), act according to other's apparent goals (Warneken & Tomasello, 2006), provide help (Melis & Tomasello, 2013), communicate their own intentions (Hobaiter & Byrne, 2011b) with flexible adjustment to the partner's social attributes (Fröhlich, Wittig, et al., 2016a), and, perhaps most strikingly, recognize false beliefs (Krupenye et al., 2016). Based on this portfolio of capacities it seems reasonable to hypothesize that they are also able to engage in collaborative activities with shared goals featuring shared intentionality. Yet still, the current theory asserts that ape interactions are based on egoistic, individual intentions, with the goal to use partners as social tools, and thus with no evidence for shared intentionality (Bullinger, Melis, & Tomasello, 2011; e.g., Völter, Rossano, & Call, 2015).

My thesis' findings challenge the current theory. When looking at shared intentionality from a joint action perspective (*as interactional achievement*), chimpanzees and bonobos reveal necessary precursors to shared intentionality – evinced by their understanding of joint commitment and precursors to face management. These findings raise many questions regarding the current perception of what constitutes a human-unique form of shared intentionality, as they suggest that apes are capable of joint action. As a reminder, joint action is the process by which two participants collaborate to achieve a shared goal (Clark, 1996), and is characterized by a state of what has been variously termed as collective intentionality (Searle, 1990), interacting in the "we" mode (Tuomela, 2007), inter-subjectivity (Zlatev, Racine, Sinha, & Itkonen, 2008), or shared intentionality (Heesen et al., 2017). Since apes structure their interactions in orderly ways comparable to the joint action in humans with entry, main body and exit (Clark, 1996), routinely reengage reluctant partners, and adjust communicative acts according to their partners' social

attributes – a potential form of precursor of face management (Brown, 2015)- my findings suggest that apes engage in joint action via shared intentionality.

Yet, one may ask the fair question, if apes have the necessary building blocks for shared intentionality as interactional achievement, then why do they fail at producing those large-scale interactions that endure in time, and increase in complexity, with participants' contribution? Admittedly, ape interactions do not reach a comparable complex, cumulative structure as seen in human joint activities, but remain rather short-termed with little evidence for increased complexity over time (e.g., grooming and play).

One explanation could be that apes are *cognitively capable* to reach higher levels of sophistication in joint action, yet lack the *social motivation* to do so, simply because they do not need to (or due to differences in neurological reward systems). The most complex interactions in apes, where participants work towards a collaborative goal, are probably group hunting (Boesch, 2002) and short-term projects such as boundary patrolling (Watts & Mitani, 2001) and consortship (Hobaiter & Byrne, 2012). It is possible that apes may only display the motivation to perform more complex joint activities if the personal gain is large enough (Bullinger et al., 2011), as it seems to be the case for these activities (Boesch, 2002). Indeed, Boesch (Boesch, 2005) argued that group hunting *must* be considered as joint activity, and criticizes the disdain by experimental psychology of natural observations from the wild.

A further explanation as to why apes do not engage in more complex joint action forms – presuming they are capable of reaching a state of shared intentionality - may be that there are other relevant cognitive abilities in humans that allow for engagement in sophisticated long-term projects, notably the ability for episodic memory (Suddendorf & Corballis, 2007). This form of mental time travel permits – among other things - the planning and simulation of future events (Suddendorf, Addis, & Corballis, 2009) and may allow humans to achieve long-term collaborative outcomes through collective planning. Indeed, so far, there is no convincing evidence that any non-human animal species can plan for the future and form long-term goals (Suddendorf & Corballis, 2007).

When comparing the degree of coordination between play and grooming activities in apes, I found that play generally involved higher levels of "jointness" than grooming. Notably, play (as opposed to grooming and mixed contexts, where dyads switch between difference activities) was the context with the highest use of mutual exit types. These types of exits involved mutual gaze

(partners looking at each other's face) or unilateral gaze by one partner, and a tactile gesture by the other. This suggests that joint play may have a more urgent need for a mutual recognition and agreement in closing, where both partner are somehow aware and share attention when disengaging. This would also make sense, because social play involves more rough movements and intense body contact, rendering its participants potentially vulnerable to aggression by a partner (i.e., play fighting may easily escalate into aggression) (Croft & Snaith, 1990; P. K. Smith & Boulton, 1990). Closing play must therefore be regulated and mutually agreed upon, especially in risky situations, to delineate the previous playful activity from whatever follows after. It is therefore that play applies as particularly adequate candidate for studying joint action coordination in non-human animal species (Heesen et al., 2017; Palagi et al., 2015).

An important final point of discussion is that bonobos and chimpanzees generally differed in their degree of coordination in joint action. Bonobos overall seemed more "socially driven" than chimpanzees, evinced by their increased likelihood to communicate in entry and exit phases, to engage in mutual exit types, to reinstate previous partners more frequently after interruptions, and to base communication efforts on social bond strength – a behavioral species difference that may be explained by the species' ecology and social organization. By comparing the joint action capacities in these two species of ape, we may therefore learn something about the kinds of societal and environmental factors promoting sophisticated joint action skills; in the next section, I will follow up on these points, and discuss the implications of my findings regarding the species differences found in bonobos and chimpanzees.

Understanding joint action by a species' ecology

My findings revealed that the joint actions of bonobos generally seem to be more coordinated, and involve higher levels of "jointness" than the joint actions of chimpanzees. They tended were more affected by the social bond with partners, more likely to exhibit entry and exit phase (especially mutual exit types), more likely to reinstate previous activity with peers, as well as to resume previous actions in grooming when reinstating grooming activities with peers. Why may this be?

Bonobos may behave more cooperatively with their partners, because they experience less competition in their environment, and because they exhibit socio-sexual behaviors that promote cooperation via oxytocinergic effects (Moscovice et al., 2019). For example, the natural habitat of

bonobos is characterized by larger food-patch sizes, and more "feed as you go" foraging (Furuichi, 2009; White & Wrangham, 1988), and their favored foods are terrestrial herbaceous vegetation – a food source which is depleted slowly and hardly monopolizable (Furuichi, 2009). Bonobo societies are also characterized by lower levels of aggression (see Gruber & Clay, 2016 for review), by their heightened levels of social tolerance even in adulthood (Hare et al., 2007; Wobber, Wrangham, & Hare, 2010). Bonobos have a more sophisticated understanding of tasks related to the social world (Herrmann, Hare, et al., 2010), and increased pro-sociality, even towards unknown individuals (Tan et al., 2017; Tan & Hare, 2013). The societal organization and environmental conditions might thus have an undue influence on the cooperative tendencies of the two species, in that cooperative environments (with less resource competition) and more egalitarian dominance styles may foster the emergence of more sophisticated joint action skills.

These species differences may be helpful for our understanding of the ecological niche in which joint action may have evolved. Overall, bonobos revealed slightly more pronounced coordinative efforts than chimpanzees, and higher levels of partner sensitivity. Although it is questionable whether bonobos truly have more sophisticated joint action skills than chimpanzees, it seems likely that they are *more socially motivated, and more appreciative about their partners' social attributes*. My findings suggest that chimpanzees and bonobos share the ability to coordinate phases, to reengage after interruptions, and to adjust the communication according to partner's social relationships, but that bonobos are just *more likely* to do so (e.g., they were, for example, *more* affected by social bond than chimpanzees). This suggests that both species share the cognitive make up for the building blocks necessary for shared intentionality, but that they differ in their socio-motivational force to create "togetherness" in joint action. It is possible that an environment / society akin to that of bonobos favors the evolution of cooperation. Indeed, bonobos presumably are less aggressive than chimpanzees as they live in environments with relaxed feeding competition, and this has been proposed to be driver that led bonobos to undergo a process of "self-domestication" (Hare, Wobber, & Wrangham, 2012), by which male aggression towards females and other group members decreases survival benefits. This process supposedly was responsible for the morphological, behavioral, psychological and neurobiological differences between bonobos and chimpanzees. These reasons may contribute to the higher levels of coordination in the joint actions of my bonobo subjects, and raise interesting questions about whether the evolution of joint action capacities, notably joint commitment and precursors to face management, may be

linked to such forms of domestication processes. In what follows, I will now take a step back, in order to discuss the broader implications of all of these findings for a theory of shared intentionality.

My findings' implications regarding the classical theory of shared intentionality

Although my findings cannot ascertain which kinds of cognitive states may be involved when participants engage in joint action, they may offer a glimpse into the kinds of alignment requirements minimally needed to engage in joint action, aiding to establish a mutual perception of joint commitment and "politeness rules". What I showed at the least is that **apes do what children do** when a partner suddenly breaks off from a social activity (Gräfenhain et al., 2009, 2013). They readily reengage, via communication, and take leave when they have to drop from a commitment – these points constitute as evidence for shared intentionality - even according to Tomasello's classical definition (Tomasello & Carpenter, 2007). Additionally, I showed that when interacting jointly, apes have some appreciation of their partners' social attributes, as evinced by their flexible adjustments and increase of coordination efforts for less well known and higher ranked partners. This pattern reflects on how politeness is used in humans: politeness efforts increase for less well known and more powerful, high-status partners (Brown & Levinson, 1987; Goffman, 1955). Although using the term face management for apes may be inadequate, I nonetheless provided evidence that communication is calibrated in similar ways as humans design their language utterances towards others when they engage in joint action; my findings thus suggest that apes have precursors of face management. These findings support the idea that apes engage in joint action – and hence shared intentionality, in both the classical sense (Tomasello & Carpenter, 2007), but also according to the joint action approach.

However, from the classical point of view, one could argue that, since my findings concern only one necessary correlate that was classically defined as *one of many cognitive abilities* governed by shared intentionality (notably, joint commitment), apes only have a less complex version of shared intentionality. There are ample other cognitive abilities claimed to be involved in the human version of shared intentionality, such as cooperative communication or joint attention (Tomasello, 1995), and indeed previous studies draw a rather skeptical picture regarding the presence of these skills in apes (see Tomasello & Carpenter, 2007 for review). So by having found evidence for joint commitment and precursors of face management, can we assume that apes have

shared intentionality, or do they need to express all of the other traits as well? Either we admit they have shared intentionality – at least in the form of the most minimal version including joint commitment, which Tomasello defines as "joint intentionality" (Tomasello, 2014, 2016), or we refine the notion of shared intentionality completely. We may instead start with a definition of shared intentionality that involves those kinds of building blocks that a species minimally needs to engage in joint action – which I suggest are joint commitment and precursors of face management, possibly achieved through lower-level alignment processes and signal exchanges in joint action. From there, the definition may become more complex, with added layers, when interactions increase in complexity. Certainly, joint commitment and precursors of face management should be considered conditional ingredients, and due to their presence in apes their evolutionary origins must be revisited.

Although this question goes beyond the scope of my thesis, and the attempt is admittedly speculative, I would like to draw on a possible evolutionary scenario, which I hope will spark ideas for future revisions. Assuming that apes have shared intentionality, one alternative evolutionary scenario could be as follows. Shared intentionality, involving both joint commitment toward a shared goal and precursors of face management in joint action, has already been in place in our last common ancestor, due to events where these skills were most needed (group hunting and managing reputation in a social group). Other, more complex aspects of shared intentionality, such as cooperative communication with mimicking and pointing, as well as cultural instructed learning, evolved later, with the rise of group size and cultural activities. Under this assumption, shared intentionality – just as human communication (Levinson & Holler, 2014), is layered and would have evolved gradually, with the first joint action capacities (joint commitment, precursors of face management) already present in our last common ancestor at about 8Mya.

To sum up, my findings raise concerns and perhaps confusion about our current understanding of what constitutes a human-unique version of shared intentionality. When studying shared intentionality as the process of joint action coordination, apes reveal similar joint action capacities to humans, notably those previously considered as definite evidence for shared intentionality – the motivation for active partner-reengagement after disrupted social activities (e.g., Gräfenhain et al., 2009, 2013; Kachel et al., 2017; Kachel & Tomasello, 2019; Warneken et al., 2006). It is timely to integrate my findings, and to revise the human-uniqueness claim.

An attempt of re-thinking human uniqueness claims

Apes produce behaviors while engaging in joint action that fulfill the necessary preconditions of joint commitment understanding and face management. Previous studies also revealed their ability to read intentions and infer others' goals (Buttelmann et al., 2012; Call et al., 2004; Call & Tomasello, 1998), understand false-beliefs (Buttelmann et al., 2017; Krupenye et al., 2016), understand complementarity of roles (Fletcher et al., 2012), offer help (Greenberg et al., 2010; Warneken et al., 2007), inform others about danger (Crockford et al., 2012), use gestures with spatial reference (Genty & Zuberbühler, 2015), and base communication on interactive histories with partners (Genty et al., 2015a).

Therefore, it is warranted to hypothesize that shared intentionality, in a minimal form, has not emerged at first in the human lineage. Instead, according to the current findings, it seems more reasonable to describe shared intentionality as a layered complex, with each of the layers having evolved at different times (see for instance Fröhlich, Sievers, et al., 2019; Levinson & Holler, 2014). The emergence of shared intentionality, at least with the aforementioned abilities, must have evolved far earlier than 400kya, and the concept of joint commitment should perhaps no longer be considered human-unique (Tomasello & Carpenter, 2007). Other abilities linked to a more complex version of shared intentionality – *collective intentionality* (Tomasello, 2016), along with other potentially human-unique abilities such as mental time travel (Suddendorf et al., 2009; Suddendorf & Corballis, 2007) may have evolved solely in the lineage of *homo*, and may be responsible for higher-level collaborative joint projects typical for human modern life.

With these points, I hope to have induced a though-provoking discussion and hinted at some ideas for a revision of the shared intentionality hypothesis. At the very least, my thesis may have sparked a motivation for scholars to revisit the shared intentionality hypothesis– if as cognitive capacity, interactional achievement, or whichever ways it may best be defined as. Although my findings might influence the current view of shared intentionality, they also feature some limitations and bring about new interesting avenues for future research, which I wish to draw attention to in the following section.

Limitations and future questions

We now know that captive bonobos and chimpanzees collaboratively coordinate the entry, reinstatement and exit of social activities, and that at the least bonobos manage partners' face

through flexible communicative adjustments to social matrices and their own responsibility for joint success. We also know that bonobos generally do these things more so than chimpanzees. What we do not know, however, is how they do this in wild settings (future avenue 1), whether these behavioral outputs involve the same socio-cognitive skills than in humans (future avenue 2), and whether these joint action capacities are unique to *Pan* and *Homo*, or whether they are shared by *other* ape species– or even more distantly related species (future avenue 3).

To the first avenue. The captive study groups are housed in a closed environment, where individuals encounter each other every day and cope with competitive situations potentially leading to aggression, such as feeding (Bloomsmith, Laule, Alford, & Thurston, 1994). In humans, a condition by which people remain in the same spatial configuration in which their interactions occur – a so-called "continuing state of incipient talk" (Schegloff & Sacks, 1973, p. 325)- can lead to the reduction of an exit phase. In this scenario, it is often superfluous to coordinate an exit phase, since people have continuing opportunities to interact (e.g., people sharing an office and not constantly saying good-bye). This continuing state may also apply to the captive groups of bonobos and chimpanzees representing my dataset. On the one hand, one could argue that, since I still found a considerable likelihood of exit phases in these zoo groups, one should definitely find exit phases in fission-fusion societies in the wild, because there individuals have less opportunities to remain in a such a continuing state.

On the other hand, however, managing relationships in an enclosed environment may actually demand for more coordination efforts by participants, precisely *because* they see each other repeatedly. The breaking of a social relationship in a zoo environment may have severe consequences for future interactions, especially during feeding. It is therefore equally possible that wild apes exhibit less of the joint action capacities as presented in my studies. Understanding whether wild and captive groups differ in their joint action capacities (joint commitment and face management) would be crucial for understanding whether these capacities have a phylogenetic or ontogenetic basis (or both). To show evidence for a phylogenetic basis, one would need to find that these capacities are expressed universally in any type of environment and across groups, and have little room for being learnt. To show evidence for an ontogenetic basis, one would need to find that that these capacities are learnt throughout individuals' development and triggered by certain environmental factors; if so, they may vary in form and function across groups (Halina et al., 2013).

To the second avenue for future research. I cannot ascertain that the signals apes use to coordinate joint action phases serve the same function as in humans. For example, I cannot say with certainty whether the apes' signals used in the exit phases of joint activities really serve a leave-taking function, so to coordinate on terminating the activity. It is well possible that these signals have other functions; they may be a result of a failed personal goal of the subject, for example to make the other partner do something else. This is a limitation of my research, and future research needs to address this. I propose two potential ways of how this may be done. In the example of "leave-taking signals", one could compare the type of signals used for *exits* of activities to those signals used for *initiation* or *reinstatement* of activities. If they differed, this would suggest that the signals indeed serve unique functions. If the signal types used in *exits* are identical to those used for *initiation* and *reinstatement*, one could compare the signals' apparently satisfactory outcomes ("ASOs") by looking at the observable change in a recipient's behavior that presumably stops the signaler from signaling further (e.g., Hobaiter & Byrne, 2014, 2017). ASOs represent some plausible functions of signalers' signals (i.e., what do signalers mean with their signals?). For instance, if a signaler communicates to terminate an activity, recipients should stop interacting after receiving the signal. If a signaler communicates to reengage a partner, recipients should come back and resume the activity after receiving the signal.

Finally, to the third future avenue for future research. One reason for why I studied chimpanzees and bonobos especially was due to their genetic closeness to humans. The comparison between these species and humans delivers informative results regarding the reconstruction of the last common ancestor, and hence the scenario in which shared intentionality might have evolved (Tomasello, 2016). But what about the joint action capacities of other great ape species, notably those of gorillas and orangutans? In our last paper (Heesen et al., 2017), we listed some signals of gorillas (*Gorilla gorilla*) with potential service for joint action coordination in both entry and exit phases. In entries, gorillas enlist social partners via gestures that have some apparent effect on the following type of action; for example, whilst *one-handed grab* gestures seem to be used for initiating contact play, *drum object* gestures seem to be used for initiating chase-play (Genty et al., 2009). To exit from play, gorillas frequently use hand-on gestures of body signals, such as pirouetting (Genty et al., 2009; Luef & Liebal, 2013). To understand whether shared intentionality uniquely evolved in the last common ancestor between *Pan* and *Homo*, it is also crucial to study the joint action capacities of other great apes. Taken to the broader level, the

joint action framework may serve to study shared intentionality in other animal species, especially in those species where the successful coordination of joint action would increase fitness benefits, such as in cooperative breeders (e.g., arabian babblers: Carlisle & Zahavi, 1986) or cooperative hunters (e.g., killer whales: Smith et al., 1981). These comparisons would provide invaluable data for the understanding of what makes human joint action special (Heesen et al., 2017), and whether joint commitment and precursors to face management may be uniquely *Pan* / *Homo*.

Conclusion

Shared intentionality – i.e., the ability to share psychological states - has previously been considered a uniquely human capacity (Tomasello, 2016), supposedly enabling us to engage in sophisticated joint action forms, and paving the way for culture and language. Previous definitions of shared intentionality have circulated around the idea that sharing intentions is manifested in cognitive abilities such as cooperative communication, joint triadic attention and joint commitment (Tomasello et al., 2005). In my thesis, I have attempted to study shared intentionality in apes by taking an alternative, ecologically more valid joint action approach. This joint action approach allows to study and demonstrate shared intentionality by the coordinative *process* of joint action – an alternative way of identifying whether a species is capable of sharing intentions, and one that goes beyond the indicators previously mentioned in the literature. My findings suggest that apes are capable of collaborating to create, re-establish and dissolve a form of joint commitment and that, at the least bonobos, are sensitive towards social relationships with partners – indicating a potential precursor of face management in humans (Brown & Levinson, 1987). These findings suggest that apes have some of the building blocks necessary for shared intentionality.

Future questions should profit from the proposed joint action framework when studying shared intentionality in the natural interactions of any species. I argue that the question of how shared intentionality may have evolved should resolve around the question of *degree*, rather distinction. In other words, shared intentionality may be gradual – a layered complex (Levinson & Holler, 2014) - which gained layers with additional cooperative abilities in modern humans. More research should focus on dissecting these kinds of human-unique cooperative abilities, to determine which of those represent the main highway toward modern human cultural life and language. I conclude by emphasizing the potential for a *gradual evolution* of shared intentionality (Fröhlich, Sievers, et al., 2019; Levinson & Holler, 2014), and that natural selection has shaped human uniqueness from ancestral primate roots. Altogether, my thesis challenges the assumption that shared intentionality

evolved uniquely with the genus *Homo*, and has delivered a theoretical approach that allows for the study of shared intentionality in a natural joint action framework in any animal species. I hope my thesis will revive the debate on what counts as shared intentionality

References

Albers, P. C. H., & de Vries, H. (2001). Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour*, *61*(2), 489–495. https://doi.org/10.1006/ANBE.2000.1571

Albert, S., & Kessler, S. (1976). Processes for ending social encounters: The conceptual archaeology of a temporal place. *Journal for the Theory of Social Behaviour*, *6*(2), 147–170. https://doi.org/10.1111/j.1468-5914.1976.tb00363.x

Albert, S., & Kessler, S. (1978). Ending social encouters. *Journal of Experimental Social Psychology*, *14*(6), 541–553. https://doi.org/10.1016/0022-1031(78)90048-3

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model. *Cognitive Science*, *26*(1), 39–83. https://doi.org/10.1207/s15516709cog2601_2

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, *49*(3), 227–266. https://doi.org/10.1163/156853974X00534

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, *49*(6), 1053–1076. https://doi.org/10.1086/593015

Arnold, K., & Whiten, A. (2003). Grooming interactions among the chimpanzees of the Budongo Forest, Uganda: tests of five explanatory models. *Behaviour*, *140*(4), 519–552. https://doi.org/https://doi.org/10.1163/156853903322127968

Arundale, R. B. (1999). An alternative model and ideology of communication for an alternative to politeness theory. *Pragmatics*, *9*(1), 119–153. https://doi.org/10.1075/prag.9.1.07aru

Baarendse, P. J. J., Counotte, D. S., O'Donnell, P., & Vanderschuren, L. J. M. J. (2013). Early social experience is critical for the development of cognitive control and dopamine modulation of prefrontal cortex function. *Neuropsychopharmacology*, *38*(8), 1485–1494. Retrieved from http://dx.doi.org/10.1038/npp.2013.47

Bangerter, A., Chevalley, E., & Derouwaux, S. (2010). Managing third-party interruptions in conversations: Effects of duration and conversational role. *Journal of Language and Social Psychology*, *29*(2), 235–244. https://doi.org/10.1177/0261927X09359591

Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science*, *27*(2), 195–225. https://doi.org/10.1016/S0364-0213(02)00118-0

Bangerter, A., Clark, H. H., & Katz, A. R. (2004). Navigating joint projects in telephone conversations. *Discourse Processes*, *37*(1), 1–23. https://doi.org/10.1207/s15326950dp3701_1

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*(1), 107–122. https://doi.org/10.1017/S0140525X00041790

Becquet, C., Patterson, N., Stone, A. C., Przeworski, M., & Reich, D. (2007). Genetic structure of chimpanzee populations. *PLoS Genetics*, *3*(4), e66. https://doi.org/10.1371/journal.pgen.0030066

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Bekoff, M. (2004). Wild justice and fair play: cooperation, forgiveness, and morality in animals. *Biology and Philosophy*, *19*(4), 489–520. https://doi.org/10.1007/sBIPH-004-0539-x

Bekoff, M., & Allen, C. (1998). Intentional communication and social play: How and why animals negotiate and agree to play. In M. Bekoff & A. Byers (Eds.), *Animal play: Evolutionary, comparative, and ecological perspectives* (pp. 97–114). Cambridge: Cambridge University Press.

Bermúdez, J., & Cahen, A. (2015). "Nonconceptual Mental Content." In *The Stanford Encyclopedia of Philosophy*.

Bloomsmith, M. A., Laule, G. E., Alford, P. L., & Thurston, R. H. (1994). Using training to moderate chimpanzee aggression during feeding. *Zoo Biology*, *13*(6), 557–566. https://doi.org/10.1002/zoo.1430130605

Boehm, C. (2018). Collective intentionality: A basic and early component of moral evolution. *Philosophical Psychology*, *31*(5), 680–702. https://doi.org/10.1080/09515089.2018.1486607

Boesch, C. (1994). Cooperative hunting in wild chimpanzees. *Animal Behaviour*, *48*(3), 653–667. https://doi.org/10.1006/anbe.1994.1285
Boesch, C. (2002). Cooperative hunting roles among taï chimpanzees. *Human Nature*, *13*(1), 27–46. https://doi.org/10.1007/s12110-002-1013-6
Boesch, C. (2005). Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously. *Behavioral and Brain Sciences*, *28*(5), 692–693. https://doi.org/10.1017/S0140525X05230121
Boesch, C. (2009). *The real chimpanzee: sex strategies in the forest*. Cambridge: Cambridge University Press.
Boesch, C., Hohmann, G., & Marchant, L. F. (Linda F. (2002). *Behavioural diversity in chimpanzees and bonobos*. Cambridge University Press.
Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171–195. https://doi.org/10.1016/0162-3095(92)90032-Y
Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327–341. https://doi.org/10.2307/2185537
Bratman, M. E. (1993). Shared Intention. *Ethics*, *104*(1), 97–113. https://doi.org/10.1086/293577
Broth, M., & Mondada, L. (2013). Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics*, *47*(1), 41–58. https://doi.org/10.1016/j.pragma.2012.11.016
Brown, P. (2015). Politeness and language. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 326–330). Elsevier. https://doi.org/http://dx.doi.org/10.1016/B978-0-08-097086-8.53072-4
Brown, P., & Levinson, S. C. (1978). Universals in language usage: Politeness phenomena. In E. N. Goody (Ed.), *Questions and politeness: Strategies in social interaction* (pp. 56–311). Cambridge University Press.
Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
Bullinger, A. F., Melis, A. P., & Tomasello, M. (2011). Chimpanzees, *Pan troglodytes*, prefer individual over collaborative strategies towards goals. *Animal Behaviour*, *82*(5), 1135–1141. https://doi.org/10.1016/J.ANBEHAV.2011.08.008
Bürkner, P.-C. (2017). brms: an R package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01
Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false beliefs in an interactive helping task. *PLOS ONE*, *12*(4), e0173793. https://doi.org/10.1371/journal.pone.0173793
Buttelmann, D., Schütte, S., Carpenter, M., Call, J., & Tomasello, M. (2012). Great apes infer others' goals based on context. *Animal Cognition*, *15*(6), 1037–1053. https://doi.org/10.1007/s10071-012-0528-4
Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition*, *20*(4), 755–769. https://doi.org/10.1007/s10071-017-1096-4
Call, J. (2009). Contrasting the social cognition of humans and nonhuman apes: The shared intentionality hypothesis. *Topics in Cognitive Science*, *1*(2), 368–379. https://doi.org/10.1111/j.1756-8765.2009.01025.x
Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Animal Cognition*, *8*(3), 151–163. https://doi.org/10.1007/s10071-004-0237-8
Call, J., Hare, B. A., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition*, *1*(2), 89–99. https://doi.org/10.1007/s100710050013
Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). "Unwilling" versus "unable": chimpanzees' understanding of human intentional action. *Developmental Science*, *7*(4), 488–498.

https://doi.org/10.1111/j.1467-7687.2004.00368.x
Call, J., & Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*, *112*(2), 192–206. https://doi.org/10.1037/0735-7036.112.2.192
Call, J., & Tomasello, M. (2007). *The gestural communication of apes and monkeys*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192. https://doi.org/10.1016/J.TICS.2008.02.010
Carlisle, T. R., & Zahavi, A. (1986). Helping at the nest, allofeeding and social status in immature arabian babblers. *Behavioral Ecology and Sociobiology*, *18*(5), 339–351. https://doi.org/10.1007/BF00299665
Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants? In J. Metcalfe & H. S. Terrace (Eds.), *Agency and Joint Attention* (pp. 49–61). New York: Oxford University Press.
Cartmill, E. A., & Byrne, R. W. (2007). Orangutans modify their gestural signaling according to their audience's comprehension. *Current Biology*, *17*(15), 1345–1348. https://doi.org/10.1016/J.CUB.2007.06.069
Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893
Chevalley, E., & Bangerter, A. (2010). Suspending and reinstating joint activities with dialogue. *Discourse Processes*, *47*(4), 263–291. https://doi.org/10.1080/01638530902959935
Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
Clark, H. H. (2006). Social actions, social commitments. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 126–150). Oxford, England: Berg.
Clark, H. H., & French, J. W. (1981). Telephone goodbyes. *Language in Society*, *10*(1), 1–19. https://doi.org/10.1017/S0047404500008393
Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press.
Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7
Clay, Z. (2011). *Vocal communication in bonobos (Pan paniscus): studies in the contexts of feeding and sex. Dissertation.* University of St Andrews.
Clay, Z., Archbold, J., & Zuberbühler, K. (2015a). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, *3*, e1124.
Clay, Z., Archbold, J., & Zuberbühler, K. (2015b). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, *3*, e1124. https://doi.org/10.7717/peerj.1124
Clay, Z., & de Waal, F. B. M. (2013). Bonobos respond to distress in others: consolation across the age spectrum. *PLoS One*, *8*(1), e55206.
Clay, Z., & de Waal, F. B. M. (2015). Sex and strife: post-conflict sexual contacts in bonobos. In *Bonobo Cognition and Behaviour* (pp. 67–88). Brill.
Clay, Z., Pika, S., Gruber, T., & Zuberbühler, K. (2011). Female bonobos use copulation calls as social signals. *Biology Letters*, rsbl20101227.
Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, *77*(6), 1387–1396.
Crockford, C., Gruber, T., & Zuberbühler, K. (2018). Chimpanzee quiet hoo variants differ according to context. *Royal Society Open Science*, *5*(5), 172066. https://doi.org/10.1098/rsos.172066
Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, *110*(3), 221–243. https://doi.org/10.1111/j.1439-0310.2004.00968.x

Crockford, C., Wittig, R. M., Langergraber, K. E., Ziegler, T. E., Zuberbühler, K., & Deschner, T. (2013). Urinary oxytocin and social bonding in related and unrelated wild chimpanzees (Vol. 280, p. 20122765). The Royal Society.

Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, *22*(2), 142–146. https://doi.org/10.1016/j.cub.2011.11.053

Crockford, C., Wittig, R. M., & Zuberbühler, K. (2015). An intentional vocalization draws others' attention: A playback experiment with wild chimpanzees. *Animal Cognition*, *18*(3), 581–591. https://doi.org/10.1007/s10071-014-0827-z

Croft, D. B., & Snaith, F. (1990). Boxing in red kangaroos, *Macropos Rufus*: Aggression or play? *International Journal of Comparative Psychology*, *4*(3), 221–236.

Cronin, K. A., De Groot, E., & Stevens, J. M. G. (2015). Bonobos show limited social tolerance in a group setting: A comparison with chimpanzees and a test of the relational model. *Folia Primatologica*, *86*(3), 164–177. https://doi.org/10.1159/000373886

de Waal, F. B. M. (1988). The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behavior*, *106*(3), 183–251. https://doi.org/https://doi.org/10.1163/156853988X00269

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: a comparative perspective. *Biological Reviews*, *89*(2), 284–301. https://doi.org/10.1111/brv.12053

Demuru, E., Ferrari, P. F., & Palagi, E. (2015). Emotionality and intentionality in bonobo playful communication. *Animal Cognition*, *18*(1), 333–344. https://doi.org/10.1007/s10071-014-0804-6

Dezecache G, Zuberbühler K, Davila-Ross M, Dahl CD (2017) Skin temperature changes in wild chimpanzees upon hearing vocalizations of conspecifics. R Soc Open Sci 4:160816 . doi: 10.1098/rsos.160816

Dobson, S. (2012). Coevolution of facial expression and social tolerance in macaques. *American Journal of Primatology*, *74*(3), 229–235. https://doi.org/10.1002/ajp.21991

Dunbar, R. I. M. (1991). Functional significance of social grooming in primates. *Folia Primatologica*, *57*(3), 121–131. https://doi.org/10.1159/000156574

Duthler, K. W. (2006). The politeness of requests made via email and voicemail: Support for the hyperpersonal model. *Journal of Computer-Mediated Communication*, *11*(2), 500–521. https://doi.org/10.1111/j.1083-6101.2006.00024.x

Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, *60*(2), 440–453. https://doi.org/10.2307/1130988

Eckerman, C. O., & Peterman, K. (2001). Peers and infant social/communicative development. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 326–350). Oxford, UK: Blackwell Publishing Ltd.

Eckerman, C. O., & Whitehead, H. (1999). How toddler peers generate coordinated action: A cross-cultural exploration. *Early Education and Development*, *10*(3), 241–266. https://doi.org/10.1207/s15566935eed1003_2

Enfield, N. J. (2006). Social consequences of common ground. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: culture, cognition, and interaction.* (pp. 399–430). London: Berg.

Enfield, N. J. (2008). Common ground as a resource for social affiliation. In I. Kecskes & J. Mey (Eds.), *Intention, common ground and the egocentric speaker hearer* (pp. 223–254). Berlin: Mouton de Gruyter.

Engelmann, J. M., & Herrmann, E. (2016). Chimpanzees trust their friends. *Current Biology*, *26*(2). https://doi.org/10.1016/j.cub.2015.11.037

Fedurek, P., & Dunbar, R. I. M. (2009). What does mutual grooming tell us about why chimpanzees groom? *Ethology*, *115*(6), 566–575. https://doi.org/10.1111/j.1439-0310.2009.01637.x

Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour*, *86*(1), 189–196. https://doi.org/10.1016/J.ANBEHAV.2013.05.010

Fedurek, P., Slocombe, K. E., Hartel, J. A., & Zuberbühler, K. (2015). Chimpanzee lip-smacking facilitates cooperative behaviour. *Scientific Reports*, *5*(1), 13460. https://doi.org/10.1038/srep13460

Fedurek, P., Zuberbühler, K., & Semple, S. (2017). Trade-offs in the production of animal vocal sequences: insights from the structure of wild chimpanzee pant hoots. *Frontiers in Zoology*, *14*(1), 50. https://doi.org/10.1186/s12983-017-0235-8

Ferraro, J. V., Plummer, T. W., Pobiner, B. L., Oliver, J. S., Bishop, L. C., Braun, D. R., … Potts, R. (2013). Earliest archaeological evidence of persistent hominin carnivory. *PLoS ONE*, *8*(4), e62174. https://doi.org/10.1371/journal.pone.0062174

Flack, J. C., Jeannotte, L. A., & de Waal, F. (2004). Play signaling and the perception of social rules by juvenile chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *118*(2), 149–159. https://doi.org/http://dx.doi.org/10.1037/0735-7036.118.2.149

Fletcher, G. E., Warneken, F., & Tomasello, M. (2012). Differences in cognitive processes underlying the collaborative activities of children and chimpanzees. *Cognitive Development*, *27*(2), 136–153. https://doi.org/10.1016/J.COGDEV.2012.02.003

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1597), 1785–1801. https://doi.org/10.1098/rstb.2011.0213

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644--R645. https://doi.org/10.1016/j.cub.2005.08.041

Fröhlich, M. (2017). Taking turns across channels: Conversation-analytic tools in animal communication. *Neuroscience & Biobehavioral Reviews*, *80*, 201–209. https://doi.org/10.1016/j.neubiorev.2017.05.005

Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, *6*(25887), 25887. https://doi.org/10.1038/srep25887

Fröhlich, M., Müller, G., Zeiträg, C., Wittig, R. M., & Pika, S. (2017). Gestural development of chimpanzees in the wild: the impact of interactional experience. *Animal Behaviour*, *134*, 271–282. https://doi.org/10.1016/j.anbehav.2016.12.018

Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & van Schaik, C. P. (2019). Multimodal communication and language origins: integrating gestures and vocalizations. *Biological Reviews*, *94*(5), 1809–1829. https://doi.org/10.1111/brv.12535

Fröhlich, M., Wittig, R. M., & Pika, S. (2016a). Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science*, *3*(8). https://doi.org/10.1098/rsos.160278

Fröhlich, M., Wittig, R. M., & Pika, S. (2016b). Should I stay or should I go? Initiation of joint travel in mother–infant dyads of two chimpanzee communities in the wild. *Animal Cognition*, *19*(3), 483–500. https://doi.org/10.1007/s10071-015-0948-z

Fröhlich, M., Wittig, R. M., & Pika, S. (2019). The ontogeny of intentional communication in chimpanzees in the wild. *Developmental Science*, *22*(1), e12716. https://doi.org/10.1111/desc.12716

Fruth, B., Hickey, J. R., André, C., Furuichi, T., Hart, J., Hart, T., … Williamson, E. A. (2016). Pan paniscus. *The IUCN Red List of Threatened Species 2016*, e.T15932A102331567. https://doi.org/http://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15932A17964305.en.

Fruth, B., & Hohmann, G. (2018). Food Sharing across Borders. *Human Nature*, *29*(2), 91–103. https://doi.org/10.1007/s12110-018-9311-9

Furuichi, T. (1989). Social interactions and the life history of female *Pan paniscus* in Wamba, Zaire. *International Journal of Primatology*, *10*(3), 173–197. https://doi.org/10.1007/BF02735199

Furuichi, T. (2009). Factors underlying party size differences between chimpanzees and bonobos: a review and hypotheses for future study. *Primates*, *50*(3), 197–209. https://doi.org/10.1007/s10329-009-0141-6

Furuichi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary*

Anthropology: Issues, News, and Reviews, *20*(4), 131–142. https://doi.org/10.1002/evan.20308
Furuichi, T., Idani, G., Ihobe, H., Kuroda, S., Kitamura, K., Mori, A., … Kano, T. (1998). Population Dynamics of Wild Bonobos *Pan paniscus* at Wamba. *International Journal of Primatology*, *19*(6), 1029–1043. https://doi.org/10.1023/A:1020326304074
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
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
Genty, E. (2019). Vocal–gestural combinations in infant bonobos: new insights into signal functional specificity. *Animal Cognition*, *22*(4), 505–518. https://doi.org/10.1007/s10071-019-01267-0
Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Animal Cognition*, *12*(3), 527–546. https://doi.org/10.1007/s10071-009-0213-4
Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PloS One*, *9*(1), e84738. https://doi.org/https://doi.org/10.1371/journal.pone.0084738
Genty, E., Heesen, R., Rossano, F., Zuberbühler, K., Guery, J., & Bangerter, A. (submitted). How apes get into and out of joint actions: precursors of shared intentionality?
Genty, E., Neumann, C., & Zuberbühler, K. (2015a). Bonobos modify communication signals according to recipient familiarity. *Scientific Reports*, *5*(16442), 1–10.
Genty, E., Neumann, C., & Zuberbühler, K. (2015b). Complex patterns of signalling to convey different social goals of sex in bonobos, Pan paniscus. *Scientific Reports*, *5*(16135), 1–13.
Genty, E., & Zuberbühler, K. (2014). Spatial Reference in a Bonobo Gesture. *Current Biology*, *24*(14), 1601–1605. https://doi.org/10.1016/J.CUB.2014.05.065
Genty, E., & Zuberbühler, K. (2015). Iconic gesturing in bonobos. *Communicative & Integrative Biology*, *8*(1), e992742. https://doi.org/10.4161/19420889.2014.992742
Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and interaction* (pp. 229–255). Oxford: Berg Publishers.
Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies In Philosophy*, *15*(1), 1–14. https://doi.org/10.1111/j.1475-4975.1990.tb00202.x
Gilbert, M. (1999). Obligation and joint commitment. *Utilitas*, *11*(02), 143–163. https://doi.org/10.1017/S0953820800002399
Gilbert, M. (2014). *Joint commitment: How we make the social world*. New York: Oxford University Press.
Goffman, Erving. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry*, *18*(3), 213–231. https://doi.org/10.1080/00332747.1955.11023008
Goffman, Erving. (1963). *Behavior in public places*. New York: Free Press.
Goffman, Erving. (1967). *Interactional ritual: Essays on face-to-face behavior*. Garden City, NY: Anchor Books.
Goffman, Erving. (1981a). Footing. In E Goffman (Ed.), *Forms of talk* (pp. 124–159). Philadelphia: University of Pennsylvania Press.
Goffman, Erving. (1981b). *Forms of talk.* Oxford, England: Blackwell.
Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1657), 699–706. https://doi.org/10.1098/rspb.2008.1324
Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge MA: Belknap Press of Harvard University Press.
Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*(5), 1430–1443. https://doi.org/10.1037/a0016122
Gräfenhain, M., Carpenter, M., & Tomasello, M. (2013). Three-year-olds' understanding of the consequences of joint commitments. *PLOS ONE*, *8*(9), e73039.

https://doi.org/10.1371/journal.pone.0073039
Graham, K. E., Furuichi, T., & Byrne, R. W. (2017). The gestural repertoire of the wild bonobo (*Pan paniscus*): a mutually understood communication system. *Animal Cognition*, *20*(2), 171–177. https://doi.org/10.1007/s10071-016-1035-9
Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLOS Biology*, *16*(2), e2004825. https://doi.org/10.1371/journal.pbio.2004825
Greenberg, J. R., Hamann, K., Warneken, F., & Tomasello, M. (2010). Chimpanzee helping in collaborative and noncollaborative contexts. *Animal Behaviour*, *80*(5), 873–880. https://doi.org/10.1016/J.ANBEHAV.2010.08.008
Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech acts* (pp. 41–58). London: Academic Press.
Gruber, T. (2014). Wild-Born Orangutans (Pongo abelii) Engage in Triadic Interactions During Play. *International Journal of Primatology*, *35*(2), 411–424. https://doi.org/10.1007/s10764-013-9745-1
Gruber, T., & Clay, Z. (2016). A comparison between bonobos and chimpanzees: A review and update. *Evolutionary Anthropology: Issues, News, and Reviews*, *25*(5), 239–252. https://doi.org/10.1002/evan.21501
Halina, M., Rossano, F., & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal Cognition*, *16*(4), 653–666. https://doi.org/10.1007/s10071-013-0601-7
Hall, K., Oram, M. W., Campbell, M. W., Eppley, T. M., Byrne, R. W., & de Waal, F. B. M. (2017). Chimpanzee uses manipulative gaze cues to conceal and reveal information to foraging competitor. *American Journal of Primatology*, *79*(3), e22622--n/a. https://doi.org/10.1002/ajp.22622
Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development*, *83*(1), 137–145. https://doi.org/10.1111/j.1467-8624.2011.01695.x
Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785. https://doi.org/10.1006/ANBE.1999.1377
Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, *17*(7), 619–623. https://doi.org/10.1016/j.cub.2007.02.040
Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573–585. https://doi.org/10.1016/J.ANBEHAV.2011.12.007
Hashimoto, C., Tashiro, Y., Kimura, D., Enomoto, T., Ingmanson, E. J., Idani, G., & Furuichi, T. (1998). Habitat use and ranging of wild bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, *19*(6), 1045–1060. https://doi.org/10.1023/A:1020378320913
Heesen, R., Genty, E., Rossano, F., Zuberbühler, K., & Bangerter, A. (2017). Social play as joint action: A framework to study the evolution of shared intentionality as an interactional achievement. *Learning and Behavior*, *45*(4), 390–405. https://doi.org/10.3758/s13420-017-0287-9
Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571
Heritage, J., & Atkinson, J. M. (1984). *Structures of social action: Studies in conversation analysis*. (J. M. Atkinson & J. Heritage, Eds.). Cambridge: Cambridge University Press.
Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the cognitive skills of bonobos and chimpanzees. *PLoS ONE*, *5*(8), e12438. https://doi.org/10.1371/journal.pone.0012438
Herrmann, E., Hare, B., Cissewski, J., & Tomasello, M. (2011). A comparison of temperament in nonhuman apes and human infants. *Developmental Science*, *14*(6), 1393–1405. https://doi.org/10.1111/j.1467-7687.2011.01082.x
Herrmann, E., Hernández-Lloreda, M. V., Call, J., Hare, B., & Tomasello, M. (2010). The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science*, *21*(1), 102–110. https://doi.org/10.1177/0956797609356511

Hobaiter, C., & Byrne, R. W. (2011a). Serial gesturing by wild chimpanzees: its nature and function for communication. *Animal Cognition*, *14*(6), 827–838. https://doi.org/10.1007/s10071-011-0416-3

Hobaiter, C., & Byrne, R. W. (2011b). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. https://doi.org/10.1007/s10071-011-0409-2

Hobaiter, C., & Byrne, R. W. (2012). Gesture use in consortship: wild chimpanzees' use of gesture for an 'evolutionarily urgent'purpose. In S. Pika & K. Liebal (Eds.), *Development in Primate Gesture Research* (pp. 129–146). Amsterdam: John Benjamins Press.

Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology*, *24*(14), 1596–1600. https://doi.org/10.1016/j.cub.2014.05.066

Hobaiter, C., & Byrne, R. W. (2017). What is a gesture? A meaning-based approach to defining gestural repertoires. *Neuroscience & Biobehavioral Reviews*, *82*, 3–12. https://doi.org/10.1016/J.NEUBIOREV.2017.03.008

Hobaiter, C., Leavens, D. A., & Byrne, R. W. (2014). Deictic gesturing in wild chimpanzees (*Pan troglodytes*)? Some possible cases. *Journal of Comparative Psychology*, *128*(1), 82–87. https://doi.org/10.1037/a0033757

Hohmann, G., & Fruth, B. (2000). Use and function of genital contacts among female bonobos. *Animal Behaviour*, *60*(1), 107–120. https://doi.org/10.1006/ANBE.2000.1451

Holekamp, K. E., Sakai, S. T., & Lundrigan, B. L. (2007). The spotted hyena (*Crocuta crocuta*) as a model system for study of the evolution of intelligence. *Journal of Mammalogy*, *88*(3), 545–554. https://doi.org/10.1644/06-MAMM-S-361R1.1

Holle, H., & Rein, R. (2015). EasyDIAg: A tool for easy determination of interrater agreement. *Behavior Research Methods*, *47*(3), 837–847. https://doi.org/10.3758/s13428-014-0506-7

Holtgraves, T. M. (2002). *Language as social action*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hopkins, W. D., Russell, J., McIntyre, J., & Leavens, D. A. (2013). Are chimpanzees really so poor at understanding imperative pointing? Some new data and an alternative view of canine and ape social cognition. *PLoS ONE*, *8*(11), e79338. https://doi.org/10.1371/journal.pone.0079338

Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2008). Observational learning in chimpanzees and children studied through 'ghost' conditions. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 835–840. https://doi.org/10.1098/rspb.2007.1542

Hostetter, A. B., Russell, J. L., Freeman, H., & Hopkins, W. D. (2006). Now you see me, now you don't: evidence that chimpanzees understand the role of the eyes in attention. *Animal Cognition*, *10*(1), 55–62. https://doi.org/10.1007/s10071-006-0031-x

Humle, T., Maisels, F., Oates, J. F., Plumptre, A., & Williamson, E. A. (2016). *Pan troglodytes*. *The IUCN Red List of Threatened Species*, e.T15933A129038584. https://doi.org/http://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15933A17964454.en.

Itakura, S., Agnetta, B., Hare, B., & Tomasello, M. (1999). Chimpanzee use of human and conspecific social cues to locate hidden food. *Developmental Science*, *2*(4), 448–456. https://doi.org/10.1111/1467-7687.00089

Jaeggi, A. V, Stevens, J. M. G., & Van Schaik, C. P. (2010). Tolerant food sharing and reciprocity is precluded by despotism among bonobos but not chimpanzees. *American Journal of Physical Anthropology*, *143*(1), 41–51.

Kaburu, S. S. K., Inoue, S., & Newton-Fisher, N. E. (2013). Death of the alpha: Within-community lethal violence among chimpanzees of the Mahale Mountains National Park. *American Journal of Primatology*, *75*(8), 789–797. https://doi.org/10.1002/ajp.22135

Kachel, U., Svetlova, M., & Tomasello, M. (2017). Three-year-olds' reactions to a partner's failure to perform her role in a joint commitment. *Child Development*, *89*(5), 1–13. https://doi.org/10.1111/cdev.12816

Kachel, U., & Tomasello, M. (2019). 3- and 5-year-old children's adherence to explicit and implicit joint commitments. *Developmental Psychology*, *55*(1), 80–88. https://doi.org/10.1037/dev0000632

Kano, T. (1992). *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford: Standford University

Press.
Kendon, A. (1976). The F-formation system: The spatial organization of social encounters. *Man-Environment Systems*, *6*, 291–296.
Kendon, A. (1990). *Conducting interaction : patterns of behavior in focused encounters*. Cambridge University Press.
Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive Ethology: A new approach for studying human cognition. *British Journal of Psychology*, *99*(3), 317–340. https://doi.org/10.1348/000712607X251243
Knapp, M. L., Hart, R. P., Friedrich, G. W., & Shulman, G. M. (1973). The rhetoric of goodbye: Verbal and nonverbal correlates of human leave-taking. *Speech Monographs*, *40*(3), 182–198. https://doi.org/10.1080/03637757309375796
Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In *Psychology of Learning and Motivation* (Vol. 54, p. 59). Burlington: Academic Press.
Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1499), 2021. https://doi.org/10.1098/rstb.2008.0006
Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: social interaction modulates action simulation. *Biology Letters*, *6*(6), 758–761. https://doi.org/10.1098/rsbl.2010.0478
Koyama, N. F., Caws, C., & Aureli, F. (2006). Interchange of Grooming and Agonistic Support in Chimpanzees. *International Journal of Primatology*, *27*(5), 1293–1309. https://doi.org/10.1007/s10764-006-9074-8
Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114. https://doi.org/10.1126/science.aaf8110
Langergraber, K. E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., Fawcett, K., … Vigilant, L. (2012). Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39), 15716–15721. https://doi.org/10.1073/pnas.1211740109
Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2017). The mismeasure of ape social cognition. *Animal Cognition*. https://doi.org/10.1007/s10071-017-1119-1
Leavens, D. A., Hopkins, W. D., & Bard, K. A. (1996). Indexical and referential pointing in chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *110*(4), 346–353. https://doi.org/10.1037/0735-7036.110.4.346
Leavens, D. A., Hostetter, A. B., Wesley, M. J., & Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, Pan troglodytes. *Animal Behaviour*, *67*(3), 467–476. https://doi.org/10.1016/j.anbehav.2003.04.007
Leavens, D. A., Reamer, L. A., Mareno, M. C., Russell, J. L., Wilson, D., Schapiro, S. J., & Hopkins, W. D. (2015). Distal communication by chimpanzees (Pan troglodytes): Evidence for common ground? *Child Development*, *86*(5), 1623–1638. https://doi.org/10.1111/cdev.12404
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, England: 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 B: Biological Sciences*, *369*(1651). https://doi.org/10.1098/rstb.2013.0302
Liebal, K., Call, J., Tomasello, M., & Pika, S. (2004). To move or not to move: How apes adjust to the attentional state of others. *Interaction Studies*, *5*(2), 199–219. https://doi.org/10.1075/is.5.2.03lie
Liebal, K., Waller, B. M., Slocombe, K. E., & Burrows, A. M. (2013). *Primate communication: A multimodal approach*. Cambridge University Press.

Limberg, H. (2009). Impoliteness and threat responses. *Journal of Pragmatics*, *41*(7), 1376–1394. https://doi.org/10.1016/J.PRAGMA.2009.02.003

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*(3), 297–307. https://doi.org/10.1111/j.1467-7687.2004.00349.x

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*(2), 173–187. https://doi.org/10.1207/s15327647jcd0702_2

Luef, E. M., & Liebal, K. (2013). The hand-on gesture in gorillas (Gorilla gorilla). *Interaction Studies*, *14*(1), 44–61. https://doi.org/10.1075/is.14.1.04lue

MacLean, E., & Hare, B. (2013). Spontaneous triadic engagement in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *127*(3), 245–255. https://doi.org/10.1037/a0030935

MacNulty, D. R., Tallian, A., Stahler, D. R., & Smith, D. W. (2014). Influence of group size on the success of wolves hunting bison. *PLoS ONE*, *9*(11), e112884. https://doi.org/10.1371/journal.pone.0112884

Maestripieri, D. (2005). Gestural communication in three species of macaques (*Macaca mulatta* , *M. nemestrina* , *M. arctoides*): Use of signals in relation to dominance and social context. *Gesture*, *5*(1–2), 57–73. https://doi.org/10.1075/gest.5.1.06mae

Mayor, E., & Bangerter, A. (2015). Managing perturbations during handover meetings: a joint activity framework. *Nursing Open*, *2*(3), 130–140. https://doi.org/10.1002/nop2.29

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

Melis, A. P., Hare, B., & Tomasello, M. (2006b). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour*, *72*(2), 275–286. https://doi.org/10.1016/J.ANBEHAV.2005.09.018

Melis, A. P., Hare, B., & Tomasello, M. (2008). Do chimpanzees reciprocate received favours? *Animal Behaviour*, *76*(3), 951–962. https://doi.org/10.1016/J.ANBEHAV.2008.05.014

Melis, A. P., & Tomasello, M. (2013). Chimpanzees' (*Pan troglodytes*) strategic helping in a collaborative task. *Biology Letters*, *9*(2). https://doi.org/10.1098/rsbl.2013.0009

Melis, A. P., & Tomasello, M. (2019). Chimpanzees (*Pan troglodytes*) coordinate by communicating in a collaborative problem-solving task. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1901), 20190408. https://doi.org/10.1098/rspb.2019.0408

Melis, A. P., Warneken, F., Jensen, K., Schneider, A.-C., Call, J., & Tomasello, M. (2011). Chimpanzees help conspecifics obtain food and non-food items. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1710), 1405–1413. https://doi.org/10.1098/rspb.2010.1735

Miller, K. (2002). Individual and joint commitments. In Meggle G. (Ed.), *Social facts & collective intentionality* (pp. 255–272). Frankfurt, Germany: German Library of Sciences.

Miller, K. (2006). Social obligation as reason for action. *Cognitive Systems Research*, *7*(2–3), 273–285. https://doi.org/10.1016/J.COGSYS.2005.11.005

Mitani, J. C., & Watts, D. (2005). Correlates of territorial boundary patrol behaviour in wild chimpanzees. *Animal Behaviour*, *70*(5), 1079–1086. https://doi.org/10.1016/J.ANBEHAV.2005.02.012

Modaff, D. P. (2003). Body movement in the transition from opening to task in doctor–patient interviews. In P. Glenn, C. D. LeBaron, & J. Mandelbaum (Eds.), *Studies in language and social Interaction: In honor of Robert Hopper* (pp. 411–422).

Mondada, L. (2006). Participants' online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. *Discourse Studies*, *8*(1), 117–129. https://doi.org/10.1177/1461445606059561

Mondada, Lorenza. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, *41*(10), 1977–1997. https://doi.org/10.1016/J.PRAGMA.2008.09.019

Morand, D. A. (1996). Dominance, deference, and egalitarianism in organizational interaction: A sociolinguistic analysis of power and politeness. *Organization Science*, *7*(5), 544–556. https://doi.org/10.1287/orsc.7.5.544

Morand, D. A. (2000). Language and power: an empirical analysis of linguistic strategies used in superior± subordinate communication. *Journal of Organizational Behavior*, *21*(3), 235–248.

Morgan, B., Adeleke, A., Bassey, T., Bergl, T., Bergl, R., Dunn, A., … Williamson, L. (2011). Regional Action Plan for the conservation of the Nigeria-Cameroon chimpanzee (Pan troglodytes ellioti). In *IUCN/SSC Primate Specialist Group and Zoological Society of San Diego, CA, USA*.

Morgan L (1981) Animal life and intelligence. Ginn, Boston, MA

Moscovice, L. R., Surbeck, M., Fruth, B., Hohmann, G., Jaeggi, A. V., & Deschner, T. (2019). The cooperative sex: Sexual interactions among female bonobos are linked to increases in oxytocin, proximity and coalitions. *Hormones and Behavior*, *116*, 104581. https://doi.org/10.1016/J.YHBEH.2019.104581

Neda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabasi, A.-L. (2000). Self-organizing processes: The sound of many hands clapping. *Nature*, *403*(6772), 849–850. https://doi.org/10.1038/35002660

Neumann, C. (2016). Socialindices: Social and association indices. Unpublished R package (v.0.46-7).

Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., … Engelhardt, A. (2011). Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour*, *82*(4), 911–921. https://doi.org/10.1016/j.anbehav.2011.07.016

Neumann, C., & Kulik, L. (2014). EloRating: Animal dominance hierarchies by Elo rating. *R Package Version*, *43*, 1–26. Retrieved from http://cran.ms.unimelb.edu.au/web/packages/EloRating/vignettes/tutorial.pdf

Neumann, C., McDonald, D. B., & Daizaburo, S. (2018). Dominance ranks, dominance ratings and linear hierarchies: a critique. *Animal Behaviour*, *144*, e1–e16.

Newton-Fisher, N. E., & Lee, P. C. (2011). Grooming reciprocity in wild male chimpanzees. *Animal Behaviour*, *81*(2), 439–446. https://doi.org/10.1016/J.ANBEHAV.2010.11.015

O'Leary, M. J., & Gallois, C. (1985). The last ten turns: Behavior and sequencing in friends' and strangers' conversational findings. *Journal of Nonverbal Behavior*, *9*(1), 8–27. https://doi.org/10.1007/BF00987556

Palagi, E. (2006). Social play in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): Implicationsfor natural social systems and interindividual relationships. *American Journal of Physical Anthropology*, *129*(3), 418–426. https://doi.org/10.1002/ajpa.20289

Palagi, E. (2008). Sharing the motivation to play: the use of signals in adult bonobos. *Animal Behaviour*, *75*(3), 887–896. https://doi.org/10.1016/J.ANBEHAV.2007.07.016

Palagi, E., Antonacci, D., & Cordoni, G. (2007). Fine-tuning of social play in juvenile lowland gorillas (*Gorilla gorilla gorilla*). *Developmental Psychobiology*, *49*(4), 433–445. https://doi.org/10.1002/dev.20219

Palagi, E., Burghardt, G. M., Smuts, B., Cordoni, G., Dall'Olio, S., Fouts, H. N., … Pellis, S. M. (2015). Rough-and-tumble play as a window on animal communication. *Biological Reviews*, *91*(2), 311–327. https://doi.org/10.1111/brv.12172

Palagi, E., Cordoni, G., Demuru, E., & Bekoff, M. (2016). Fair play and its connection with social tolerance, reciprocity and the ethology of peace. *Behaviour*, *153*(9–11), 1195–1216.

Palagi, E., & Paoli, T. (2007). Play in adult bonobos (*Pan paniscus*): modality and potential meaning. *American Journal of Physical Anthropology*, *134*(2), 219–225.

Parr, L. A., Hopkins, W. D., & DeWaal, F. B. M. (1998). The perception of facial expressions by chimpanzees, Pan troglodytes. *Evolution of Communication*, *2*(1), 1–23. https://doi.org/10.1075/eoc.2.1.02par

Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion (Washington, D.C.)*, *7*(1), 172–181. https://doi.org/10.1037/1528-3542.7.1.172

Pika, S., & Fröhlich, M. (2019). Gestural acquisition in great apes: the Social Negotiation Hypothesis. *Animal Cognition*, *22*(4), 551–565. https://doi.org/10.1007/s10071-017-1159-6

Pika, S., Liebal, K., Call, J., & Tomasello, M. (2005). Gestural communication of apes. *Gesture*, *5*(1–2), 41–56. https://doi.org/10.1075/gest.5.1.05pik

Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, learning, and use. *American Journal of Primatology*, *60*(3), 95–111. https://doi.org/10.1002/ajp.10097

Pika, S., & Mitani, J. C. (2009). *The directed scratch: Evidence for a referential gesture in chimpanzees?* (R. Botha & C. Knight, Eds.), *The prehistory of language*. Oxford, UK: Oxford University Press.

Pika, S., & Zuberbühler, K. (2008). Social games between bonobos and humans: evidence for shared intentionality? *American Journal of Primatology*, *70*(3), 207–210. https://doi.org/10.1002/ajp.20469

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed effects models in S and S-Plus*. New York: Springer.

Pitman, R. L., & Durban, J. W. (2012). Cooperative hunting behavior, prey selectivity and prey handling by pack ice killer whales (*Orcinus orca*), type B, in Antarctic Peninsula waters. *Marine Mammal Science*, *28*(1), 16–36. https://doi.org/10.1111/j.1748-7692.2010.00453.x

Port, M., Clough, D., & Kappeler, P. M. (2009). Market effects offset the reciprocation of grooming in free-ranging redfronted lemurs, Eulemur fulvus rufus. *Animal Behaviour*, *77*(1), 29–36. https://doi.org/10.1016/j.anbehav.2008.08.032

Pozis-Francois, O., Zahavi, A., & Zahavi, A. (2004). Social play in Arabian babblers. *Behaviour*, *141*(4), 425–450. https://doi.org/10.1163/156853904323066720

Prüfer, K., Munch, K., Hellmann, I., Akagi, K., Miller, J. R., Walenz, B., … Pääbo, S. (2012). The bonobo genome compared with the chimpanzee and human genomes. *Nature*, *486*(7404), 527–531. https://doi.org/10.1038/nature11128

Rakoczy, H. (2007). Play, games, and the development of collective intentionality. *New Directions for Child and Adolescent Development*, *2007*(115), 53–67. https://doi.org/10.1002/cd.182

Rakoczy, H., & Schmidt, M. F. H. (2013). The early ontogeny of social norms. *Child Development Perspectives*, *7*(1), 17–21. https://doi.org/10.1111/cdep.12010

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PLOS ONE*, *8*(8), e71182. https://doi.org/10.1371/journal.pone.0071182

Roberts, A. I., Vick, S.-J., Roberts, S. G. B., & Menzel, C. R. (2014). Chimpanzees modify intentional gestures to coordinate a search for hidden food. *Nature Communications*, *5*(3088), 1–8. https://doi.org/DOI: 10.1038/ncomms4088

Rosas, A., & Bermúdez, J. P. (2018). Viewing others as equals: the non-cognitive roots of shared intentionality. *Review of Philosophy and Psychology*, *9*(3), 485–502. https://doi.org/10.1007/s13164-018-0394-3

Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2007). The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology*, *17*(19), 1663–1668. https://doi.org/10.1016/J.CUB.2007.08.033

Ross, H. S., & Lollis, S. P. (1987). Communication within infant social games. *Developmental Psychology*, *23*(2), 241–248. https://doi.org/10.1037/0012-1649.23.2.241

Rossano, F. (2013). Sequence organization and timing of bonobo mother-infant interactions. *Interaction Studies*, *14*(2), 160–189. https://doi.org/10.1075/is.14.2.02ros

Rossano, F., & Liebal, K. (2014). "Requests" and "offers" in orang-utans and human infants. In *Requesting in Social Interaction* (Vol. 26, pp. 333–362). Amsterdam/Philadelphia: John Benjamins.

Sacks, H., Schegloff, E. A., & 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

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE*, *8*(10), e76674.

https://doi.org/10.1371/journal.pone.0076674
Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x
Schino, G., di Sorrentino, E. P., & Tiddi, B. (2007). Grooming and coalitions in Japanese macaques (Macaca fuscata): Partner choice and the time frame reciprocation. *Journal of Comparative Psychology*, *121*(2), 181–188. https://doi.org/10.1037/0735-7036.121.2.181
Searle, J. R. (1979). What Is an Intentional State? *Mind*, *88*, 74–92.
Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401–415). Cambridge MA: MIT Press.
Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009
Seyfarth, R. M., & Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 249–273). Cambridge, UK: Cambridge University Press.
Sidnell, J., & Stivers, T. (2012). *The handbook of conversation analysis* (Vol. 121). John Wiley & Sons.
Silk, J., Alberts, S. C., & Altmann, J. (2006). Social relationships among adult female baboons (Papio cynocephalus) II. Variation in the quality and stability of social bonds. *Behavioral Ecology and Sociobiology*, *61*(2), 197–204. https://doi.org/10.1007/s00265-006-0250-9
Silk, J., Cheney, D., & Seyfarth, R. (2013). A practical guide to the study of social relationships. *Evolutionary Anthropology: Issues, News, and Reviews*, *22*(5), 213–225. https://doi.org/10.1002/evan.21367
Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17228–17233. https://doi.org/10.1073/pnas.0706741104
Smith, P. K., & Boulton, M. (1990). Rough-and-tumble play, aggression and dominance: Perception and behaviour in children's encounters. *Human Development*, *33*(4–5), 271–282. https://doi.org/DOI:10.1159/000276524
Smith, T. G., Siniff, D. B., Reichle, R., & Stone, S. (1981). Coordinated behavior of killer whales, *Orcinus orca* , hunting a crabeater seal, *Lobodon carcinophagus*. *Canadian Journal of Zoology*, *59*(6), 1185–1189. https://doi.org/10.1139/z81-167
Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
Spinka, M., Newberry, R. C., & Bekoff, M. (2001). Mammalian play: Training for the unexpected. *The Quarterly Review of Biology*, *76*(2), 141–168. https://doi.org/10.1086/393866
Stadler, S. A. (2006). *Multimodal im/politeness: The verbal, prosodic and non-verbal realization of disagreement in German and New Zealand English. Dissertation*. University of Auckland.
Stander, P. E. (1992). Cooperative hunting in lions: the role of the individual. *Behavioral Ecology and Sociobiology*, *29*(6), 445–454. https://doi.org/10.1007/BF00170175
Stevens, J. M. G., Vervaecke, H., de Vries, H., & van Elsacker, L. (2007). Sex differences in the steepness of dominance hierarchies in captive bonobo groups. *International Journal of Primatology*, *28*(6), 1417–1430. https://doi.org/10.1007/s10764-007-9186-9
Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The Morality of Knowledge in Conversation* (pp. 3–24). Cambridge: Cambridge University Press.
Suchak, M., Eppley, T. M., Campbell, M. W., Feldman, R. A., Quarles, L. F., & de Waal, F. B. M. (2016). How chimpanzees cooperate in a competitive world. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(36), 10215–10220. https://doi.org/10.1073/pnas.1611826113
Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1317–1324. https://doi.org/10.1098/rstb.2008.0301

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*(3), 299–313. https://doi.org/10.1017/S0140525X07001975

Sugiyama, Y. (1988). Grooming interactions among adult chimpanzees at bossou, Guinea, with special reference to social structure. *International Journal of Primatology*, *9*(5). Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF02736216.pdf

Surbeck, M., Deschner, T., Schubert, G., Weltring, A., & Hohmann, G. (2012). Mate competition, testosterone and intersexual relationships in bonobos, *Pan paniscus*. *Animal Behaviour*, *83*(3), 659–669. https://doi.org/10.1016/J.ANBEHAV.2011.12.010

Surbeck, M., Girard-Buttoz, C., Boesch, C., Crockford, C., Fruth, B., Hohmann, G., … Mundry, R. (2017). Sex-specific association patterns in bonobos and chimpanzees reflect species differences in cooperation. *Royal Society Open Science*, *4*(5). https://doi.org/10.1098/rsos.161081

Surbeck, M., & Hohmann, G. (2015). Social preferences influence the short-term exchange of social grooming among male bonobos. *Animal Cognition*, *18*(2), 573–579. https://doi.org/10.1007/s10071-014-0826-0

Tan, J., Ariely, D., & Hare, B. (2017). Bonobos respond prosocially toward members of other groups. *Scientific Reports*, *7*(1), 14733. https://doi.org/10.1038/s41598-017-15320-w

Tan, J., & Hare, B. (2013). Bonobos share with strangers. *PLoS One*, *8*(1), e51922.

Tanner, J. E. (2004). Gestural phrases and gestural exchanges by a pair of zoo-living lowland gorillas. *Gesture*, *4*(1), 1–24. https://doi.org/10.1075/gest.4.1.02tan

Team, R. C. (2013). R: A language and environment for statistical computing. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.470.5851&rep=rep1&type=pdf

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

Theall, L. A., & Povinelli, D. J. (1999). Do chimpanzees tailor their gestural signals to fit the attentional states of others? *Animal Cognition*, *2*(4), 207–214. https://doi.org/10.1007/s100710050041

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

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M. (2014). *A natural history of human thinking*. Cambridge, Massachusetts: Harvard University Press.

Tomasello, M. (2016). *A natural history of human morality*. Cambridge, Massachusetts: Harvard University Press.

Tomasello, M., & Call, J. (2006). Do chimpanzees know what others see—or only what they are looking at? In *Rational Animals?* (pp. 371–384). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528272.003.0017

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*(4), 1063–1069. https://doi.org/10.1006/anbe.1997.0636

Tomasello, M., Call, J., Warren, J., Frost, G. T., Carpenter, M., & Nagell, K. (1997). The ontogeny of chimpanzee gestural signals: A comparison across groups and generations. *Evolution of Communication*, *1*(2), 223–259. https://doi.org/10.1075/eoc.1.2.04tom

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., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(05), 675–735.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(03), 495. https://doi.org/10.1017/S0140525X0003123X

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation the interdependence hypothesis. *Current Anthropology*, *53*(6). https://doi.org/10.1086/668207

Tomasello, M., & Moll, H. (2010). The gap is social: human shared intentionality and culture. In P. M. Kappeler & J. B. Silk (Eds.), *Mind the gap. Tracing the origins of human universals* (pp. 331–349). Berlin: Springer.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language*, *18*(2), 121–147. https://doi.org/10.1111/1468-0017.00217

Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., … Manser, M. B. (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews*, *92*(3), 1427–1433. https://doi.org/10.1111/brv.12289

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, *58*(5), 583–603. https://doi.org/10.1016/S1071-5819(03)00023-5

Trimboli, C., & Walker, M. B. (1984). Switching pauses in cooperative and competitive conversations. *Journal of Experimental Social Psychology*, *20*(4), 297–311. https://doi.org/10.1016/0022-1031(84)90027-1

Tuomela, R. (2005). We-Intentions Revisited. *Philosophical Studies*, *125*(3), 327–369. https://doi.org/10.1007/s11098-005-7781-1

Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. New York: Oxford University Press.

Vail, A. L., Manica, A., & Bshary, R. (2013). Referential gestures in fish collaborative hunting. *Nature Communications*, *4*(1), 1765. https://doi.org/10.1038/ncomms2781

van Schaik, C. P., & Burkart, J. M. (2018). The moral capacity as a biological adaptation: A commentary on Tomasello. *Philosophical Psychology*, *31*(5), 703–721. https://doi.org/10.1080/09515089.2018.1486608

Vehtari, A., Gelman, A., & Gabry, J. (2015). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., … Nijssen, S. (2016). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*, *7*, 2039. https://doi.org/https://doi.org/10.3389/fpsyg.2016.02039

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

Vlamings, P. H. J. M., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3–5-year-old children. *Animal Cognition*, *13*(2), 273–285. https://doi.org/10.1007/s10071-009-0265-5

Völter, C. J., Rossano, F., & Call, J. (2015). From exploitation to cooperation: social tool use in orang-utan mother–offspring dyads. *Animal Behaviour*, *100*, 126–134. https://doi.org/10.1016/J.ANBEHAV.2014.11.025

Waller, B. M., Caeiro, C. C., & Davila-Ross, M. (2015). Orangutans modify facial displays depending on recipient attention. *PeerJ*, *3*, e827. https://doi.org/10.7717/peerj.827

Waller, B. M., & Cherry, L. (2012). Facilitating play through communication: Significance of teeth exposure in the gorilla play face. *American Journal of Primatology*, *74*(2), 157–164. https://doi.org/10.1002/ajp.21018

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640–663. https://doi.org/10.1111/j.1467-8624.2006.00895.x

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental*

Science, *15*(1), 54–61. https://doi.org/10.1111/j.1467-7687.2011.01107.x
Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLOS Biology*, *5*(7), e184. https://doi.org/10.1371/journal.pbio.0050184
Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301. https://doi.org/10.1126/science.1121448
Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology*, *44*(6), 1785–1788. https://doi.org/10.1037/a0013860
Watts, D. (2016). Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: variation, social learning, and possible functions. *Primates*, *57*(1), 61–72. https://doi.org/10.1007/s10329-015-0497-8
Watts, D., & Mitani, J. C. (2001). Boundary patrols and intergroup encounters in wild champanzees. *Behaviour*, *138*(3), 299–327.
Watts, D., Muller, M., Amsler, S. J., Mbabazi, G., & Mitani, J. C. (2006). Lethal intergroup aggression by chimpanzees in Kibale National Park, Uganda. *American Journal of Primatology*, *68*(2), 161–180. https://doi.org/10.1002/ajp.20214
Wellman, H. M., Cross, D., & Watson, J. (2001). 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
White, F. J., & Wrangham, R. W. (1988). Feeding competition and patch site in the chimpanzee species *Pan paniscus* and *Pan troglodytes*. *Behaviour*, *105*(1–2), 148–164. https://doi.org/https://doi.org/10.1163/156853988X00494
Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *110*(1), 3–14. https://doi.org/10.1037/0735-7036.110.1.3
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.
Wildman, D. E., Uddin, M., Liu, G., Grossman, L. I., & Goodman, M. (2003). Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus Homo. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(12), 7181–7188. https://doi.org/10.1073/pnas.1232172100
Williams, J. M., Liu, H. Y., & Pusey, A. E. (2002). Costs and benefits of grouping in female chimpanzees at Gombe. In C. Boesch, G. Hohmann, & L. F. Marchant (Eds.), *Behavioral diversity in Pan* (pp. 192–203). Cambridge: Cambridge University Press.
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
Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN : a professional framework for multimodality research, 1556–1559. Retrieved from http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:60436
Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, *20*(3), 226–230. https://doi.org/10.1016/J.CUB.2009.11.070
Yamamoto, S., Humle, T., & Tanaka, M. (2009). Chimpanzees help each other upon request. *PLoS ONE*, *4*(10), e7416. https://doi.org/10.1371/journal.pone.0007416
Yamamoto, S., Humle, T., & Tanaka, M. (2012). Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3588–3592. https://doi.org/10.1073/pnas.1108517109
Zlatev, J., Racine, T. P., Sinha, C., & Itkonen, E. (Eds.). (2008). *The shared mind: Perspectives on intersubjectivity*. Amsterdam: John Benjamins Publishing Company.

List of figures

List of tables

Appendices

The first Appendix, Appendix S1, contains the theoretical paper that was published to introduce the joint action framework, as part of my thesis. The remaining appendices (2-5) contain the electronic supplementary materials (ESMs) of respective articles.

Appendix 1

Manuscript is published and reprinted with adaptions by permission from Springer Nature and Copyright Clearance Center: Springer Nature, Learning & Behavior, Heesen, R., Genty, E., Rossano, F., Zuberbühler, K., & Bangerter, A. (2017). Social play as joint action: A framework to study the evolution of shared intentionality as an interactional achievement. Learning & Behavior, 45, 390–405. https://doi.org/10.3758/s13420-017-0287-9. License number: 43746806407.

Social Play as Joint Action:

A Framework to Study the Evolution of Shared Intentionality as an Interactional Achievement

Raphaela Heesen[1], Emilie Genty[1], Federico Rossano[2], Klaus Zuberbühler[1,3],

Adrian Bangerter[1]

[1]University of Neuchâtel, [2]University of California San Diego, [3] University of St Andrews

Author Note

Raphaela Heesen, Institute of Work and Organizational Psychology, University of Neuchâtel, Switzerland; Emilie Genty, Institute of Work and Organizational Psychology, University of Neuchâtel, Switzerland; Federico Rossano, Department of Cognitive Science, University of California San Diego, CA, USA; Klaus Zuberbühler, Department of Comparative Cognition, University of Neuchâtel, Switzerland, and School of Psychology and Neuroscience, University of St Andrews, Scotland (UK); Adrian Bangerter, Institute of Work and Organizational Psychology, University of Neuchâtel, Switzerland.

The present research was supported by the Swiss National Science Foundation (Grant No. CR31I3_166331 awarded to AB and KZ). The authors declare no conflicts of interest.

Correspondence concerning this article should be addressed to Raphaela Heesen, Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, Neuchâtel, CH-2000, phone: +41 32 718 21 13, email: raphaela.heesen@unine.ch

Abstract

Social play has a complex, cooperative nature that requires substantial coordination. This has led researchers to use social games to study cognitive abilities like shared intentionality, the skill and motivation to share goals and intentions with others during joint action. We expand this proposal by considering play as a joint action and examining how shared intentionality is achieved during human joint action. We describe how humans get into, conduct and get out of joint actions together in an orderly way, thereby constructing the state of "togetherness" characteristic of shared intentionality. These processes play out as three main phases, the opening (where participants are ratified and joint commitments are established), the main body (where progress, ongoing commitments and possible role reversals are coordinated), and the closing (where the intention to terminate the action is coordinated and where participants take leave of each other). We use this process in humans as a framework for examining how various animal species get into, maintain, and get out of play bouts. This comparative approach constitutes an alternative measure of those species' possession of shared intentionality. Using this framework, we review the play literature on human children and different social species of mammals and birds in search of behavioral markers of shared intentionality in the coordination of play bouts. We discuss how our approach could shed light on the evolution of the special human motivation to cooperate and share psychological states with others.

Keywords: social play, joint action, shared intentionality, communication, cooperation

Social Play as Joint Action: A Framework to Study the Evolution of Shared Intentionality as an Interactional Achievement

Social play is widespread throughout the animal kingdom. One aspect that has recurrently fascinated scholars is its complex, cooperative nature that requires substantial on-the-fly coordination and improvisation (Bekoff, 2001; Bekoff & Allen, 1998; Palagi, 2006), in comparison to other social activities like grooming or sex that involve more stereotyped activity-specific patterns. To play together, partners need to recognize each other's playful intentions, anticipate each other's movements, adjust the timing and nature of individual acts (Bekoff, 2001; Bekoff, 2004; Palagi, 2006; Palagi, 2008) and adapt their moves to the strength and age of their partner (Fröhlich, Wittig, & Pika, 2016b; Fry, 1987).

Playing together thus seems to require a particular kind of focused attunement to one's partner. This has led some researchers to use social play as a means to study complex cognitive abilities like *shared intentionality*, the skill and motivation to share goals and intentions with others during collaborative interactions (Tomasello, Carpenter, Call, Behne, & Moll, 2005a). To test the existence of shared intentionality in human children and chimpanzees, Warneken, Chen and Tomasello (2006) used interruptions to study how participants valued their joint commitments during social play. They found that human children, but not chimpanzees, attempted to re-engage reluctant human play partners after an interruption. They concluded that chimpanzees lacked an awareness of the joint commitment with a social partner toward a common goal - one of the critical requisites taken as evidence for shared intentionality in humans.

While fully-fledged shared intentionality is probably unique to humans, some related component abilities may be present in other species. Indeed, one way of assessing the differences in such abilities between species is to investigate how coordination is achieved when individuals interact together and the processes involved in the regulation of joint activities in various species (Call, 2009). Accordingly, in this article, we suggest that social play is a form of joint action that constitutes a unique testbed for studying the evolution of shared intentionality.

We take a different tack from previous research, however. We start from an analysis of joint action in humans, describing the step-by-step processes involved in how humans get into, conduct and get out of focused joint actions together in an orderly way (H. Clark, 1996), thereby

jointly constructing the state of "togetherness" characteristic of shared intentionality (Tomasello & Moll, 2010). In human joint actions, participants typically exchange communicative signals to build up a participation framework (Goffman, 1981), which defines the participants of the interaction, the terms on which they are to interact, and the particular content of the action, among other things (H. Clark, 2006). This constitutes the opening phase of the joint action. When engaging in the action proper, or main body, participants continue to exchange signals to progress through the action in a coordinated manner, or to signal their ongoing commitment to the action. Participants also signal to each other their readiness to terminate the action, proceeding through parting rituals like well-wishing, before disbanding in what is known as the closing phase of the joint action. Taken together, the opening, main body and closing constitute three macro-level phases that are common to all joint actions and constitute the behavioral embodiment of the process by which participants achieve and maintain shared intentionality.

We propose to use this process in humans as a systematic framework and yardstick for examining how various animal species get into, maintain, and get out of play bouts. Such a framework could, in turn, shed light on the evolution of shared intentionality, the unique human motivation for sharing psychological states with others (Tomasello et al., 2005a), and the specific human "cognition for interaction" (Levinson, 2006a; Levinson, 2006b). That is, the extent to which the signal exchanges used to coordinate the interactional achievement of shared intentionality in humans are observable in other animal species may constitute a measure of those species' possession of shared intentionality.

In the next section, we describe what a joint action is, how it is coordinated in humans and how the coordination of joint action reveals abilities necessary for shared intentionality. We then outline our framework to comparatively study social play as joint action and how animals and young children solve the coordination problems arising in joint actions (entering into the action, maintaining it, and exiting from it), thereby potentially achieving states of shared intentionality. Using this framework, we review the literature in search for evidence on how human children and different social species of mammals and birds solve these coordination problems. Finally, we discuss the implications of studying play as joint action for understanding the evolution of shared intentionality.

Joint Action: Shared Intentionality as an Interactional Achievement

What is Joint Action?

Joint action involves two or more individuals collaborating to achieve a shared goal, often corresponding to an outcome that no individual could attain alone. The term potentially includes situations ranging from small-scale, brief, ad hoc collaborations (e.g., Ed enlisting John to help him move a bench) to large-scale, long-term actions involving organized social groups (e.g., Hannibal and his army crossing the Alps to invade Rome). In this paper, we will be concerned with small-scale, brief joint actions that involve a handful of participants who share a joint focus of attention and who attempt to coordinate their individual behaviors towards a goal that they all share. Such joint actions feature a collective state of being that has variously been termed intersubjectivity (Merleau-Ponty, 1962), togetheness or shared intentionality (Reddish, Fischer, & Bulbulia, 2013; Tomasello & Carpenter, 2007; Tomasello et al., 2005a; Zlatev, Racine, Sinha, & Itkonen, 2008). Our analytical focus will be on describing how these states are established, maintained, and dissolved.

Several disciplines have investigated joint action in human interaction. In the words of Levinson (2006b, p. 39) "…human interaction belongs in an interdisciplinary no-man's land: it belongs equally to anthropology, sociology, biology, psychology, ethology, but is owned by none of them". Philosophy and pragmatics (Bratman, 1992; Grice, 1975; Sperber & Wilson, 1986) typically analyzes the intentional structure of joint action. Psychology focuses on experimental explorations of the neural and cognitive processes involved (H. Clark, 1996; Sebanz, Bekkering, & Knoblich, 2006). Approaches from the social sciences like ethnomethodology and conversation analysis have described the linguistic and bodily coordination of joint action in natural settings (Sacks, Schegloff, & Jefferson, 1974; Sidnell & Stivers, 2012). Economics and biology have explored its game-theoretical structure (Henrich et al., 2004). Converging evidence from all these fields has led to the establishment of a sophisticated understanding of the phenomena that human interaction produces as well as the cognitive underpinnings entailed by it (Vesper et al., 2016). Increasingly, proposals are emerging for integrative approaches to bridge methodological and epistemological divides (De Ruiter & Albert, 2017; Galantucci & Sebanz, 2009; H. Clark, 1996; H. Clark & Bangerter, 2004; Levinson, 2006a).

In spontaneous joint actions, participants need to accomplish two things together. First,

they need to establish a sense of *joint commitment* by ensuring they are all able, ready and willing to commit to the action (H. Clark, 2006), that they share their goals and intend to do their part, i.e., that they will not free-ride by not contributing their share to the joint effort. Specifically, establishing joint commitment entails coordinating on a number of generic elements: who is to participate, in what roles, what actions will be performed, and when and where they will be performed (H. Clark, 2006). Second, they need to coordinate their individual actions so that these fit together in time and space to bring about the desired outcomes. This is achieved by exchanging signals in real time that help partners to adapt to each other. For example, in the case of joint actions accomplished mainly via talk (e.g. everyday conversations), speakers design their utterances to display their intentions (Grice, 1975) and facilitate their interpretation, whereas adressees display evidence of how they have interpreted those intentions (H. Clark & Schaefer, 1989). In joint actions involving physical actions (e.g., assembling a Lego model), participants design those actions to be visible and informative to their partners (H. Clark & Krych, 2004) while observing their partners' actions to extract information from them (Vesper et al., 2016). This process is called grounding, and is achieved by the exchange of signals that often are produced incidentally or implicitly, in parallel to the main track of conversation (H. Clark & Schaefer, 1989). Grounding is thus the process by which intersubjectivity (Merleau-Ponty, 1962) or shared intentionality (Bratman, 1992; Tomasello & Carpenter, 2007) is attained. Indeed, attributes of shared intentionality include *joint commitment* to a goal, *mutual responsiveness* in the pursuit of the commitment*, role-reversal*, and *mutual support* (Bratman, 1992; Tomasello & Moll, 2010). Thus, shared intentionality can be construed as a transient state of collective being that participants in joint action strive to attain and maintain, or in the terminology of conversation analysis, as an *interactional achievement* (Schegloff, 1982; Schegloff, 1995). As a result, shared intentionality is an ongoing process in joint action. At the same time, in most joint actions, certain phases can be distinguished that are particularly important to its achievement.

Entering into, Maintaining and Dissolving Shared Intentionality: Phases in Joint Action Joint actions are typically initiated with participants working towards establishing joint commitments. These initial steps of recruiting and ratifying participants (consider Ed approaching John on the street and saying *excuse me, do you have a minute?*), co-constructing the content and nature of the interaction and deciding on action location and timing (Goffman, 1981; H. Clark, 2006; Kendon,

1976; Kendon, 2004) lead to the emergence of a section of the interaction variously termed the initiation, entry or (hereafter) *opening* phase. This can be divided into two sub-phases. In the *pre-entry* (attracting attention, checking availability, ratifying participation), participants establish a joint commitment to interact. In the *entry*, they establish a joint commitment to engage in a specific joint action and to the details of its timing and implementation. In spontaneous joint actions, like picking up a bench, these commitments typically emerge incrementally. If Ed asks John to help him pick up a bench, John might first commit to the action overall, and then they might each commit to picking up one side of the bench and finally, lifting up their sides at the exact same time.

Once participants engage in joint action (e.g., carrying the bench from A to B), they must coordinate progress between and within its different steps (Bangerter & H. Clark, 2003) in what is called the *main body* (H. Clark, 1996). Participants coordinate to put together their efforts in an optimal way, which involves correctly anticipating what one's partners will do and timing one's own actions, i.e., doing the right thing at the right time (Sebanz & Knoblich, 2009). Progress within the main body typically is accomplished via ad hoc turn-taking (Levinson, 2016). For example, two children engaging in pretend play might coordinate switching roles within the play session. Or teams might switch sides at halftime in a soccer match. Sometimes, participants may agree to suspend turn-taking rules to allow some of them to take the initiative for a while, as when one participant tells a story to others (Mandelbaum, 2012). Signals are also exchanged during the main body to re-affirm joint commitments, i.e., to reassure partners that the intent of the action is the same. In rough-and-tumble play, for example, ongoing signal exchange is instrumental to reduce the risk of escalation into real aggression (Bekoff, 2001). Joint actions are often interrupted by some external event; when this happens, participants collaborate to coordinate suspending it in an orderly way, and re-establish a sense of joint commitment when reinstating it after having dealt with the interruption. For example, they may ask permission to suspend the interaction, apologize for keeping their partners waiting, check availability when attempting to re-engage or justify the necessity to suspend before reconstructing the topic (Bangerter, Chevalley, & Derouwaux, 2010; Chevalley & Bangerter, 2010).

Finally, to complete a joint action, participants first need to arrive at the mutual conviction that they are both indeed ready to terminate it. In human interaction, participants communicate

this readiness through the exchange of signals, such as *okay*, ensuring that potentially unraised topics can be addressed (Bolden, 2008; Schegloff & Sacks, 1973). Then, they progress through several steps including well-wishing or suggesting continuity of the relationship, reminiscing about the encounter, exchanging leave-taking signals such as *good-bye* and finally, taking leave of each other by hanging up a telephone or walking away (Albert & Kessler, 1976; Bangerter, H. Clark, & Katz, 2004; Broth & Mondada, 2013; H. Clark & French, 1981; Schegloff & Sacks, 1973). These steps collectively comprise the termination, exit or (hereafter) *closing* phase of a joint action. We distinguish two steps of *pre-exit* (establishing mutual awareness of the readiness of participants to end the encounter) and *exit* (terminating the encounter) (Schegloff & Sacks, 1973). This step-wise closing process allows participants to maintain inter-personal relationships beyond the encounter. Violations of conventions, as in unanswered good-byes, can pose threats to relationships and thus participants cooperate to avoid such failures (Schegloff & Sacks, 1973).

Sometimes, opening and closing phases are reduced or even absent. This does not mean that shared intentionality is not achieved or that it is attained automatically; rather such cases reflect the operation of interpersonal conventions or institutionalized procedures. For example, in rule-based games (as opposed to free play), pre-existing common ground provides players with shared behavioral routines and rules that pre-empt many of the coordination problems. Games thus feature reduced entry and exit phases because participants share understanding about the features of the joint actions they are engaged in (H. Clark, 1996). This can happen also in everyday situations in which individuals share the same physical environment for an extended period of time (e.g., co-passengers in a car, workers sharing an open plan office, or toddlers spending their day together in a kindergarten). These situations create a state of *incipient talk* (Schegloff & Sacks, 1973). It is easier for participants to initiate and terminate focused interactions without necessarily engaging in fully fledged opening and closing procedures. For example, an activity may lapse and be picked up again later or an extended pause in a conversation may occur without being interpreted as inappropriate (Berger, Viney, & Rae, 2016), and so on. While states of incipient talk may simplify the requirements of overt communication, they do not completely obviate the need for them, and participants often still mark encounters within such environments, however fleetingly (González-Martínez, Bangerter, & Lê Van, 2017).

Is the Interactional Achievement of Shared Intentionality Unique to Humans? Many species of animals engage in joint actions on a daily basis with members of their group. For example, primates groom each other (Fedurek & Dunbar, 2009), chimpanzees hunt collectively (Boesch, 2002), and many species of animals engage in rough-and-tumble play (Palagi et al., 2015). These activities qualify as cooperative since they require that participants coordinate their behaviours both in time and space. However, human joint action seems unique in the animal kingdom. According to Tomasello et al. (2005a), it is the ability to engage in shared intentionality, or "togetherness", that constitutes the crucial difference between humans and other species. Indeed, participation in interactions involving shared intentionality has transformed human cognition in fundamental ways and underlies other unique human abilities such as language, cultural learning, and pretense (Tomasello & Moll, 2010).

In a similar vein, Levinson (2006b) suggests that the properties of human joint action are expressions of a uniquely human set of capabilities and motivations for social interaction, the human "interaction engine". These include special communicative abilities, such as multimodal signal use (Levinson & Holler, 2014) and alternations of speaker-recipient turns in conversation (Levinson, 2016), special cognitive abilities, such as shared intentionality, and other ethological outputs, such as leave-taking rituals (Levinson, 2006a). Taken together, the elements of this engine enable human "cognition-for-interaction" (Levinson, 2006a) in a way that is independent of language, even though language has evolved as one of the primary means by which humans coordinate joint action. By enabling sophisticated forms of joint action, the human interaction engine paves the way for the emergence of cumulative culture and the development of social institutions. Levinson and Holler (2014) further suggested that the human interaction engine emerged as a step in an increasingly stratified system of communicative competencies that mark the evolution of modern human communication and has potentially evolved around 2 mya with the early forms of Homo.

Thus, it seems unlikely that the ability to engage in shared intentionality appeared suddenly with the genus *Homo* and seems likely that some components of this ability may be found in other species, at least in our ape ancestors. Indeed, apes seem to possess some abilities necessary for understanding shared intentionality, like reading others' attention (Tomasello, Call, & Hare, 1998) and intentions (Call, Hare, Carpenter, & Tomasello, 2004; Call & Tomasello, 1998). They are

also capable of communicating multi-modally to convey meaning (Genty, Clay, Hobaiter, & Zuberbühler, 2014; Hobaiter, Byrne, & Zuberbühler, 2017) and engaging in gestural turn-taking (Fröhlich et al., 2016a; Rossano, 2013). But they have difficulties participating in activities involving shared attention (Melis & Tomasello, 2013; Tomasello & Carpenter, 2005; Tomasello et al., 2005a). Other abilities related to shared intentionality have been claimed to be missing as well, like the ability to communicate declaratively to share attention and interest about external objects or events (Call & Tomasello, 2007; Plooij, 1984; but see Hobaiter, Leavens, & Byrne, 2014), to engage in triadic joint attention (Tomasello, Carpenter, & Hobson, 2005; Tomonaga et al., 2004; Warneken et al., 2006; but see Pika & Zuberbühler, 2008), to engage in active teaching (but see Boesch, 1991) and to offer unprompted help (Warneken, Hare, Melis, Hanus, & Tomasello, 2007). All of these cases taken together suggest that the major difference between apes and humans seems to be the ability and motivation to share psychological states with others (Call, 2009; Tomasello et al., 2005a), which emerges early in human ontogeny at around one year of age, as infants start understanding and then participating in joint actions with others (Carpenter, 2009).

Much of the empirical evidence concerning the lack of shared intentionality in nonhuman animals is based on how they understand shared goals. This is tested for instance by establishing whether, when playing games with human partners, apes attempt to re-engage reluctant partners after interruptions (Warneken et al., 2006). But reinstating an interrupted joint action is only one aspect of the more global interactive process described above by which individuals enter into, maintain, and dissolve a sense of 'togetherness'. We thus call for a more holistic approach to study whether various animal species exhibit phases of opening, main body and closing when engaged in naturally occurring joint activities with peers. The extent to which these phases are observable and the complexity of their coordination could constitute a yardstick for a systematic comparison of the various components of shared intentionality among different species.

We suggest that social play among peers is an ideal testbed for this approach since it is a complex cooperative process that requires substantial on-the-fly coordination (Bekoff, 2001; Bekoff & Allen, 1998; Palagi, 2006) and improvisation, in comparison to other forms of social activities, and is widely shared (and thus comparable) across social species. Moreover, play is only possible given a shared understanding that the behavior implemented is playful and non-

serious. In other words, while in many social activities, one participant might force a partner into doing something (e.g., food sharing or sex), engaging in social play requires both participants' understanding of the irrealis nature of the joint action (E. Clark, 2009) and therefore that they both have a shared intention to play.

So far, however, communicative signals in play have typically been studied in isolation as means to signal particular intentions (e.g., the intention to initiate or to terminate play), but the role of those intentions within the overall cooperative activity of play, i.e. as means to articulate the different phases of the activity to gradually reach a state of shared intentionality, has not been investigated. We thus propose studying how species enter into, maintain and exit from play as a means of gauging shared intentionality. We outline our framework in the next section.

A Comparative Framework for the Study of Shared Intentionality in Social Play

To enable systematic inter-species comparisons, play needs to be analyzed by means of a consistent framework based on human interaction. We now propose such a framework specifically for the study of naturally occurring play bouts. This approach is designed to encompass results from unimodal research (e.g., gestures only: Fröhlich et al., 2016b) and research concentrating on specific moments of the bout (e.g., opening: Fröhlich et al., 2016b; closing: Luef & Liebal, 2013; main body: Palagi, 2008). The framework takes a multimodal perspective (Levinson & Holler, 2014), including all communicative means, such as gaze, body orientation, behaviors, body postures, facial expressions, vocalizations and gestures and analyzes play bouts in their entirety. This framework can produce a more holistic picture of play coordination and a more complete understanding of how shared intentionality is achieved.

As described earlier, human participants' efforts to coordinate joint action emerge as a sequence of moves that can be divided into macro-level phases of *opening*, *main body* and *closing*. These phases allow constructing a framework for comparing data in social play of animals and human children concerning the presence of phases, the complexity of communicative or behavioral means deployed to articulate those phases, and the presence of markers of shared intentionality: *joint attention* and *joint commitment*, *mutual responsiveness* and coordination of *role reversal,* communicative *turn-taking*, *re-engagement* after interruption, *mutual support*, and *leave-taking*

rituals (Bangerter et al., 2010; Bratman, 1992; Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Levinson, 2016; Schegloff & Sacks, 1973; Tomasello & Moll, 2010).

In Table 1 we provide a description of each phase, as well as sub-phases, the coordination problems that lead to the emergence of each sub-phase, and the behaviors and communicative signals typically or potentially deployed to solve those coordination problems. We distinguish two sub-phases for the opening (pre-entry and entry), five for the main body (continuation, type change, role reversal, and suspension and re-engagement after interruption), and two for the closing (pre-exit and exit). In Table 1, sub-phases occur in the sequence depicted in opening and closing, but may vary in their exact sequence in the main body, i.e., they may occur in a different order, get repeated, or not occur at all depending on the circumstances. The framework embodied in the table is derived from the phase structure of adult human joint action as described above, and provides a yardstick for analyzing the achievement of shared intentionality with which to compare animal and human play.

The opening is a phase in which participants establish joint commitment to engage into play (Table 1). We distinguish two sub-phases of *pre-entry* and *entry*. Pre-entry involves selecting participants, orienting toward and approach with the aim of attaining a state of joint attention, and making sure they are ready and willing to interact by establishing joint commitment to an as yet unspecified action. Next, in the entry sub-phase, they jointly commit to the nature (through species-typical play intiation signals) and location (through potentially deictic signals and behaviors) of the play bout, and time its actual beginning.

The main body involves the actual play bout. It begins when play starts, with the first body contact for contact play (e.g., Rough-and-tumble play: Palagi et al., 2015) or the first chase movement for chase play (e.g., Pozis-Francois, Zahavi, & Zahavi, 2004). The main body is itself composed of various sub-phases, depending on how the play bout unfolds. Play *continuation* signals serve to coordinate on the willingness to continue the bout and may occur when participants encourage less active partners. Laughter and play faces in primates (Davila Ross, Owren, & Zimmermann, 2010; Demuru, Ferrari, & Palagi, 2015) are examples. Participants may coordinate a *play type change*, e.g., from contact to chase play. They may reverse roles, e.g., from being the chaser to being chased (Table 1). If an interruption occurs, for example through a loud noise, participants need to coordinate on the *suspension* and the possible *re-engagement* of the play bout.

Finally, closings allow participants to exit from their joint commitments, thereby preventing possible hostile escalations and maintaining social bonds beyond the play action. Participants typically articulate intentions to end play before actually doing so (*pre-exit*), for instance through behavioral or communicative efforts that reduce play intensity or tempo. Pre-exits may be followed by terminations of the interaction (*exit*). Participants might use *leave-taking* signals, to signal the termination of the interaction. Closing processes may allow participants to maintain inter-personal relationships and to prolong the feeling of "togetherness" beyond the encounter (Albert & Kessler, 1976). An exception is for abovementioned states of incipient talk, where individuals share the same environment for prolonged time periods (such as animals living in captive conditions or infants spending their day together in a kindergarten), and where opening and closing phases might be attenuated or simply disappear. Compared to other phases, the structural features of closings in animal play may be attenuated (or absent). If this is the case, it would point to the lack of a sense of "togetherness" as a feature of animal play or children's play.

In the next section, we review the human infant and animal play literature in search of evidence that social play is organized into macro-level phases of opening, main body and closing, and for behaviours and communicative signals used to coordinate them. Guided by the idea that the sequential coordination of joint action in macro-level phases is how people incrementally get into and out of the state of shared intentionality, we aim at looking for behavioural markers of shared intentionality in human children and animal social play coordination.

Table 1

Descriptions of macro-level phases and subphases of play, coordination problems, observable behavioral outputs deployed to solve these coordination problems as potential markers of shared intentionality

Phase	**Subphase**	**Coordination problem**	**Observable behavioral outputs**
Opening	Pre-entry	Selecting participants to establish joint attention	Initiator talks to, gazes at, orients toward or approaches partner, uses attention-

		and joint commitment to interact.	getter. Joint attention is established via mutual gaze.
	Entry	Jointly committing to content, timing, nature and location of the interaction.	Participants use play-initiation signals or play type-specific signals. Participants may also use signals to indicate location, e.g., deictic gestures.
Main body	Continuation	Maintaining joint commitment to the activity.	Participants use behaviors/signals of ongoing engagement, e.g., laughter and play face.
	Play type change	Agreeing to change the activity type, e.g., from chase play to contact play.	Participants use behaviors/signals to change play type.
	Role reversal	Agreeing to change roles or positions during activity, e.g., chaser and chased participants switch roles during chase play.	Participants use behaviors/signals to reverse roles. Participants may use behaviours to help partners fulfil their role.
	Suspension	An interruption occurs. Coordinating on the suspension of the play bout.	Some or all participants switch attention to the interruption stimulus. They agree to suspend the bout.
	Re-engagement after interruption	Coordinating on reinstating the bout.	Participants signal availability to resume the bout. Participants may either use initiation signals or distinct re-engagement signals.

Closing	Pre-exit	Establishing mutual conviction that both participants are ready to stop playing.	Participants use behaviors/signals to reduce tempo and intensity of bout before the termination of play.
	Exit	Disengaging from the interaction.	Participants use leave-taking signals before joint attention is dissolved and participants disengage by walking away or suspending bodily contact (but may remain in proximity, creating state of incipiency).

Applying the Framework: Social Play in Children and Animals

Social Play in Children

Children express social interest in one another from a very young age, but before the age of 18 months, peer social interactions are rare and poorly coordinated (Brownell & Brown, 1992; Eckerman & Peterman, 2001). Cooperative forms of play among peers, such as imitative games, increase between 20 and 24 months of age alongside the emergence of more sophisticated interactional skills to initiate, maintain and coordinate activities (Eckerman, Davis, & Didow, 1989). Between 24 and 30 months of age, children reliably cooperate with each other in problem-solving tasks, but younger children do not (Brownell & Carriger, 1990). It is only during the third year of life, as the child's social understanding and language about self and other develops and they begin to care about social norms and rules of games (Rakoczy & Schmidt, 2013; Rakoczy, Warneken, & Tomasello, 2008), that social games become more coordinated and cooperative (Brownell, Ramani, & Zerwas, 2006; Eckerman & Didow, 1996; Verba, 1994). Indeed, by taking into account their partners' intentions and by monitoring, timing and sequencing their own and their partner's actions, children can adjust their behavior appropriately to attain a shared goal (Barresi & Moore, 1996; Brownell et al., 2006; Smiley, 2001). In terms of the cognitive abilities necessary for shared intentionality understanding, from 12 months on children possess the

motivation to inform others and to share attention and interest via declarative pointing (Liszkowski, Carpenter, Striano, & Tomasello, 2006), but also understand and engage in role reversal (Carpenter, Tomasello, & Striano, 2005). Furthermore, unlike chimpanzees, children from 18 months on attempt to re-engage reluctant partners after interruption of a shared game (Warneken et al., 2006) and show mutual support by helping others achieve their goal (Warneken & Tomasello, 2006). Taken together, then, children, already possess a "we" intentionality (shared intentionality) at least from 14 months of age and act cooperatively, but it is only by 3 years that they become sensitive to joint commitments and begin to understand the obligations and conventions involved in joint action (Gräfenhain et al., 2009; Gräfenhain, Carpenter, & Tomasello, 2013; Kachel, Svetlova, & Tomasello, 2017).

It is unclear how children interactionally achieve shared intentionality via the coordination of phases in naturally occurring joint action with peers. Thus, we now review evidence for how children communicate, verbally or non verbally, to coordinate social play into macro-level phases of opening, main body and closing, focusing on two types of play: rough-and-tumble play (hereafter R&T) and pretend play. We focus on these types because they pose particularly intricate coordination problems. R&T requires coordinating physical actions on the fly and carries the risk of escalating into aggression, whereas pretend play requires coordinating the pretense. Indeed, pretend play has been argued to constitute one of the earliest ontogenetic cases of true shared intentionality (Rakoczy, 2008). We highlight potential behavioural markers of shared intentionality in the coordination of play phases, such as establishment of joint attention and joint commitment in the opening phase, coordination of role reversal, communicative turn-taking, mutual support, and re-engagement after interruption in the main body phase, and leave-taking rituals in the closing phase.

The general characteristic of R&T is that the behaviors performed seem agonistic but are performed in a non-serious context (Smith, 1997) with friends (Blurton-Jones, 1972; Smith, Smees, & Pellegrini, 2004). Human R&T consists of chasing, fleeing, wrestling, grappling, pinning down and delivering restrained blows (Blurton-Jones, 1972; Fry, 2005). To coordinate R&T and avoid escalation into real aggression, children need to metacommunicate to ensure mutual awareness of playful intentions (Fry, 2005; Smith & Boulton, 1990). The patterns of R&T and the play-signals involved appear to be widely comparable across cultures (Fry, 2005). In the opening phase, children often initiate R&T and ratify participants by hitting, pulling hair, using

verbal insult, but at the same time exhibit playful signals, including smiles, laughter or giggles (Fry, 2005; Smith & Boulton, 1990). In the main body, ongoing signal exchange is instrumental to reduce the risk of escalation into real aggression (Bekoff, 2001). For example, smiles and giggles increase when physical contact becomes rougher (Fry, 1987). The cooperative aspect of alternation of turns (role reversal) and reciprocity (mutual responsiveness) is also essential to attenuate competition (Pellis & Pellis, 2017). Role reversals may involve the stronger partner giving advantage to the weaker participant (Fry, 1987). The termination of R&T requires active cooperation by players to de-escalate fights, e.g., by turning their body away from partner (Fry, 2005) or using linguistic markers like "mercy" (Sluckin, 1981). Moreover, after R&T and in contrast to real aggression, partners tend to remain in each other's company (Aldis, 1975; Fry, 1990; Humphreys & Smith, 1984; Smith & Lewis, 1985), suggesting continuation of the relationship beyond the interaction.

Social pretend play, or make-believe play, refers to an activity in which children transform » (…) the Here and Now, You and Me, or the action potential in these features of the situation » (Garvey & Berndt, 1975, p. 4) into some shared imaginative framework (Rakoczy, 2006). Behaviors in pretense are nonliteral or simulative (Fein, 1981; Lillard, 1993). For example, children may engage in object substitution, e.g., pretending a banana is a telephone, or object imagination, e.g., pretending there is a pillow although there is none (Lillard, 1993). The ability to interpret behaviors as "not real" appears in toddlers as early as 18 months of age (Lillard & Witherington, 2004). Mothers play an especially active role in assisting the interpretation of pretend behaviors (Haight, Wang, Fung, Williams, & Mintz, 1999; Haight & Miller, 1992; Lillard, 2007; Miller & Garvey, 1984). The acquisition of symbolic understanding provides an essential fundament to engage in shared pretense with peers at a later stage in development (Bretherton, 1984; Lillard, 1993). Due to its symbolic nature, pretend play coordination is challenging and requires ample metacommunication between players to regulate shared symbolic frameworks (Bretherton, 1984; Stambak & Sinclair, 1990).

In the opening phase, children often initiate pretend play though imitation of the peer's actions, by performing an action complementary to the peer's one, by joining the peer's manipulation of material, or by offering appropriate objects to assist a partner setting up a shared pretense scene (Garvey, 1977; Giffin, 1984; Nelson & Seidman, 1984; Ramani & Brownell, 2013; Schwartzman, 1978; Stockinger Forys & McCune-Nicolich, 1984). Before the development of

full speech, imitation of a peer's nonverbal actions represents an important behavioral strategy to achieve coordination (Eckerman et al., 1989). Later, at the pre-school stage, coordinating pretend play involves more complex forms of cooperation, as linguistic and socio-cognitive skills advance (Garvey, 1977; Miller & Garvey, 1984). Prior to the start of the game, children can establish joint commitment by determining the type and location of the game, and the assignment of roles with the use of explicit verbal social bids, e.g. "Let's play house"; "this is the kitchen"; "I'll be the patient and you'll be the doctor, ok ?" (Bretherton, 1984; Garvey, 1974; Giffin, 1984). Common ground plays a major role in setting up pretend scenarios, especially when activities reflect an event structure borrowed from real-life cultural activities, e.g., baking a cake. To ensure smooth coordination during the main body of pretend play, children need to cooperate by communicating their own - but also accepting other's – ideas (Ramani & Brownell, 2013). Pre-schoolers engage in a continous process of negotiating, discussing, improvising and proposing new features within the game (Ramani & Brownell, 2013). In order to introduce new ideas or to change rules to the current scene, children produce metacommunication or stage directions (E. Clark, 2009; H. Clark, 2016). For example, children may use the past tense to suggest future actions, e.g. "you said you were going to the ball" (E. Clark, 2009), or negotiate roles (e.g., "I'm the mommy now") (Bretherton, 1984; Giffin, 1984). Cooperation by players to maintain joint pretense is especially manifested in their efforts to avoid interruptions, suggesting that children have a mutual understanding that their actions contribute to the joint action (Schwartzman, 1978). In the closing phase, children terminate play by displaying leave-taking signals such as meaningful looks, gestures, verbal markers (Gräfenhain et al., 2009), such as "Let's not play this anymore" (Garvey, 1974; Schwartzman, 1978), or statements helping a player to abandon the play identity, such as "I'm not the dragon anymore" (Garvey & Berndt, 1975; Schwartzman, 1978).

According to Tomasello and Moll (2010), the participation in joint action involving shared intentionality underlies several unique human abilities including pretense. Pretend play is indeed a complex form of play that is specific to humans (Gómez & Martin-Andrade, 2005; Vygotsky, 1967) although it may not be common to all cultures (Gaskins, 2013). Although children seem to be able to engage in joint action from 12 months on (Carpenter, 2009) and to understand shared intentionality from 14 months (Tomasello & Moll, 2010), it is only around 18 months that they start understanding pretense (e.g. Lillard & Witherington, 2004). And it is even later, around the third year of age, as they start understanding social conventions and developing more sophisticated

linguistic competences, that they engage in more coordinated social pretend games (Brownell et al., 2006; Eckerman & Didow, 1996; Verba, 1994). Because the coordination of pretend play relies mostly on verbalization, it is easier to identify markers of shared intentionality in this form of play i.e. joint commitment (e.g. "Let's play doctor together"), role reversal (e.g. "Now it's your turn to be doctor"), re-engagement after interruption (e.g. "Come back, we're not finished!"), mutual support (e.g. "This is the right box") and leave-taking signals (e.g. "I don't want to play anymore").

R&T play, on the other hand, is widely comparable across human cultures (Fry, 2005) and also common to many animal social species (Palagi et al., 2015; Pellis & Pellis, 2017). Although R&T relies less on verbalization, its coordination still requires metacommunication to avoid escalation into real aggression. We thus suggest that the comparative study of how different species solve the various coordination problems inherent in R&T could provide a promising tool to shed light on the evolution of the human unique motivation to share psychological state with others and its special "cognition for interaction". In the next section, we review evidence on how different species of mammals and birds achieve coordination in R&T, and what behaviours and communicative signals are used to articulate the structure of play into opening, main body and closing phases, looking for markers of shared intentionality.

Social Play in Non-Human Animals

In animals, R&T involves behaviours that resemble fighting (e.g., wrestling, tumbling, or chasing) but lack key characteristics of agonistic behaviours: threats are rare or absent, muscles are relaxed, biting is inhibited and non serious intent can be communicated via playfaces and play vocalizations (Palagi et al., 2015; Palagi, Antonacci, & Cordoni, 2007; Pellis, 1984; Smith, 1997). For R&T to be distinguishable from competition and to remain enjoyable, it requires a certain degree of reciprocity (Pellis & Pellis, 2017). Reciprocity is achieved by partners through cooperation, for example by giving the advantage to a currently overpowered partner (e.g. Pellis & Pellis, 2017). We review the evidence starting with species for which communicative signals and behaviors have been documented in the coordination of each phase of play, including openings, main bodies and closings. These species include gorillas, black bears, red-necked wallabies, red kangaroos, and Arabian babblers.

In the opening phase, gorillas (*Gorilla gorilla*) select play partners with ample, silent gestures or audible gestures to attract their attention, e.g. *arm shake* (Tanner, 2004). Once shared

attention is established, *one-handed grab* is commonly used to initiate contact play, whilst *drum object* is used to initiate chase-play (Genty, Breuer, Hobaiter, & Byrne, 2009). During the main body phase of play, *playfaces* and *laughter* serve to maintain play (Palagi et al., 2007). Playfaces are more frequent and intense compared to gentle play, and bouts that include full playfaces (upper teeth exposed) are longer than those with playfaces (upper teeth covered) (Waller & Cherry, 2012). To re-engage partners following interruptions, individuals animate objects and show them to partners to re-establish mutual attention toward the game (Tanner & Byrne, 2010). In the closing phase, to exit from play, gorillas use *hand-on* and *pirouette* gestures (Genty et al., 2009; Luef & Liebal, 2013).

Black bears (*Ursus americanus*) select play partners via approaches in which they communicate intent through subtle positioning of their ears, i.e. *crescent ears* (ears visible but turned laterally from the head) (Henry & Herrero, 1974). Once *joint attention* is established, play is initiated with signals, such as *pawing, biting, rearing* (i.e. holding the fore-paws off the ground in a sitting or standing position) and *head butting.* In the main body, play is maintained via signals such as the *relaxed open-mouth face* and breathing/panting sounds. When physical contact becomes intense, one of the participant faces its partner and moans. If moans are followed by further play, moaners often *flatten their ears*; if this is ignored, players risk being attacked by partners (Henry & Herrero, 1974; Pruitt, 1976), suggesting that flattening of ears signals an intention to terminate the interaction. In the closing phase, before leaving their partner by walking away or running away, bears often *look away*, *lick partner, extend neck and head* or *shrug away* (Henry & Herrero, 1974).

Red-necked wallabies (*Macropus rufogriseus banksianus*) select prospective play partners by *approaching, high-stance posturing* and *orienting toward* (Watson & Croft, 1993). Once joint attention is established, they initiate play by *sniffing, skipping* and *grabbing* the partner (Watson & Croft, 1993). Wallabies coordinate role reversal by self-handicapping their defense and giving the subordinate partner a chance to take advantage through *standing-flat-footed* (Watson & Croft, 1996). In the closing phase, the termination of play is coordinated by one of the participants removing itself from the bout, e.g. by orienting away or moving away from their partners. After termination of R&T, partners often remain close or even face each other, suggesting potential awareness of mutual participation (Watson & Croft, 1993; Watson & Croft, 1996).

Red kangaroos (*Macropus rufus*) exhibit similar behaviors as wallabies to select participants: *Approaching*, *high-stance posturing* and *upright body position* (Croft & Snaith, 1990). They initiate play with *pawing*, *head arching* and *kicking* (Croft & Snaith, 1990; Watson, 1998). To maintain reciprocity and coordinate role reversal during the main body phase, both partners self-handicap their movements, e.g., lowering their kicking rates, to increase the opponents chances (Croft & Snaith, 1990). Kangaroos terminate play by *pushing-away* or *pushing-down* their partners, a signal that reliably leads to terminations of fights between participants (Croft & Snaith, 1990).

Arabian babblers (*Turdoides squamiceps*) select partners via gaze alternation and initiate play with *pendulums* (one participant pushing the other from a branch and grab-holding its foot), *crouching, holding up a twig* on the ground or *bowing* signals (Pozis-Francois et al., 2004). Both *crouching* and *bowing* are also used to re-engage partners after interruptions. If play becomes too rough, abrupt terminations are often signalled via vocal signals (Pozis-Francois et al., 2004). Play is usually terminated when participants stop moving and exit is often followed by affiliative behaviors (*allopreening*) between players (Pozis-Francois et al., 2004).

Evidence for coordination of only some of the phases of play is also available for chimpanzees, bonobos, dogs, coyotes, wolves, dolphins, lemurs, rats, Visayan warty pigs, kakas and keas.

In the opening phase, chimpanzees (*Pan troglodytes*) select prospective participants with audible gestures to attract attention, such as *drum object* (Hobaiter & Byrne, 2014). Once joint attention is established, they initate play with gestures such as *arm shake, dangle*, *gallop, head nod, head stand*, *object in mouth approach, poke, roll over* and *stomp other with two feet* (Hobaiter & Byrne, 2011; Hobaiter & Byrne, 2014). When soliciting play with same-age or younger individuals, participants cooperate by using self-handicapping gestures (Fröhlich et al., 2016b). During the main body, signals such as *play faces*, *laughter* (Davila Ross et al., 2010; Preuschoft, 1992), and *feet shaking* (Hobaiter & Byrne, 2014) serve to maintain play and distinguish it from potentially serious actions. *Galloping* is occasionally used to decrease the intensity from chase-play to contact-play and conversely *hand shaking* to increase intensity from contact-play to chase-play (Hobaiter & Byrne, 2014). If interruptions occur during the bout, participants re-engage their partners by gesturing with *feet shake*, *object in mouth approach*, *head stand* or *roll over* (Hobaiter & Byrne, 2014).

Like chimpanzees, bonobos (*Pan paniscus*) maintain play bouts with *play faces* and *laughter* (Enomoto, 1990; Palagi, 2008). Play faces are more frequent when participants match in age and size and if play includes more physical contact (Palagi, 2008; Palagi & Paoli, 2007). Experimental evidence shows that bonobos re-engage reluctant partners in a social triadic game if conditions are ecologically relevant, e.g. include species-specific gestures and naturalistic play-objects (Pika & Zuberbühler, 2008). Re-engagement signals used to reinstate play in bonobos are *begging* or *grabbing* gestures, often combined with facial expressions, such as *protruded lip* displays (Pika & Zuberbühler, 2008).

To establish joint attention with prospective play partners, dogs (*Canis familiaris*) use attention getters such as *barking* (Bekoff, 1974; Horowitz, 2009). To initiate play they use signals such as *face-pawing*, *bowing* (Bekoff, 1977; also used by coyotes, *Canis latrans* and wolves, *Canis lupus*; Bekoff, 1995), alternating *approach-withdrawals* (Bekoff, 1974), *leap on*, *bow head* and *play slap* (Horowitz, 2009). In the main body, participants coordinate on changes of the play type with the use of *body biting, side-to-side head shaking* and *roll over* to switch from chase to wrestle-play (Bekoff, 1974, p.332). To regulate role reversals and promote reciprocity during the bout, dogs self-handicap via *inhibitory bites* (Bauer & Smuts, 2007). To re-engage a reluctant partner dogs use *biting*, *pawing*, *barking*, *nosing*, *bumping*, *exaggerated retreating* and *presenting* (Horowitz, 2009).

Kakas and keas (*Nestor notabilis* and *Nestor meridionalis*) select prospective play partners with *bouncy hopping* toward them and initiate play with signals such as *head cock* and *roll over* displays (Diamond & Bond, 2003). Keas also further initiate play with signals such as *stiff-leg walk*, directed gaze, or *vertically toss objects* (sometimes in direction of the partner). During play, both keas and kakas coordinate role reversals and maintain reciprocity through self-handicapping signals, such as *rolling over* and *foot pushing* (Diamond & Bond, 2004). In ring-tailed lemurs, (*Lemur catta*), the *relaxed open mouth* display serves to regulate and maintain social play (Palagi, Norscia, & Spada, 2014).

In bottlenose dolphins, (*Tursiops truncatus*), *signature whistles* are characteristic sounds given during R&T with the possible function of giving feedback, promoting a playful mood and distinguishing play from aggression (Blomqvist, Mello, & Amundin, 2005).

In rats, (*Rattus norvegicus*), role reversal is coordinated in the main body through self-handicapping behaviours, such as standing on the partner with four paws once the attacker has overpowered the subordinate player (Foroud & Pellis, 2003; Pellis & Pellis, 2017).

In Visayan warty pigs (*Sus cebifrons*), role reversal and reciprocity in the main body phase is regulated through the production of submissive signals, such as *crouching* (Pellis & Pellis, 2016; Pellis & Pellis, 2017). *Crouching* and *swerving laterally by 90° or more* also often leads to the termination of play (Pellis & Pellis, 2016).

Taken together, across species, this review shows that R&T seems to be organized in macro-level phases of opening, main body and closing. Coordinating these phases relies on species-specific behaviors and communication. However, systematic comparisons are not yet possible, for several reasons. First, much research on animal play signals has mainly focused on communicative signals *per se* (Palagi, 2006; Palagi et al., 2015; Pellis & Pellis, 1996) and not as means to coordinate joint action. Second, the literature often only analyzes single phases of the bout and not the entire sequence as a potential achievement of shared intentionality. Many studies have analysed the signals used to communicate the intention to intiate (e.g., Fröhlich et al., 2016b) or maintain play (e.g., Palagi, 2006), but there is less evidence for the existence of a closing phase. This might indicate the absence of a "we" intentionality or togetherness feeling which would motivate individuals to maintain relationships beyond the encounter. However, there is suggestive evidence for markers of shared intentionality, such as establishment of joint attention, re-engagement after interruption, role reversal and potential leave-taking signals. This evidence is summarized in Table 2, which also includes evidence from children's R&T. We find this promising and call for more comparative research on play as joint action according to our framework to shed light on the building blocks that gradually led to the evolution of the human unique motivation to interact cooperatively and share psychological states with others.

Table 2

Summary of the evidence on communicative signals and behavioral means used to coordinate phases in R&T play in human children and animals. Only subphases and species for which evidence is available are shown

Phase	Subphase	Species	Communicative signals and behaviors
Opening	Pre-entry and entry*	Human children	Verbal utterances (insults, provocation statements), laughter, smiling, body contact (hitting, pulling hair)
		Chimpanzees	Drum object, arm shake, dangle, gallop, head nod, head stand, object in mouth approach, poke at, roll over, stomp other with two feet
		Gorillas	Arm shake, one-handed grab, drum object
		Dogs	Barking, face-pawing, bowing
		Coyotes	Bowing
		Wolves	Bowing
		Black bears	Crescent ear positioning, pawing, biting, rearing, head butting
		Red kangaroos	Approaching, high-stance posturing, positioning body upright, pawing, head arching, kicking
		Red-necked wallabies	Approaching, high-stance posturing, orienting toward, sniffing, skipping, grabbing
		Kakas, Keas	Bouncy hopping, directed gaze, head cock, roll over, stiff-leg walk, toss object
		Arabian babblers	Gaze alternation, pendulums, crouching, holding up twig, bowing
Main body	Continuation	Human children	Smiling, laughter, giggles
		Bonobos	Play face, laughter
		Chimpanzees	Play face, laughter, feet shake
		Gorillas	Play face, laughter
		Ring-tailed lemurs	Play face
		Black bears	Play face, breathing/panting sounds
		Dolphins	Signature whistles

	Play type change	Chimpanzees	Gallop (decrease intensity), hand shake (increase intensity)
		Dogs	Body biting, side-to-side head shaking, roll over (switch from chase to contact play)
	Role reversal	Human Children	Restrained blows (self-handicapping to reduce combat strength)
		Chimpanzees	Self-handicapping to reduce combat strength (unspecified)
		Dogs	**Inhibitory bites (self-handicapping to reduce combat strength)**
		Rats	Standing on partner with four paws (self-handicapping to reduce combat strength)
		Red kangaroos	Lowering of kicking rates (self-handicapping to reduce combat strength)
		Red-necked wallabies	Standing flat-footed (self-handicapping to reduce combat strength)
		Visayan warty pigs	Submissive crouching
		Kakas, Keas	Roll over, foot-pushing (self-handicapping to reduce combat strength)
	Re-engagement after interruption	Bonobos	Begging, grabbing, protruded lip display**
		Chimpanzees	Feet shake, object in mouth approach, head stand, roll over
		Gorillas	Animate/move objects, show objects to partner
		Dogs	Biting, pawing, barking, nosing, bumping, exaggerated retreating, presenting
		Arabian babblers	Crouching, bowing

Closing	Pre-exit and exit*	Human children	Use of linguistic markers ("mercy"), turning body away, remaining in proximity
		Gorillas	Hand-on, pirouette
		Black bears	Moaning, flattened ear positioning, looking away, licking, extending neck/head, shrugging away
		Red kangaroos	Pushing-away, pushing-down
		Red-necked wallabies	Orienting away, moving away, remaining in proximity/facing each other
		Visayan warty pigs	Submissive crouching, swerving laterally by 90°
		Arabian babblers	Vocalization (unspecified), remaining in proximity, allopreening

Note. *Since the literature does not systematically distinguish between subphases of openings and closings in R&T, we collapse the distinction in the table. **Evidence comes from an experimental study (Pika & Zuberbühler, 2008).

Conclusion: Implications for Studying Play as Joint Action Across Species

The study of joint action in humans has led to a rich understanding of the interplay between cognition and communication in the coordination of interdependencies between individuals cooperating to achieve a shared goal. Here we reviewed interdisciplinary research on joint action, which has revealed the importance of shared intentionality as a key feature of joint action in humans. We also described the interactive process by which shared intentionality is achieved, distinguishing between opening, main body and closing phases. Social play, especially R&T play, represents an ideal testbed for a systematic comparative analysis of the interactional achievement of shared intentionality because it requires on-the-fly coordination and improvisation in comparison to other social activities and because it is widely shared across species. Applying a joint action framework to comparatively study social play could offer some insight into the evolutionary significance of social play and shed light on the evolution of human unique

motivation to interact (cognition-for-interaction: Levinson, 2006a) and share psychological states with others (shared intentionality: Tomasello et al., 2005a).

Our framework allows testing the relationship between species' abilities to solve the different coordination problems in play (Table 1) and their overall cooperativeness. It suggests a principled approach to explore the existence of potential components of shared intentionality and how it is achieved in the interactions of nonhuman animals. This could expand the range of situations where evidence of shared intentionality has been looked for. For example, an influential test is based on experimental evidence obtained from chimpanzees and children playing games with experimenters. When cooperative games were interrupted, children tried to re-engage experimenters, but chimpanzees did not (Warneken et al., 2006). This constitutes evidence that chimpanzees do not have a sense of being jointly committed to the same activity and sharing the goals (Warneken et al., 2006). While this study uses an interruption in a joint action to draw conclusions about shared intentionality, our framework theorizes shared intentionality as an interactional achievement, suggesting a range of occasions potentially related to the establishment, maintenance, change/negotiation, interruption, re-engagement and dissolution of joint actions that may constitute situations for exploring shared intentionality. Moreover, while the Warneken et al. (2006) study is experimental and features interactions with a human caretaker in the context of an artificial game, our framework suggests the potential for investigating the achievement of shared intentionality in naturally-occurring joint actions between conspecifics. Play is a prime example of such an activity because it is intrinsically cooperative. Finally, by enabling the systematic study of different species, our framework opens up the possibility to discover more nuanced aspects of shared intentionality. For example, using more naturalistic triadic play interactions in bonobos, Pika and Zuberbühler (2008) found that subjects were very active in their attempts to reengage a human partner, something that is regularly reported from dogs interacting with humans in such ways.

Of course, the fact that species perform coordinative behaviors superficially similar to those humans perform in opening and closing phases does not necessarily constitute evidence of shared intentionality. For example, animal species that produce leave-taking signals may not have the same understanding or interpretation of what they are doing that humans would have. In other words, similarities in behavior do not necessarily reflect similar shared understandings of the

situation (in this regard, experimental studies that test the flexibility of the behaviors remain crucial).

Taken together, our analysis of play as joint action reveals insights into species' capacities to co-construct a state of shared intentionality through the orderly process of play coordination. Such an insight permits to recreate the building blocks that may have led to the fully-fledged cognition-for-interaction (including shared intentionality) underpinning human joint action. Since many of the key attributes taken as evidence for shared intentionality in humans, i.e. *joint commitment*, *mutual responsiveness* and *role reversal* (Bratman, 1992; Tomasello & Moll, 2010) also characterize R&T play in many species outside of our own species, we conclude that the extensive practice of social play may have contributed to the evolution of cognition-for-interaction in humans (Levinson, 2006a).

References

Albert, S., & Kessler, S. (1976). Processes for ending social encounters: The conceptual archaeology of a temporal place. *Journal for the Theory of Social Behaviour*, *6*, 147–170. https://doi.org/10.1111/j.1468-5914.1976.tb00363.x

Aldis, O. (1975). *Play fighting*. New York: Academic Press.

Bangerter, A., Chevalley, E., & Derouwaux, S. (2010). Managing third-party interruptions in conversations: Effects of duration and conversational role. *Journal of Language and Social Psychology*, *29*, 235–244. https://doi.org/10.1177/0261927X09359591

Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science*, *27*, 195–225. https://doi.org/10.1016/S0364-0213(02)00118-0

Bangerter, A., Clark, H. H., & Katz, A. R. (2004). Navigating joint projects in telephone conversations. *Discourse Processes*, *37*, 1–23. https://doi.org/10.1207/s15326950dp3701_1

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*, 107–122. https://doi.org/10.1017/S0140525X00041790

Bauer, E. B., & Smuts, B. B. (2007). Cooperation and competition during dyadic play in domestic dogs, *Canis familiaris*. *Animal Behaviour*, *73*, 489–499. https://doi.org/10.1016/j.anbehav.2006.09.006

Bekoff, M. (1974). Social play and play-soliciting by infant canids. *American Zoologist*, *14*, 323. https://doi.org/10.1093/icb/14.1.323

Bekoff, M. (1977). Social communication in canids: Evidence for the evolution of a stereotyped mammalian display, *197*, 1097–1099. Retrieved from http://animalstudiesrepository.org/acwp_ena/39/

Bekoff, M. (1995). Play signals as punctuation: the structure of social play in canids. *Behaviour*, *132*, 419–429. https://doi.org/10.1163/156853995X00649

Bekoff, M. (2001). Social play behaviour. Cooperation, fairness, trust, and the evolution of morality. *Journal of Consciousness Studies*, *8*, 81–90. Retrieved from http://www.ingentaconnect.com/content/imp/jcs/2001/00000008/00000002/1075

Bekoff, M. (2004). Wild justice and fair play: cooperation, forgiveness, and morality in animals. *Biology and Philosophy*, *19*, 489–520. https://doi.org/10.1007/sBIPH-004-0539-x

Bekoff, M., & Allen, C. (1998). Intentional communication and social play: How and why animals negotiate and agree to play. In M. Bekoff & A. Byers (Eds.), *Animal play: Evolutionary, comparative, and ecological perspectives* (pp. 97–114). Cambridge: Cambridge University Press.

Berger, I., Viney, R., & Rae, J. P. (2016). Do continuing states of incipient talk exist? *Journal of Pragmatics*, *91*, 29–44. https://doi.org/10.1016/j.pragma.2015.10.009

Blomqvist, C., Mello, I., & Amundin, M. (2005). An acoustic play-fight signal in bottlenose dolphins (*Tursiops truncatus*) in human care. *Aquatic Mammals*, *31*, 187–194. https://doi.org/10.1578/AM.31.2.2005.187

Blurton-Jones, N. G. (1972). Categories of child-child interaction. In N. G. Blurton-Jones (Ed.), *Ethological studies of child behaviour* (pp. 97–127). Camrbidge: Cambridge University Press.

Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour*, *41*, 530–532.

Boesch, C. (2002). Cooperative hunting roles among taï chimpanzees. *Human Nature*, *13*, 27–46. https://doi.org/10.1007/s12110-002-1013-6

Bolden, G. B. (2008). Reopening russian conversations: The discourse particle –to and the negotiation of interpersonal accountability in closings. *Human Communication Research*, *34*, 99–136. https://doi.org/10.1111/j.1468-2958.2007.00315.x

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*, 327–341. https://doi.org/10.2307/2185537

Bretherton, I. (Ed.). (1984). *Symbolic play: The development of social understanding*. Orlando, FL: Academic Press.

Broth, M., & Mondada, L. (2013). Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics*, *47*, 41–58. https://doi.org/10.1016/j.pragma.2012.11.016

Brownell, C. A., & Brown, E. (1992). Peers and play in infants and toddlers. In V. B. Van Hasselt & M. Hersen (Eds.), *Handbook of social development: A lifespan perspective* (pp. 183–200). New York: Plenum.

Brownell, C. A., & Carriger, M. S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development*, *61*, 1164–1174. https://doi.org/10.2307/1130884

Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a social partner with peers: Cooperation and social understanding in one- and two-year-olds. *Child Development*, *77*, 803–821. https://doi.org/10.1111/j.1467-8624.2006.t01-1-.x-i1

Call, J. (2009). Contrasting the social cognition of humans and nonhuman apes: The shared intentionality hypothesis. *Topics in Cognitive Science*, *1*, 368–379. https://doi.org/10.1111/j.1756-8765.2009.01025.x

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). "Unwilling" versus "unable": chimpanzees' understanding of human intentional action. *Developmental Science*, *7*, 488–498. https://doi.org/10.1111/j.1467-7687.2004.00368.x

Call, J., & Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, *112*, 192–206. https://doi.org/10.1037/0735-7036.112.2.192

Call, J., & Tomasello, M. (2007). *The gestural communication of apes and monkeys*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, *1*, 380–392. https://doi.org/10.1111/j.1756-8765.2009.01026.x

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy*, *8*, 253–278. https://doi.org/10.1207/s15327078in0803_4

Chevalley, E., & Bangerter, A. (2010). Suspending and reinstating joint activities with dialogue. *Discourse Processes*, *47*, 263–291. https://doi.org/10.1080/01638530902959935

Clark, E. V. (2009). First language acquisition. Cambridge: Cambridge University Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2006). Social actions, social commitments. In N. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 126–150). Oxford, England: Berg.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review*, *123*, 324–347. https://doi.org/10.1037/rev0000026

Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 25–49). London: Palgrave Macmillan UK. Retrieved from http://dx.doi.org/10.1057/9780230524125_2

Clark, H. H., & French, J. W. (1981). Telephone goodbyes. *Language in Society*, *10*, 1–19. https://doi.org/10.1017/S0047404500008393

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81. https://doi.org/10.1016/j.jml.2003.08.004

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259–294. https://doi.org/10.1207/s15516709cog1302_7

Croft, D. B., & Snaith, F. (1990). Boxing in red kangaroos, *Macropos Rufus*: Aggression or play? *International Journal of Comparative Psychology*, *4*, 221–236.

Davila Ross, M., Owren, M. J., & Zimmermann, E. (2010). The evolution of laughter in great apes and humans. *Communicative and Integrative Biology*, *3*, 191–194. https://doi.org/10.4161/cib.3.2.10944

De Ruiter, J. P., & Albert, S. (2017). An appeal for a methodological fusion of conversation analysis and experimental psychology. *Research on Language and Social Interaction*, *50*, 90–107. https://doi.org/10.1080/08351813.2017.1262050

Demuru, E., Ferrari, P. F., & Palagi, E. (2015). Emotionality and intentionality in bonobo playful communication. *Animal Cognition*, *18*, 333–344. https://doi.org/10.1007/s10071-014-0804-6

Diamond, J., & Bond, A. B. (2003). A comparative analysis of social play in birds. *Behaviour*, *140*, 1091–1115. https://doi.org/https://doi.org/10.1163/156853903322589650

Diamond, J., & Bond, A. B. (2004). Social play in kaka (*Nestor meridionalis*) with comparisons to kea (*Nestor notabilis*). *Behaviour*, *141*, 777–798. https://doi.org/10.1163/1568539042265680

Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, *60*, 440–453. https://doi.org/10.2307/1130988

Eckerman, C. O., & Didow, S. M. (1996). Nonverbal imitation and toddlers' mastery of verbal means of achieving coordinated action. *Developmental Psychology*, *32*, 141–152. https://doi.org/10.1037/0012-1649.32.1.141

Eckerman, C. O., & Peterman, K. (2001). Peers and infant social/communicative development. *Blackwell Handbook of Infant Development*, 326–350.

Enomoto, T. (1990). Social play and sexual behavior of the bonobo (*Pan paniscus*) with special reference to flexibility. *Primates*, *31*, 469–480. https://doi.org/10.1007/BF02382531

Fedurek, P., & Dunbar, R. I. M. (2009). What does mutual grooming tell us about why chimpanzees groom? *Ethology*, *115*, 566–575. https://doi.org/10.1111/j.1439-0310.2009.01637.x

Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, *52*, 1095–1118. https://doi.org/10.2307/1129497

Foroud, A., & Pellis, S. M. (2003). The development of "roughness" in the play fighting of rats: A Laban Movement Analysis perspective. *Developmental Psychobiology*, *42*, 35–43. https://doi.org/10.1002/dev.10088

Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016a). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, *6*, 25887. https://doi.org/10.1038/srep25887

Fröhlich, M., Wittig, R. M., & Pika, S. (2016b). Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science*, *3*. https://doi.org/10.1098/rsos.160278

Fry, D. P. (1987). Differences between playfighting and serious fighting among Zapotec children. *Ethology and Sociobiology*, *8*, 285–306. https://doi.org/10.1016/0162-3095(87)90029-X

Fry, D. P. (1990). Play aggression among Zapotec children: Implications for the practice hypothesis. *Aggressive Behavior*, *16*, 321–340. https://doi.org/10.1002/1098-2337(1990)16:5<321::AID-AB2480160504>3.0.CO;2-D

Fry, D. P. (2005). Rough-and-tumble social play in humans. In A. D. Pellegrini & P. K. Smith (Eds.), *The nature of play: Great apes and humans* (pp. 54–85). New York: Guilford Press.

Galantucci, B., & Sebanz, N. (2009). Joint action: Current perspectives. *Topics in Cognitive Science*, *1*, 255–259. https://doi.org/10.1111/j.1756-8765.2009.01017.x

Garvey, C. (1974). Some properties of social play. *Merrill-Palmer Quarterly of Behavior and Development*, *20*, 163–180.

Garvey, C. (1977). Play with language and speech. *Child Discourse*, 27–47.

Garvey, C., & Berndt, R. (1975). The organization of pretend play. Presented at the Paper presented at the Annual Meeting of the American Psychological Association, August, Chicago. Retrieved from https://eric.ed.gov/?id=ED114891

Gaskins, S. (2013). Pretend play as culturally constructed activity. In M. Taylor (Ed.), *The Oxford handbook of the development of imagination* (pp. 224–247). New York, NY: Oxford University Press.

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Animal Cognition*, *12*, 527–546. https://doi.org/10.1007/s10071-009-0213-4

Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PloS One*, *9*, e84738. https://doi.org/https://doi.org/10.1371/journal.pone.0084738

Giffin, H. (1984). The coordination of meaning in the creation of a shared make-believe reality. In *Symbolic play: The development of social understanding* (pp. 73–100). Orlando, FL: Academic Press.

Goffman, E. (1981). Footing. In E. Goffman (Ed.), *Forms of talk* (pp. 124–159). Philadelphia: University of Pennsylvania Press.

Gómez, J.-C., & Martin-Andrade, B. (2005). Fantasy play in apes. In A. D. Pellegrini & P. K. Smith (Eds.), *The nature of play: Great apes and humans* (pp. 139–172). New York: Guilford Press.

González-Martínez, E., Bangerter, A., & Lê Van, K. (2017). Passing-by "Ça va?" checks in clinic corridors. *Semiotica*, 1–42. https://doi.org/10.1515/sem-2015-0107

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*, 1430–1443. https://doi.org/10.1037/a0016122

Gräfenhain, M., Carpenter, M., & Tomasello, M. (2013). Three-year-olds' understanding of the consequences of joint commitments. *PLOS ONE*, *8*, e73039. https://doi.org/10.1371/journal.pone.0073039

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (Vol. 3, pp. 41–58). New York: Academic Press.

Haight, W. L., Wang, X., Fung, H. H., Williams, K., & Mintz, J. (1999). Universal, developmental, and variable aspects of young children's play: A cross-cultural comparison of pretending at home. *Child Development*, *70*, 1477–1488. https://doi.org/10.1111/1467-8624.00107

Haight, W., & Miller, P. J. (1992). The development of everyday pretend play: A longitudinal study of mothers' participation. *Merrill-Palmer Quarterly*, *38*, 331–349.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (Eds.). (2004). *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. New York: Oxford University Press.

Henry, J. D., & Herrero, S. (1974). Social play in the american black bear: Its similarity to canid social play and an examination of its identifying characteristics. *American Zoologist*, *14*, 371–389. https://doi.org/https://doi.org/10.1093/icb/14.1.371

Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*, 745–767. https://doi.org/10.1007/s10071-011-0409-2

Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology*, *24*, 1596–1600. https://doi.org/10.1016/j.cub.2014.05.066

Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, *71*, 96. https://doi.org/10.1007/s00265-017-2325-1

Hobaiter, C., Leavens, D. A., & Byrne, R. W. (2014). Deictic gesturing in wild chimpanzees (*Pan troglodytes*)? Some possible cases. *Journal of Comparative Psychology*, *128*, 82–87. https://doi.org/10.1037/a0033757

Horowitz, A. (2009). Attention to attention in domestic dog (*Canis familiaris*) dyadic play. *Animal Cognition*, *12*, 107–118. https://doi.org/10.1007/s10071-008-0175-y

Humphreys, A. P., & Smith, P. K. (1984). Rough-and-tumble in preschool and playground. In P. K. Smith (Ed.), *Play in animals and humans* (pp. 241–266). Oxford, England: Blackwell.

Kachel, U., Svetlova, M., & Tomasello, M. (2017). Three-year-olds' reactions to a partner's failure to perform her role in a joint commitment. *Child Development*, 1–13. https://doi.org/10.1111/cdev.12816

Kendon, A. (1976). The F-formation system: The spatial organization of social encounters. *Man-Environment Systems*, *6*, 291–296.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.

Levinson, S. C. (2006a). Cognition at the heart of human interaction. *Discourse Studies*, *8*, 85–93. https://doi.org/10.1177/1461445606059557

Levinson, S. C. (2006b). On the human" interaction engine." In N. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford, England: Berg.

Levinson, S. C. (2016). Turn-taking in human communication – origins and implications for language processing. *Trends in Cognitive Sciences*, *20*, 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 B: Biological Sciences*, *369*. https://doi.org/10.1098/rstb.2013.0302

Lillard, A. (2007). Guided participation: How mothers structure and children understand pretend play. In A. Göncü & S. Gaskins (Eds.), *Play and development: Evolutionary, sociocultural, and functional perspectives* (pp. 131–153). Mahwah, NJ: Lawrence Erlbaum.

Lillard, A. S. (1993). Pretend play skills and the child's theory of mind. *Child Development*, *64*, 348–371. https://doi.org/10.1111/j.1467-8624.1993.tb02914.x

Lillard, A. S., & Witherington, D. C. (2004). Mothers' behavior modifications during pretense and their possible signal value for toddlers. *Developmental Psychology*, *40*, 95–113. https://doi.org/10.1037/0012-1649.40.1.95

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*, 173–187. https://doi.org/10.1207/s15327647jcd0702_2

Luef, E. M., & Liebal, K. (2013). The hand-on gesture in gorillas (*Gorilla gorilla*). *Interaction Studies*, *14*, 44–61. https://doi.org/10.1075/is.14.1.04lue

Mandelbaum, J. (2012). Storytelling in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 492–507). Chichester: Wiley-Blackwell.

Melis, A. P., & Tomasello, M. (2013). Chimpanzees' (*Pan troglodytes*) strategic helping in a collaborative task. *Biology Letters*, *9*. https://doi.org/10.1098/rsbl.2013.0009

Merleau-Ponty, M. (1962). *Phenomenology of perception*. (C. Smith, Trans.). London: Routledge and Kegan Paul.

Miller, P., & Garvey, C. (1984). Mother-baby role play: Its origins in social support. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 101–130). Orlando, FL: Academic Press.

Nelson, K., & Seidman, S. (1984). Playing with scripts. In *Symbolic Play: The development of social understanding* (pp. 45–71). Orlando, FL: Academic Press.

Palagi, E. (2006). Social play in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): Implicationsfor natural social systems and interindividual relationships. *American Journal of Physical Anthropology*, *129*, 418–426. https://doi.org/10.1002/ajpa.20289

Palagi, E. (2008). Sharing the motivation to play: the use of signals in adult bonobos. *Animal Behaviour*, *75*, 887–896.

Palagi, E., Antonacci, D., & Cordoni, G. (2007). Fine-tuning of social play in juvenile lowland gorillas (*Gorilla gorilla gorilla*). *Developmental Psychobiology*, *49*, 433–445. https://doi.org/10.1002/dev.20219

Palagi, E., Burghardt, G. M., Smuts, B., Cordoni, G., Dall'Olio, S., Fouts, H. N., … Pellis, S. M. (2015). Rough-and-tumble play as a window on animal communication. *Biological Reviews*, *91*, 311–327. https://doi.org/10.1111/brv.12172

Palagi, E., Norscia, I., & Spada, G. (2014). Relaxed open mouth as a playful signal in wild ring-tailed lemurs. *American Journal of Primatology*, *76*, 1074–1083. https://doi.org/10.1002/ajp.22294

Palagi, E., & Paoli, T. (2007). Play in adult bonobos (*Pan paniscus*): modality and potential meaning. *American Journal of Physical Anthropology*, *134*, 219–225.

Pellis, S. M. (1984). Two aspects of play-fighting in a captive group of oriental small-clawed otters *Amblonyx cinerea*. *Zeitschrift Für Tierpsychologie*, *65*, 77–83. https://doi.org/10.1111/j.1439-0310.1984.tb00374.x

Pellis, S. M., & Pellis, V. C. (1996). On knowing it's only play: The role of play signals in play fighting. *Aggression and Violent Behavior*, *1*, 249–268. https://doi.org/10.1016/1359-1789(95)00016-X

Pellis, S. M., & Pellis, V. C. (2016). Play fighting in Visayan warty pigs (*Sus cebifrons*): insights on restraint and reciprocity in the maintenance of play. *Behaviour*, *153*, 727–747. https://doi.org/10.1163/1568539X-00003346

Pellis, S. M., & Pellis, V. C. (2017). What is play fighting and what is it good for? *Learning & Behavior*, 1–12. https://doi.org/10.3758/s13420-017-0264-3

Pika, S., & Zuberbühler, K. (2008). Social games between bonobos and humans: evidence for shared intentionality? *American Journal of Primatology*, *70*, 207–210. https://doi.org/10.1002/ajp.20469

Plooij, F. X. (1984). The behavioral development of free-living chimpanzee babies and infants. *Monographs on Infancy*, 207.

Pozis-Francois, O., Zahavi, A., & Zahavi, A. (2004). Social play in arabian babblers. *Behaviour*, *141*, 425–450. https://doi.org/10.1163/156853904323066720

Preuschoft, S. (1992). "Laughter" and "smile" in barbary macaques (*Macaca sylvanus*). *Ethology*, *91*, 220–236. https://doi.org/10.1111/j.1439-0310.1992.tb00864.x

Pruitt, C. H. (1976). Play and agonistic behavior in young captive black bears. *Bears: Their Biology and Management*, *3*, 79–86. https://doi.org/10.2307/3872757

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognition, Joint Action and Collective Intentionality*, *7*, 113–127. https://doi.org/10.1016/j.cogsys.2005.11.008

Rakoczy, H. (2008). Pretence as individual and collective intentionality. *Mind & Language*, *23*, 499–517. https://doi.org/10.1111/j.1468-0017.2008.00357.x

Rakoczy, H., & Schmidt, M. F. H. (2013). The early ontogeny of social norms. *Child Development Perspectives*, *7*, 17–21. https://doi.org/10.1111/cdep.12010

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, *44*, 875–881. https://doi.org/10.1037/0012-1649.44.3.875

Ramani, G. B., & Brownell, C. A. (2013). Preschoolers' cooperative problem solving: Integrating play and problem solving. *Journal of Early Childhood Research*, *12*, 92–108. https://doi.org/10.1177/1476718X13498337

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PLOS ONE*, *8*, e71182. https://doi.org/10.1371/journal.pone.0071182

Rossano, F. (2013). Sequence organization and timing of bonobo mother-infant interactions. *Interaction Studies*, *14*, 160–189. https://doi.org/10.1075/is.14.2.02ros

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735. https://doi.org/10.2307/412243

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh"and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk. Georgetown university round table on languages and linguistics 1981* (pp. 71–93). Washington, DC: Georgetown University Press.

Schegloff, E. A. (1995). Discourse as an interactional achievement III: The omnirelevance of action. *Research on Language and Social Interaction*, *28*, 185–211. https://doi.org/10.1207/s15327973rlsi2803_2

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*, 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Schwartzman, H. B. (1978). *Transformations: The anthropology of children's play*. New York: Plenum Press.

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

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*, 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x

Sidnell, J., & Stivers, T. (Eds.). (2012). *The handbook of conversation analysis*. Chichester: Wiley-Blackwell.

Sluckin, A. (1981). *Growing up in the playground: The social development of children* (Vol. 51). London: Routledge and Kegan Paul.

Smiley, P. A. (2001). Intention understanding and partner-sensitive behaviors in young children's peer interactions. *Social Development*, *10*, 330–354. https://doi.org/10.1111/1467-9507.00169

Smith, P. K. (1997). Play fighting and real fighting. In A. Schmitt, K. Atzwanger, K. Grammer, & K. Schäfer (Eds.), *New aspects of human ethology* (pp. 47–64). Boston, MA: Springer US. https://doi.org/10.1007/978-0-585-34289-4_3

Smith, P. K., & Boulton, M. (1990). Rough-and-tumble play, aggression and dominance: Perception and behaviour in children's encounters. *Human Development*, *33*, 271–282. https://doi.org/DOI:10.1159/000276524

Smith, P. K., & Lewis, K. (1985). Rough-and-tumble play, fighting, and chasing in nursery school children. *Ethology and Sociobiology*, *6*, 175–181. https://doi.org/10.1016/0162-3095(85)90029-9

Smith, P. K., Smees, R., & Pellegrini, A. D. (2004). Play fighting and real fighting: Using video playback methodology with young children. *Aggressive Behavior*, *30*, 164–173. https://doi.org/10.1002/ab.20013

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Stambak, M., & Sinclair, H. (Eds.). (1990). *Pretend play among 3-year-olds*. Hillsdale, NJ: Lawrence Erlbaum.

Stockinger Forys, S. K., & McCune-Nicolich, L. (1984). Shared pretend: Sociodramatic play at 3 years of age. In I. Bretherton (Ed.), *Symbolic Play: The development of social understanding* (pp. 159–191). Orlando, FL: Academic Press.

Tanner, J. E. (2004). Gestural phrases and gestural exchanges by a pair of zoo-living lowland gorillas. *Gesture*, *4*, 1–24. https://doi.org/10.1075/gest.4.1.02tan

Tanner, J. E., & Byrne, R. W. (2010). Triadic and collaborative play by gorillas in social games with objects. *Animal Cognition*, *13*, 591–607. https://doi.org/10.1007/s10071-009-0308-y

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*, 1063–1069. https://doi.org/10.1006/anbe.1997.0636

Tomasello, M., & Carpenter, M. (2005). Intention reading and imitative learning. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 133–148). MIT Press.

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*, 121–125. https://doi.org/10.1111/j.1467-7687.2007.00573.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005a). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–735.

Tomasello, M., Carpenter, M., & Hobson, R. P. (2005b). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, *70*, i-152.

Tomasello, M., & Moll, H. (2010). The gap is social: human shared intentionality and culture. In P. M. Kappeler & J. B. Silk (Eds.), *Mind the gap. Tracing the origins of human universals* (pp. 331–349). Berlin: Springer.

Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., Kosugi, D., Mizuno, Y., … Bard, K. A. (2004). Development of social cognition in infant chimpanzees (*Pan troglodytes*): Face recognition, smiling, gaze, and the lack of triadic interactions. *Japanese Psychological Research*, *46*, 227–235. https://doi.org/10.1111/j.1468-5584.2004.00254.x

Verba, M. (1994). The beginnings of collaboration in peer interaction. *Human Development*, *37*, 125–139.

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., … Nijssen, S. (2016). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*, *7*, 2039. https://doi.org/https://doi.org/10.3389/fpsyg.2016.02039

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet Psychology*, *5*, 6–18. https://doi.org/10.2753/RPO1061-040505036

Waller, B. M., & Cherry, L. (2012). Facilitating play through communication: Significance of teeth exposure in the gorilla play face. *American Journal of Primatology*, *74*, 157–164. https://doi.org/10.1002/ajp.21018

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*, 640–663. https://doi.org/10.1111/j.1467-8624.2006.00895.x

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLOS Biology*, *5*, e184. https://doi.org/10.1371/journal.pbio.0050184

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*, 1301. https://doi.org/10.1126/science.1121448

Watson, D. M. (1998). Kangaroos at play: play behaviour in the Macropodoidea. In M. Bekoff & J. A. Byers (Eds.), *Animal play: Evolutionary, comparative, and ecological perspectives* (pp. 61–95). Cambridge: Cambridge University Press.

Watson, D. M., & Croft, D. B. (1993). Playfighting in captive red-necked wallabies, *Macropus rufogriseus banksianus*. *Behaviour*, *126*, 219–245. https://doi.org/10.1163/156853993X00128

Watson, D. M., & Croft, D. B. (1996). Age-related differences in playfighting strategies of captive male red-necked wallabies (*Macropus rufogriseus banksianus*). *Ethology*, *102*, 336–346. https://doi.org/10.1111/j.1439-0310.1996.tb01129.x

Zlatev, J., Racine, T. P., Sinha, C., & Itkonen, E. (Eds.). (2008). *The shared mind: Perspectives on intersubjectivity*. Amsterdam: John Benjamins Publishing Company.

Appendix 2

ESM for Chapter 3

S2.1 Information on site, species, focal individuals and hours of observation.

Site	Species	Focal ID	Birthdate	Sex	Hours observation
San Diego Zoo, USA	*Bonobo*	Belle	2013-12-21	F	*30*
		Erin	1991-12-23	M	*30*
		Kalli	2005-03-15	F	*30*
		Lisa	1991-06-14	F	*30*
		Loretta	1974-01-22	F	*30*
		Maddie	2009-03-24	F	*30*
		Makasi	2004-04-22	M	*30*
		Mali	2007-09-04	F	*30*
		Vic	2001-06-27	M	*30*
		***n* = 9**			***270***
La Vallée des Singes, France	*Bonobo*	Daniela	1968-06-17	F	*20.75*
		David	2001-07-27	M	*20.25*
		Diwani	1996-08-11	M	*20.25*
		Kelele	2004-07-22	M	*22*
		Khalessi	2012-12-12	F	*20.25*
		Khaya	2001-10-19	F	*20.5*
		Lingala	2003-07-17	F	*21*
		Lokoro	2015-05-22	M	*21.25*
		Loto	2009-09-02	M	*20.5*
		Lucy	2003-12-01	F	*20.25*
		Moko	2012-08-04	M	*20*
		Swahili	2014-09-21	F	*20.75*
		Ukela	1985-12-19	F	*20.25*
		Ulindi	1993-10-10	F	*21*
		Yahimba	2009-08-07	F	*21*
		Yuli	2014-07-14	F	20.25
		***n* = 16**			**330.25**
	Chimpanzee	Cauna	2007-05-05	F	*37.1*
		Conan	1996-03-20	M	*37.4*
		Jorg	1995-01-01	M	*37.1*
		Lila	2009-06-20	F	*37.4*
		Panya	2008-07-14	F	*36.1*
		Roy	1997-09-27	M	*37.2*
		Wonder	1997-03-19	M	*36.8*
		***n* = 7**			***259.1***
	Chimpanzee	Benga	1979-09-17	F	*30*

Basel Zoo, Switzerland		Colebe	2005-07-12	M	*30*
		Fahamu	2008-08-19	M	*30*
		Fifi	1993-05-12	F	*30*
		Garissa	2009-04-06	F	*30*
		Jacky	1967-06-30	F	*30*
		Kitoko	1993-05-20	F	*30*
		Kume	2003-10-04	M	*30*
		Xindra	1975-10-23	F	*30*
		***n* = 9**			***270***
La Réserve Africaine de Sigean, France	*Chimpanzee*	Ann	2002-07-21	F	*18.9*
		Goldie	1989-08-29	F	*19*
		Inongo	2009-06-19	M	*18.8*
		Jessica	1973	F	*18.8*
		Macourie	1999-12-12	F	*18.6*
		Madingo	2013-09-02	M	*17.8*
		Pablo	1999-11-12	M	*19.3*
		Penny	1989	F	*19.3*
		Poppy	2013-04-13	F	*18.7*
		n=9			***169.2***
Grand total		***n= 50***			1298.55

S2.2 Summary of Results for Bayesian GLMMs and LMMs

Model 1 (RQ2): Use of mutual exit types (DV, *n*=937)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	-0.24	0.18	[-0.61;0.12]	40480	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.43	0.19	[-0.81;-0.05]	35793	1.00	Student t (3,0,10)
Age difference in years	0.00	0.01	[-0.01;0.02]	21286	1.00	Student t (3,0,10)
Sex						
Different	*Reference*					
Same	-0.16	0.15	[-0.45;0.13]	49520	1.00	Student t (3,0,10)
Context						
Grooming	*Reference*					
Mixed	0.77	0.32	[0.14;1.41]	55958	1.00	Student t (3,0,10)
Play	1.01	0.18	[0.67;1.36]	48363	1.00	Student t (3,0,10)
Random effects						
Initiator	0.27	0.15	[0.02;0.57]	6697	1.00	Student t (3,0,10)
Partner	0.19	0.13	[0.01;0.47]	9175	1.00	Student t (3,0,10)

Model 2 (RQ3): Entry phase presence (dependent variable, DV, *n*=1215)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	2.77	0.29	[2.22;3.37]	19147	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-1.65	0.32	[-2.32;-1.04]	16661	1.00	Student t (3,0,10)
Age difference in years	-0.01	0.01	[-0.02;0.01]	22506	1.00	Student t (3,0,10)
Sex						
Different	*Reference*					
Same	-0.57	0.18	[-0.92;-0.23]	32492	1.00	Student t (3,0,10)
DSI	-0.38	0.16	[-0.69;-0.06]	21819	1.00	Student t (3,0,10)
Rank difference	-0.01	0.29	[-0.59;0.55]	16592	1.00	Student t (3,0,10)
DSI * Species	0.33	0.19	[-0.05;0.70]	20838	1.00	Student t (3,0,10)
Rank difference * Species	-0.02	0.30	[-0.61;0.57]	20216	1.00	Student t (3,0,10)

Random effects						
Initiator	0.26	0.15	[0.01;0.59]	6490	1.00	Student t (3,0,10)
Partner	0.72	0.18	[0.40;1.12]	7973	1.00	Student t (3,0,10)

Model 3 (RQ3): Entry phase duration (DV, *n*=912)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	1.97	0.09	[1.80;2.15]	32827	1.00	Student t (3,2,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.15	0.11	[-0.36;0.06]	27753	1.00	Student t (3,2,10)
Age difference in years	-0.00	0.00	[-0.01;0.00]	30663	1.00	Student t (3,2,10)
Sex						
Different	*Reference*					
Same	-0.04	0.08	[-0.20;0.11]	47612	1.00	Student t (3,2,10)
DSI	-0.18	0.06	[-0.29;-0.06]	33385	1.00	Student t (3,2,10)
Rank difference	-0.19	0.09	[-0.38;-0.01]	16671	1.00	Student t (3,2,10)
DSI * Species	0.17	0.08	[0.01;0.33]	32688	1.00	Student t (3,2,10)
Rank difference * Species	0.13	0.10	[-0.07;0.33]	20834	1.00	Student t (3,2,10)
Random effects						
Initiator	0.13	0.07	[0.01;0.27]	7014	1.00	Student t (3,0,10)
Partner	0.23	0.06	[0.10;0.35]	8137	1.00	Student t (3,0,10)

Model 4 (RQ3): Exit phase presence (DV, *n*=1129)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	3.28	0.36	[2.63;4.03]	19494	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-1.50	0.37	[-2.26;-0.83]	21585	1.00	Student t (3,0,10)
Age difference in years	-0.00	0.01	[-0.02;0.01]	28826	1.00	Student t (3,0,10)
Sex						
Different	*Reference*					
Same	0.27	0.22	[-0.15;0.70]	34391	1.00	Student t (3,0,10)
DSI	-0.74	0.22	[-1.18;-0.32]	18007	1.00	Student t (3,0,10)
Rank difference	0.18	0.35	[-0.49;0.86]	20094	1.00	Student t (3,0,10)
DSI * Species	0.69	0.25	[0.20;1.19]	18864	1.00	Student t (3,0,10)
Rank difference * Species	-0.44	0.36	[-1.15;0.28]	19765	1.00	Student t (3,0,10)
Random effects						
Initiator	0.29	0.17	[0.01;0.66]	7585	1.00	Student t (3,0,10)
Partner	0.27	0.18	[0.01;0.69]	7313	1.00	Student t (3,0,10)

Model 5 (RQ3): Exit phase duration (DV, *n*=1002)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	2.27	0.08	[2.11;2.42]	29823	1.00	Student t (3,2,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.09	0.09	[-0.26;0.08]	23105	1.00	Student t (3,2,10)
Age difference in years	0.00	0.00	[-0.00;0.01]	21687	1.00	Student t (3,2,10)
Sex						
Different	*Reference*					
Same	0.01	0.07	[-0.12;0.14]	40400	1.00	Student t (3,2,10)
DSI	-0.22	0.06	[-0.34;-0.10]	22748	1.00	Student t (3,2,10)
Rank difference	-0.17	0.08	[-0.32;-0.02]	19435	1.00	Student t (3,2,10)

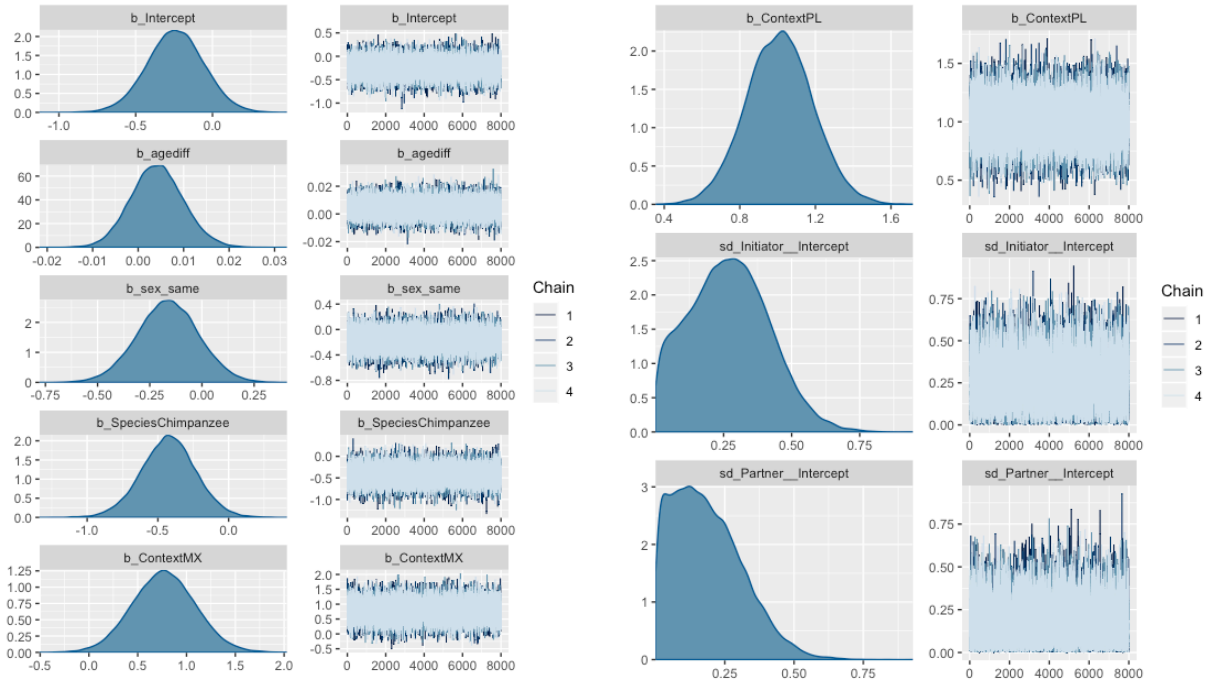
DSI * Species	0.27	0.07	[0.13;0.42]	22660	1.00	Student t (3,2,10)
Rank difference * Species	0.14	0.09	[-0.03;0.31]	19955	1.00	Student t (3,2,10)
Random effects						
Initiator	0.09	0.06	[0.00;0.22]	5608	1.00	Student t (3,0,10)
Partner	0.15	0.06	[0.02;0.27]	5020	1.00	Student t (3,0,10)

Note. All datasets used for the mixed models excluded *n*=118 observations of exit phases that were identified as "mutually intitiated" phases, due to the problem of determining an ID of an initiator and partner for random effects.

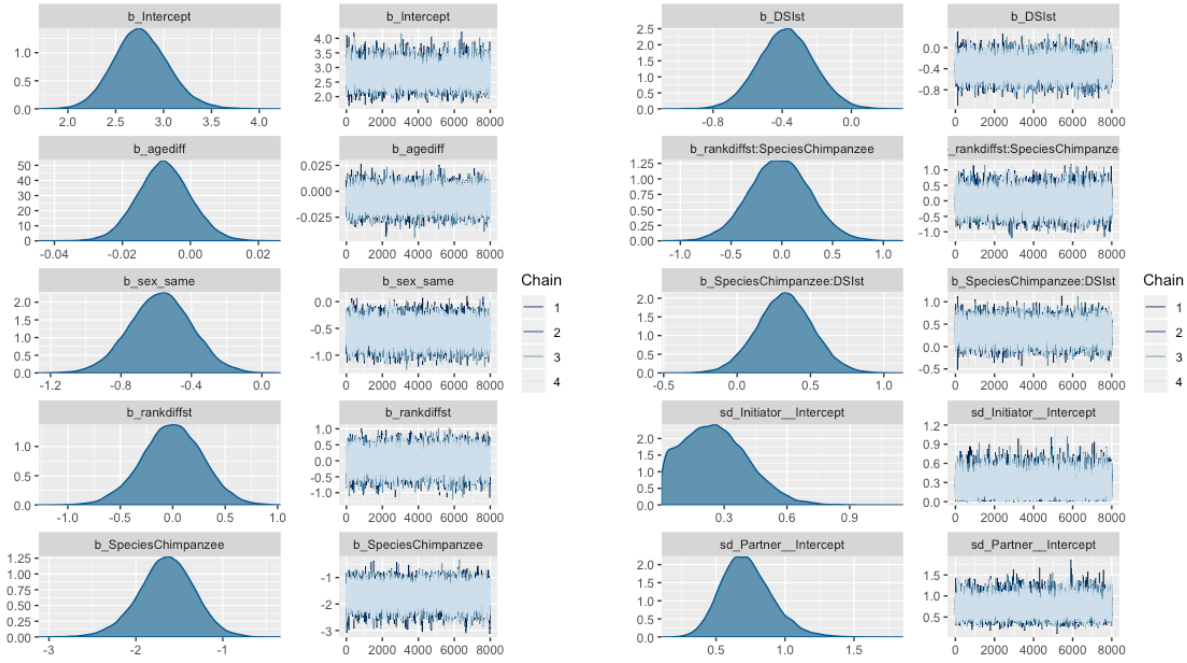
[A]Bürkner, P.C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. https://doi.org/10.18637/jss.v080.i01.

S2.3 Posterior Distributions and MCMC Trace-plots of Bayesian GLMMs and LMMs including only intercepts

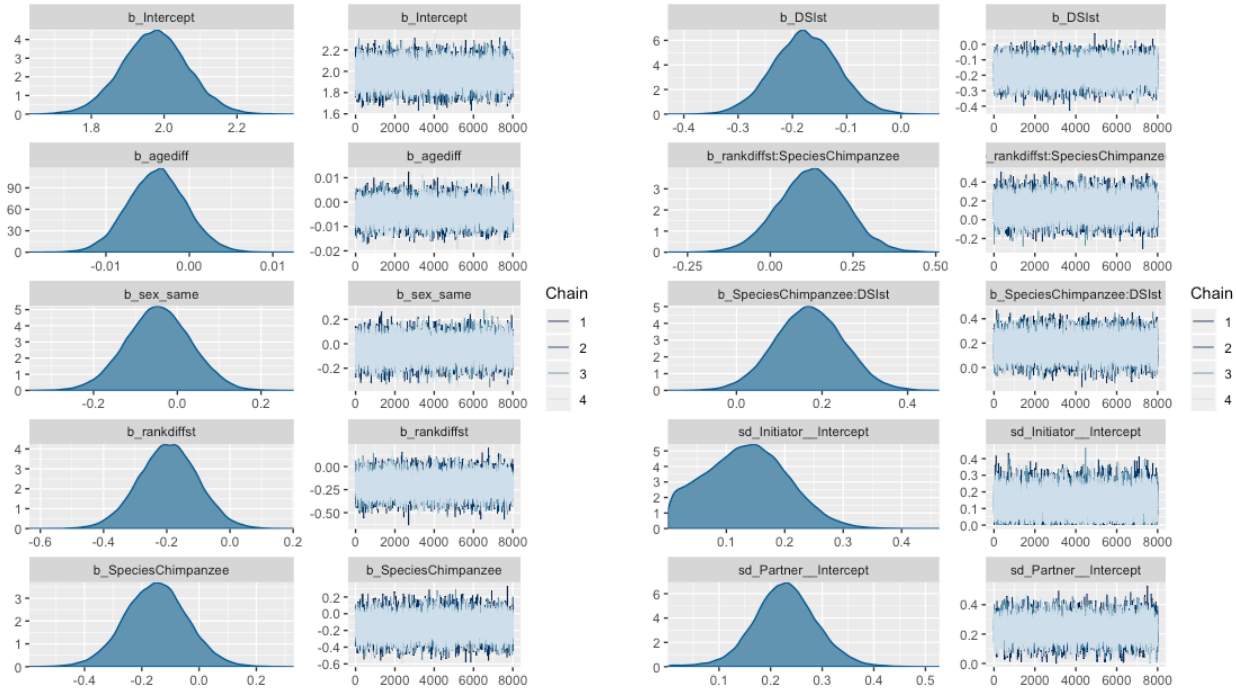
Model 1

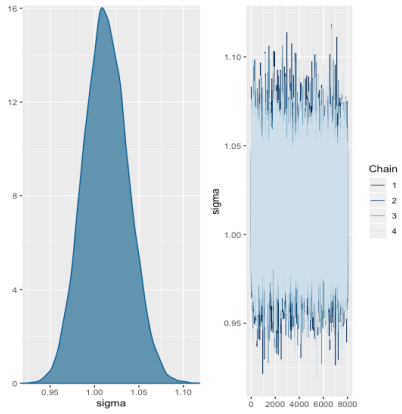


Model 2

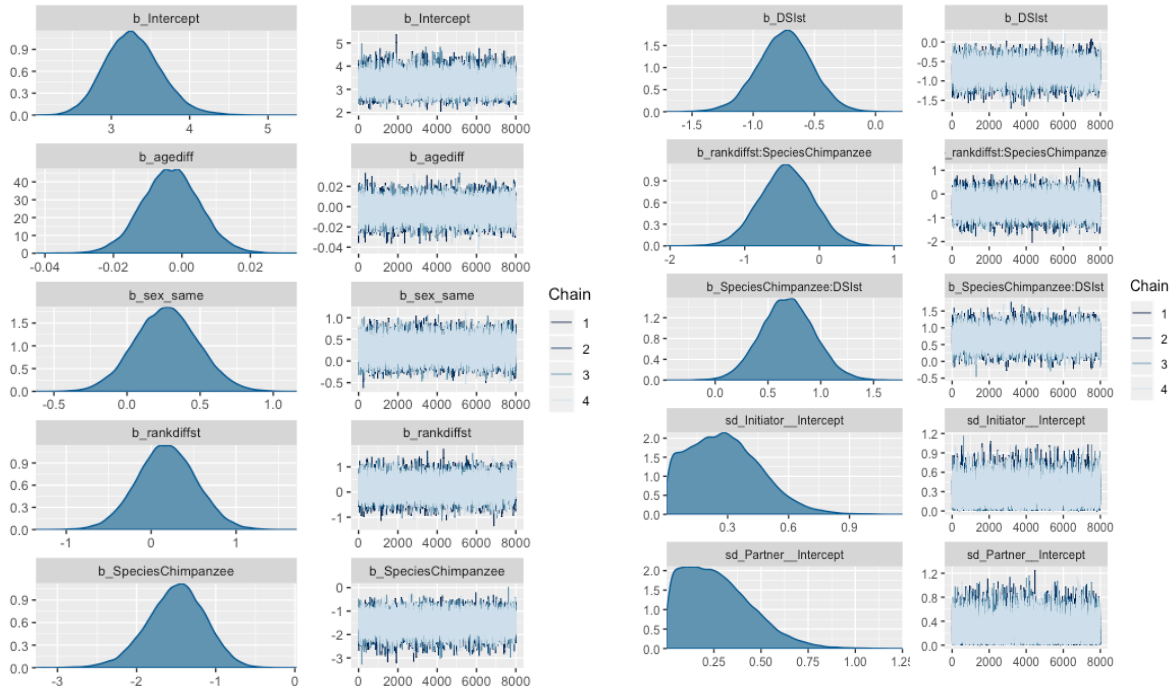


Model 3

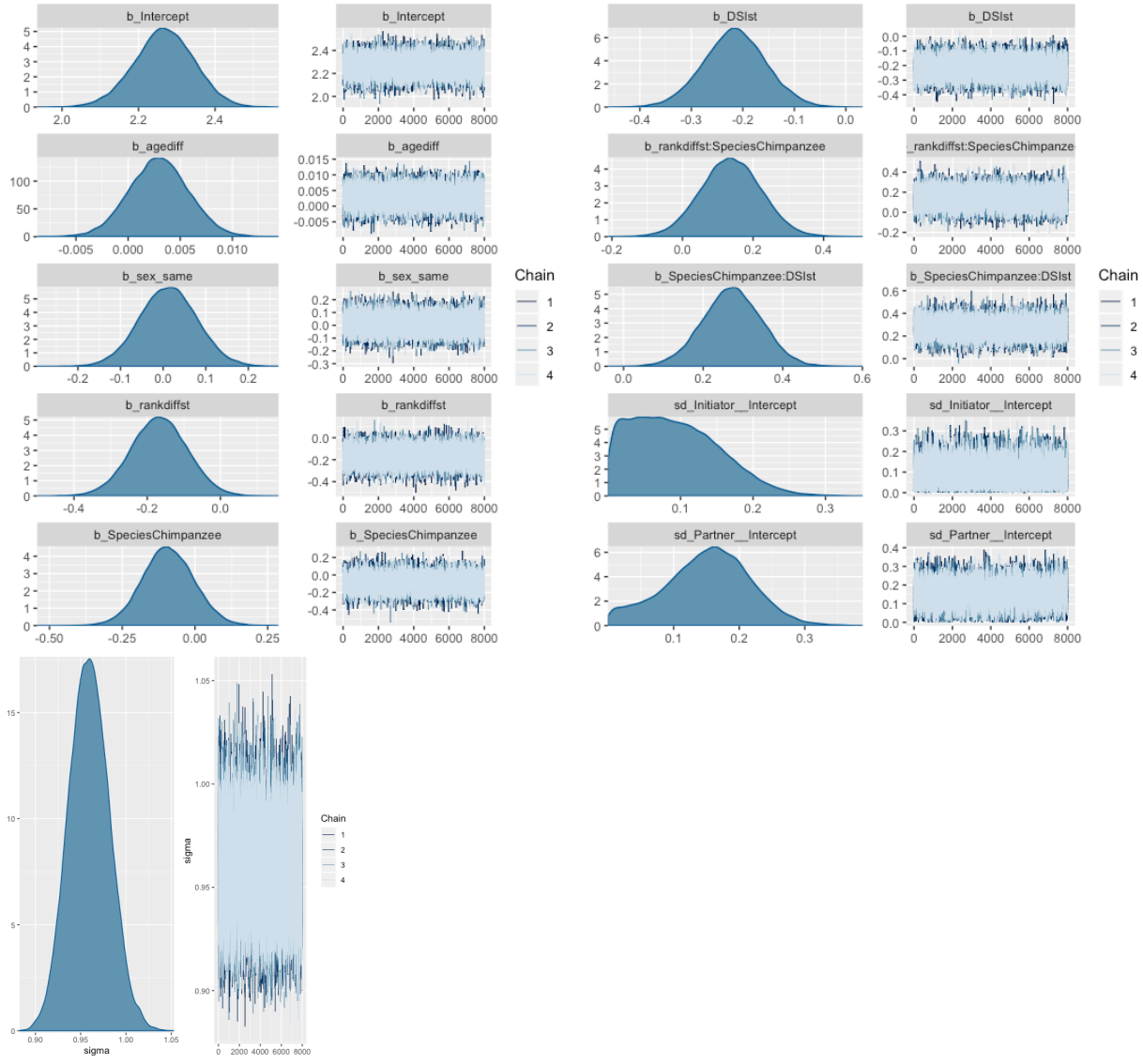



sigma
sigma
Chain
1
2
3
4

Model 4


b_Intercept
b_agediff
b_sex_same
b_rankdiffst
b_SpeciesChimpanzee
b_DSIst
b_rankdiffst:SpeciesChimpanzee
b_SpeciesChimpanzee:DSIst
sd_Initiator__Intercept
sd_Partner__Intercept
Chain
1
2
3
4

Model 5


b_Intercept
b_DSIst
b_agediff
b_rankdiffst:SpeciesChimpanzee
b_sex_same
b_SpeciesChimpanzee:DSIst
b_rankdiffst
sd_Initiator__Intercept
b_SpeciesChimpanzee
sd_Partner__Intercept
Chain
1
2
3
4
sigma

S2.4 Leave-one-out cross-validation comparing intercept-only vs. random slope models. When the difference in the expected log predictive density (ELPD) is positive, the expected predictive accuracy for the second model is higher. A negative ELPD difference favors the reduced, fist model. Only ELPD differences twice as large as SE difference can be considered true differences in accuracy.

Model	*LOO IC*	*SE*	*ELPD difference*	*SE difference*	*Pareto k diagnostic*	*Divergent chains?*
"Model 2": Dependent variable: Entry phase presence						
Intercept only model: initiator/partner intercepts only	1225.6	36.6	3.8	4.2	Good (<0.5)	No
Random slope model: DSI and Rank slopes for partner and initiator	1217.9	37.1			OK (<0.7)	No
"Model 3": Dependent variable: Entry phase duration						
Intercept only model: initiator/partner intercepts only	6114.6	84.7	5.2	3.9	Good (<0.5)	No
Random slope model: DSI and Rank slopes for partner and initiator	6104.3	84.8			Good (<0.5)	No
"Model 4": Dependent variable: Exit phase presence						
Intercept only model: initiator/partner intercepts only	766.9	43.8	1.5	3.9	Good (<0.5)	No
Random slope model: DSI and Rank slopes for partner and initiator	764.0	44.2			OK (<0.7)	No
"Model 5": Dependent variable: Exit phase duration						
Intercept only model: initiator/partner intercepts only	7217.6	76.1	4.6	3.2	Good (<0.5)	No
Random slope model: DSI and Rank slopes for partner and initiator	7208.5	76.0			OK (<0.7)	No

S2.5 Summary of Results for Bayesian GLMMs and LMMs including Random Slopes

Model 2 (RQ3): Entry phase presence (dependent variable, DV, *n*=1215)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	2.74	0.29	[2.21;3.34]	31901	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-1.85	0.31	[-2.50;-1.27]	29561	1.00	lkj_corr_cholesky(1)
Age difference in years	-0.01	0.01	[-0.02;0.01]	35215	1.00	lkj_corr_cholesky(1)
Sex						
Different	*Reference*					
Same	-0.39	0.20	[-0.78;-0.01]	41689	1.00	lkj_corr_cholesky(1)
DSI	-0.23	0.23	[-0.66;0.22]	21188	1.00	lkj_corr_cholesky(1)
Rank difference	0.04	0.34	[-0.65;0.70]	25335	1.00	lkj_corr_cholesky(1)
DSI * Species	0.25	0.27	[-0.31;0.77]	23474	1.00	lkj_corr_cholesky(1)
Rank difference * Species	0.04	0.40	[-0.72;0.85]	24064	1.00	lkj_corr_cholesky(1)
Random effects						
Initiator						
sd(intercept)	0.26	0.17	[0.01;0.65]	7308	1.00	Student t (3,0,10)
sd(rank difference)	0.21	0.16	[0.01;0.58]	13244	1.00	Student t (3,0,10)
sd(DSI)	0.30	0.17	[0.02;0.67]	11603	1.00	Student t (3,0,10)
cor(intercept, rank difference)	-0.20	0.49	[-0.93;0.80]	20565	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	0.20	0.47	[-0.78;0.92]	14353	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.00	0.49	[-0.87;0.87]	16923	1.00	lkj_corr_cholesky(1)
Partner						
sd(intercept)	0.52	0.23	[0.07;0.98]	5372	1.00	Student t (3,0,10)
sd(rank difference)	0.70	0.25	[0.21;1.22]	9339	1.00	Student t (3,0,10)
sd(DSI)	0.25	0.17	[0.01;0.66]	10171	1.00	Student t (3,0,10)
cor(intercept, rank difference)	0.46	0.35	[-0.39;0.94]	6201	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	-0.28	0.46	[-0.93;0.74]	16971	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.30	0.44	[-0.93;0.69]	19997	1.00	lkj_corr_cholesky(1)

Model 3 (RQ3): Entry phase duration (DV, *n*=921)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	1.91	0.09	[1.74;2.09]	37544	1.00	Student t (3,2,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.12	0.11	[-0.33;0.09]	35279	1.00	lkj_corr_cholesky(1)
Age difference in years	-0.00	0.00	[-0.01;0.00]	3262	1.00	lkj_corr_cholesky(1)
Sex						
Different	*Reference*					
Same	0.00	0.08	[-0.16;0.17]	20113	1.00	lkj_corr_cholesky(1)
DSI	-0.17	0.07	[-0.31;-0.03]	30992	1.00	lkj_corr_cholesky(1)
Rank difference	-0.17	0.12	[-0.41;0.06]	22096	1.00	lkj_corr_cholesky(1)
DSI * Species	0.17	0.10	[-0.02;0.36]	30382	1.00	lkj_corr_cholesky(1)
Rank difference * Species	0.09	0.15	[-0.21;0.38]	24113	1.00	lkj_corr_cholesky(1)
Random effects						
Initiator						
sd(intercept)	0.08	0.06	[0.00;0.22]	9990	1.00	Student t (3,0,10)
sd(rank difference)	0.28	0.09	[0.12;0.46]	8995	1.00	Student t (3,0,10)
sd(DSI)	0.09	0.06	[0.00;0.23]	10946	1.00	Student t (3,0,10)
cor(intercept, rank difference)	-0.29	0.44	[-0.93;0.72]	4160	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	0.03	0.49	[-0.86;0.88]	19216	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.08	0.45	[-0.87;0.79]	26116	1.00	lkj_corr_cholesky(1)
Partner						
sd(intercept)	0.17	0.08	[0.01;0.32]	6950	1.00	Student t (3,0,10)

sd(rank difference)	0.11	0.07	[0.01;0.25]	10784	1.00	Student t (3,0,10)
sd(DSI)	0.08	0.06	[0.00;0.21]	13191	1.00	Student t (3,0,10)
cor(intercept, rank difference)	-0.21	0.46	[-0.92;0.75]	18094	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	-0.04	0.47	[-0.87;0.84]	27993	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	0.12	0.49	[-0.84;0.91]	24158	1.00	lkj_corr_cholesky(1)

Model 4 (RQ3): Exit phase presence (DV, *n*=1129)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	3.48	0.38	[2.78;4.30]	34628	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-1.55	0.39	[-2.36;-0.81]	41135	1.00	lkj_corr_cholesky(1)
Age difference in years	-0.00	0.01	[-0.02;0.02]	41170	1.00	lkj_corr_cholesky(1)
Sex						
Different	*Reference*					
Same	0.22	0.24	[-0.26;0.69]	56032	1.00	lkj_corr_cholesky(1)
DSI	-0.71	0.32	[-1.34;-0.07]	28996	1.00	lkj_corr_cholesky(1)
Rank difference	0.04	0.41	[-0.77;0.83]	28061	1.00	lkj_corr_cholesky(1)
DSI * Species	0.87	0.39	[0.12;1.67]	27899	1.00	lkj_corr_cholesky(1)
Rank difference * Species	-0.26	0.44	[-1.11;0.63]	26776	1.00	lkj_corr_cholesky(1)
Random effects						
Initiator						
sd(intercept)	0.27	0.19	[0.01;0.69]	11387	1.00	Student t (3,0,10)
sd(rank difference)	0.37	0.23	[0.02;0.88]	9757	1.00	Student t (3,0,10)
sd(DSI)	0.49	0.23	[0.06;0.96]	10870	1.00	Student t (3,0,10)
cor(intercept, rank difference)	0.24	0.47	[-0.76;0.94]	12608	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	-0.08	0.47	[-0.88;0.81]	10024	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.35	0.45	[-0.95;0.69]	10095	1.00	lkj_corr_cholesky(1)
Partner						
sd(intercept)	0.21	0.16	[0.01;0.59]	16264	1.00	Student t (3,0,10)
sd(rank difference)	0.18	0.14	[0.01;0.53]	20026	1.00	Student t (3,0,10)
sd(DSI)	0.41	0.24	[0.02;0.93]	8350	1.00	Student t (3,0,10)
cor(intercept, rank difference)	-0.03	0.50	[-0.89;0.87]	39875	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	0.09	0.49	[-0.84;0.90]	15120	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.04	0.50	[-0.89;0.86]	14986	1.00	lkj_corr_cholesky(1)

Model 5 (RQ3): Exit phase duration (DV, *n*=1002)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***[A]
Intercept	2.26	0.08	[2.10;2.42]	26187	1.00	Student t (3,2,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.09	0.09	[-0.27;0.10]	23089	1.00	lkj_corr_cholesky(1)
Age difference in years	0.00	0.00	[-0.00;0.01]	23569	1.00	lkj_corr_cholesky(1)
Sex						
Different	*Reference*					
Same	0.00	0.07	[-0.14;0.14]	34499	1.00	lkj_corr_cholesky(1)
DSI	-0.22	0.08	[-0.39;-0.06]	18948	1.00	lkj_corr_cholesky(1)
Rank difference	-0.17	0.09	[-0.35;0.01]	17983	1.00	lkj_corr_cholesky(1)
DSI * Species	0.28	0.11	[0.07;0.50]	18088	1.00	lkj_corr_cholesky(1)
Rank difference * Species	0.15	0.11	[-0.06;0.37]	16719	1.00	lkj_corr_cholesky(1)
Random effects						
Initiator						
sd(intercept)	0.09	0.06	[0.00;0.22]	7871	1.00	Student t (3,0,10)
sd(rank difference)	0.07	0.05	[0.00;0.20]	11223	1.00	Student t (3,0,10)
sd(DSI)	0.17	0.08	[0.02;0.33]	5741	1.00	Student t (3,0,10)

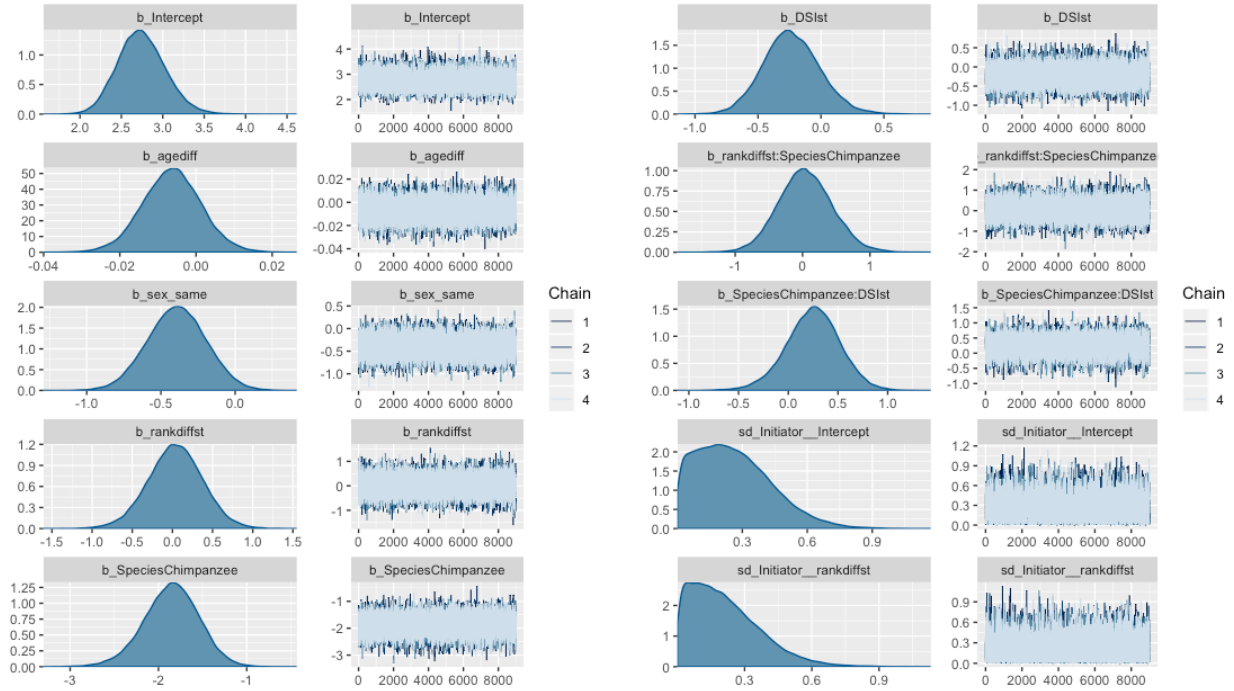
cor(intercept, rank difference)	0.06	0.50	[-0.86;0.89]	17225	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	-0.02	0.47	[-0.86;0.84]	6479	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	-0.20	0.48	[-0.92;0.79]	7024	1.00	lkj_corr_cholesky(1)
Partner						
sd(intercept)	0.13	0.07	[0.01;0.27]	4999	1.00	Student t (3,0,10)
sd(rank difference)	0.12	0.08	[0.01;0.29]	6619	1.00	Student t (3,0,10)
sd(DSI)	0.11	0.07	[0.01;0.25]	7668	1.00	Student t (3,0,10)
cor(intercept, rank difference)	-0.11	0.46	[-0.88;0.81]	10586	1.00	lkj_corr_cholesky(1)
cor(intercept, DSI)	0.33	0.45	[-0.71;0.94]	8065	1.00	lkj_corr_cholesky(1)
cor(rank difference, DSI)	0.11	0.49	[-0.83;0.90]	11673	1.00	lkj_corr_cholesky(1)

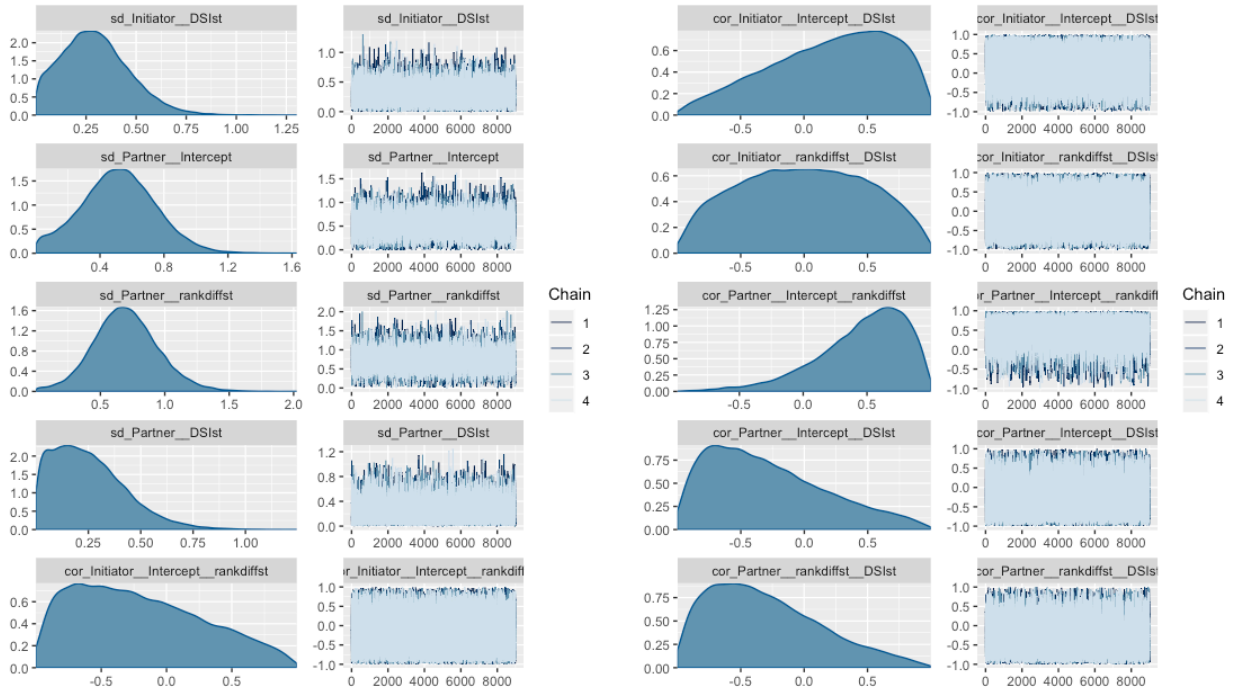
Note. All datasets used for the mixed models excluded observations of exit phases that were identified as "mutually intitiated" phases, due to the problem of determining an ID of an initiator and partner for random effects.

[A]Bürkner, P.C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. https://doi.org/10.18637/jss.v080.i01.

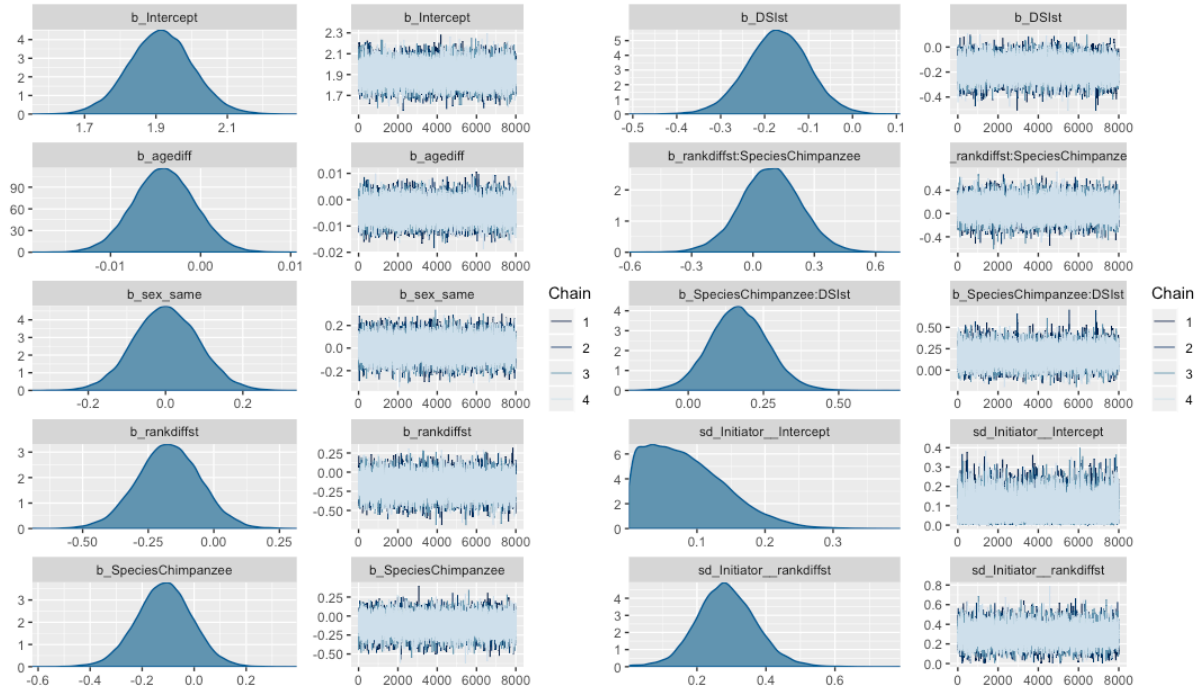
S2.6 Posterior Distributions and MCMC Trace-plots of Bayesian GLMMs and LMMs including random slopes

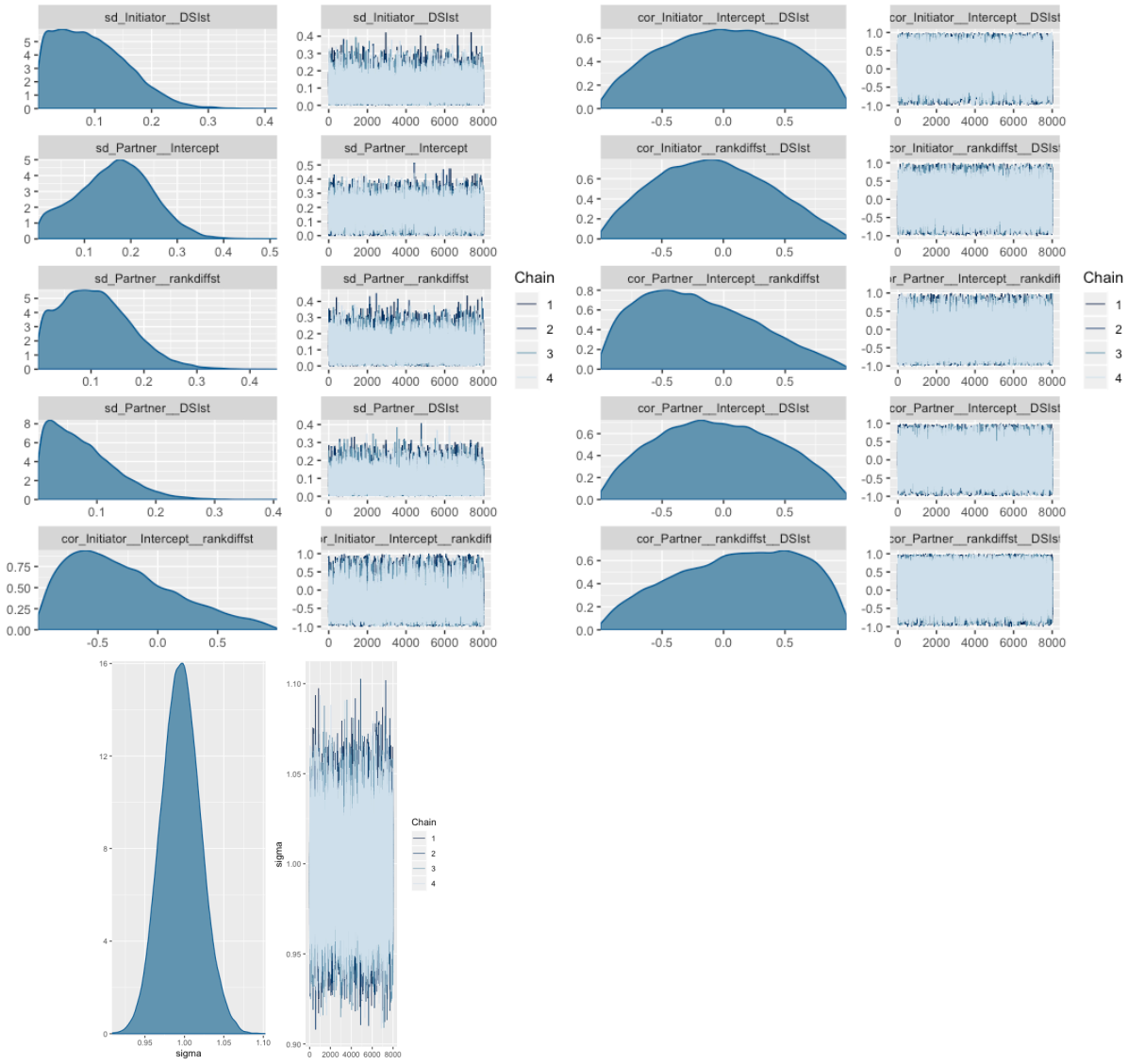
Model 2



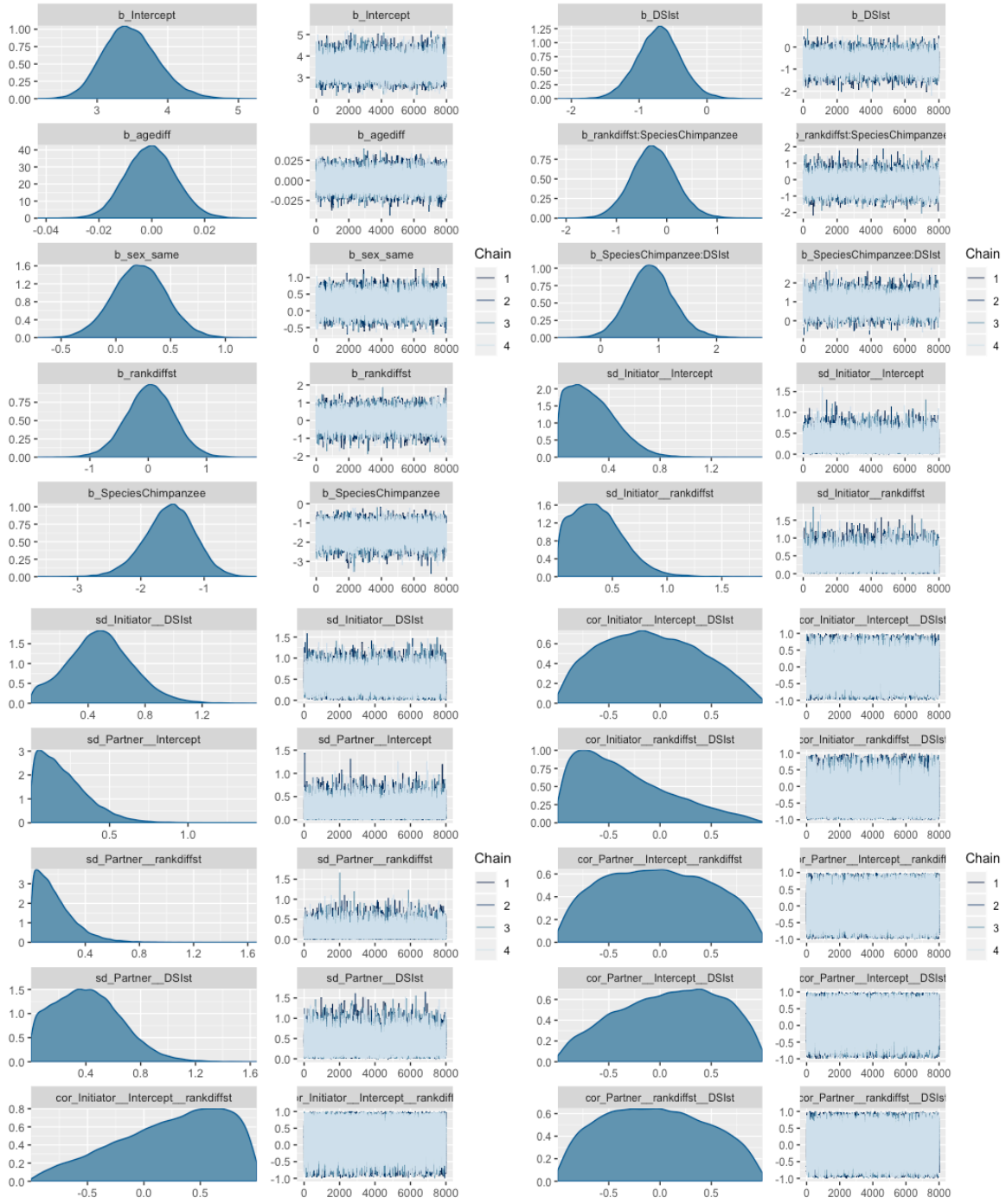

sd_Initiator__DSIst
sd_Partner__Intercept
sd_Partner__rankdiffst
sd_Partner__DSIst
cor_Initiator__Intercept__rankdiffst
cor_Initiator__Intercept__DSIst
cor_Initiator__rankdiffst__DSIst
cor_Partner__Intercept__rankdiffst
cor_Partner__Intercept__DSIst
cor_Partner__rankdiffst__DSIst
Chain
1
2
3
4

Model 3

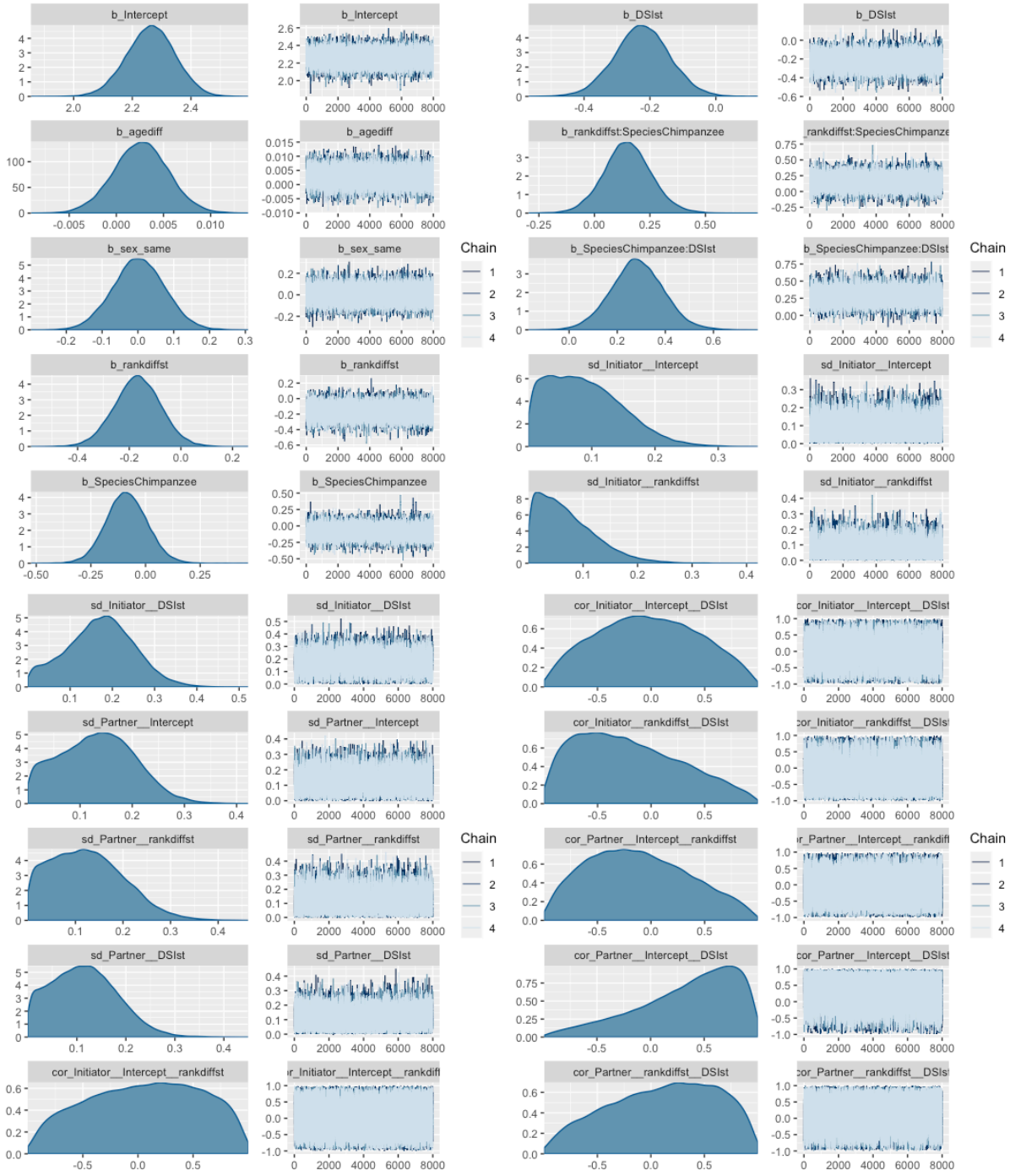

b_Intercept
b_agediff
b_sex_same
b_rankdiffst
b_SpeciesChimpanzee
b_DSIst
b_rankdiffst:SpeciesChimpanzee
b_SpeciesChimpanzee:DSIst
sd_Initiator__Intercept
sd_Initiator__rankdiffst
Chain
1
2
3
4

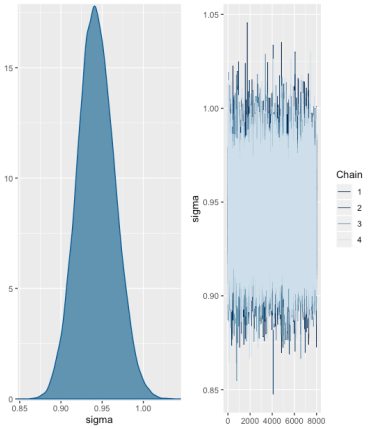

sd_Initiator__DSIst
sd_Partner__Intercept
sd_Partner__rankdiffst
sd_Partner__DSIst
cor_Initiator__Intercept__rankdiffst
cor_Initiator__Intercept__DSIst
cor_Initiator__rankdiffst__DSIst
cor_Partner__Intercept__rankdiffst
cor_Partner__Intercept__DSIst
cor_Partner__rankdiffst__DSIst
Chain
1
2
3
4
sigma

Model 4


b_Intercept
b_DSIst
b_agediff
b_rankdiffst:SpeciesChimpanzee
b_sex_same
b_SpeciesChimpanzee:DSIst
b_rankdiffst
sd_Initiator__Intercept
b_SpeciesChimpanzee
sd_Initiator__rankdiffst
sd_Initiator__DSIst
cor_Initiator__Intercept__DSIst
sd_Partner__Intercept
cor_Initiator__rankdiffst__DSIst
sd_Partner__rankdiffst
cor_Partner__Intercept__rankdiffst
sd_Partner__DSIst
cor_Partner__Intercept__DSIst
cor_Initiator__Intercept__rankdiffst
cor_Partner__rankdiffst__DSIst
Chain
1
2
3
4

Model 5


b_Intercept
b_DSIst
b_agediff
b_rankdiffst:SpeciesChimpanzee
b_sex_same
b_SpeciesChimpanzee:DSIst
b_rankdiffst
sd_Initiator__Intercept
b_SpeciesChimpanzee
sd_Initiator__rankdiffst
sd_Initiator__DSIst
cor_Initiator__Intercept__DSIst
sd_Partner__Intercept
cor_Initiator__rankdiffst__DSIst
sd_Partner__rankdiffst
cor_Partner__Intercept__rankdiffst
sd_Partner__DSIst
cor_Partner__Intercept__DSIst
cor_Initiator__Intercept__rankdiffst
cor_Partner__rankdiffst__DSIst
Chain
1
2
3
4


15
10
5
0
0.85
0.90
0.95
1.00
sigma

1.05
1.00
0.95
0.90
0.85
sigma
0
2000
4000
6000
8000
Chain
1
2
3
4

Appendix 3

ESM for Chapter 4

S3.1 Detailed Information about Study Subjects

Subject	Species	Study	Site	Estimated age (years)	Age class (after Kano, 1992)	Sex	N° Trials still-faced condition	N° Trials back-turned condition
BA	Bonobo	1	Lola Ya Bonobo	3.0	Infant	Male	4	3
BI	Bonobo	1	Lola Ya Bonobo	4.0	Infant	Male	2	2
KW	Bonobo	1	Lola Ya Bonobo	2.5	Infant	Male	3	2
LA	Bonobo	1	Lola Ya Bonobo	5.0	Infant	Female	3	1
LU	Bonobo	1	Lola Ya Bonobo	2.5	Infant	Female	1	2
BO	Bonobo	2	Lola Ya Bonobo	9.0	Subadult	Female	1	1
EL	Bonobo	2	Lola Ya Bonobo	6.0	Juvenile	Female	5	3
IS	Bonobo	2	Lola Ya Bonobo	21.0	Adult	Female	1	1
KE	Bonobo	2	Lola Ya Bonobo	25.0	Adult	Male	6	3
KI	Bonobo	2	Lola Ya Bonobo	8.0	Subadult	Female	1	1
LI	Bonobo	2	Lola Ya Bonobo	8.0	Subadult	Female	1	1
LO	Bonobo	2	Lola Ya Bonobo	11.0	Subadult	Male	8	7
LU	Bonobo	2	Lola Ya Bonobo	6.0	Juvenile	Female	1	0
MI	Bonobo	2	Lola Ya Bonobo	8.0	Subadult	Female	4	0
MO	Bonobo	2	Lola Ya Bonobo	11.0	Subadult	Male	3	1
OP	Bonobo	2	Lola Ya Bonobo	23.0	Adult	Female	1	1
SI	Bonobo	2	Lola Ya Bonobo	10.0	Subadult	Male	0	2
TC	Bonobo	2	Lola Ya Bonobo	30.0	Adult	Female	0	1
CO	Chimpanzees	1	La Vallée des Singes	22.0	Adult	Male	6	3
JO	Chimpanzees	1	La Vallée des Singes	23.0	Adult	Male	12	20
LL	Chimpanzees	1	La Vallée des Singes	9.0	Subadult	Female	1	1

RO	Chimpanzees	1	La Vallée des Singes	21.0	Adult	Male	5	4
WO	Chimpanzees	1	La Vallée des Singes	21.0	Adult	Male	3	3

S3.2 Definition of signals and behaviors observed to reengage a partner during interruption periods. Definition taken and adapted from (Byrne et al., 2017; Clay, Archbold, & Zuberbühler, 2015b; Crockford et al., 2018; de Waal, 1988).

Type	Modality	Description
Gestures		
Bang cage	audible	Hand(s) (or foot/feet) brought into hard audible contact with the cage mesh (can be repetitive)
Dangle	visual	Subject hangs from arm(s) above another, from the top of the cage, with feet shaking loosely
Grab experimenter	tactile	Hand(s) firmly closed over part of experimenter's body
Hit object	audible	Hand(s) brought into short hard contact with the surface of an object
Hit object with object	audible	Object brought into short hard contact with another object
Jump	visual	While in bipedal position both feet leave ground simultaneously with horizontal displacement
Knock mesh/object	audible	Back of hand/knuckles brought into short hard audible contact with cage mesh/object (can be repetitive)
Move hose	visual	Hose is displaced in one direction, contact with hose is maintained throughout
Present body	visual	Body part moved to deliberately expose an area to experimenter's attention
Present sex	visual	Genitals moved to deliberately expose them to experimenter's attention
Reach	visual	Arm is extended to the experimenter with hand in open, palm exposed position
Shake experimenter	tactile	Hand(s) is firmly closed over part of experimenter's body with repeated back and forth motion of hand
Shake head	visual	Head moved back and forth in small repeated side motion
Shake hose	visual	Hand(s) firmly closed over hose with repeated up and down or side movements
Shake hand	visual	Small repeated back and forth motion of hand(s) from wrist
Stomp	audible	Sole of the foot/feet brought into short hard audible contact with surface
Touch experimenter	tactile	Palm of hand and/or body is brought to light contact over part of experimenter's body, for under 2 s
Touch hose	visual	Palm of hand and/or body is brought to light contact over hose, for under 2 s
Behaviors		
Drop hose out of cage	visual	Hose pushed through the mesh outside of cage and out of reach
Hand back hose	visual	One end of hose pushed through the cage mesh outside of cage, while hand is firmly closed over other end of hose
Prompt	visual	Body/hand is moved in rapid, short, and tense movements, while hand is firmly closed over hose and while looking at experimenter
Simulate game action	visual	Hose pushed back and forth through the mesh (mimicking the game's pull and relax movements)
Touch with hose	tactile	Hose pushed through the cage mesh outside of cage and brought to light contact over part of experimenter's body

Vocalizations		
Laughter	audible	Voiceless breathing sounding like low-pitched grunting
Hoo	audible	Quiet, relatively inconspicuous vocalization sounding like «*hoo*», mostly emitted as a single call
Peep	audible	High-frequency, closed mouth vocalization; short in duration and flat, simple acoustic structure
Facial expressions		
Pout face	visual	Lips pursed forward and curled outward in front (circular opening); lips pressed together at mouth corners
Tightened lips	visual	Lips horizontally tensed; drawn inward or slightly baring teeth
Playface	visual	Mouth opened with lips either in relaxed position covering the upper teeth completely and the lower teeth partially, or retracted, showing both upper and lower frontal teeth

S3.3 Modality adjustment for Experiment 1-3

Results

To what extent do adult chimpanzees adapt the signal or behavioral modality to the condition (i.e., attentive state of the partner)?

In the still-faced condition, the chimpanzees most often produced visual behaviors/signals (mean proportion 91.03%), and second audible behaviors/signals (7.69%); the subjects did not deploy tactile behaviors/signals in this condition (fig. S3.3.1). In the back-turned condition, the chimpanzees produced less visual behaviors/signals as compared with the still-faced condition (68.65%), but produced more tactile (15.41%) and audible (15.14%) behaviors/signals (fig. S4.1).

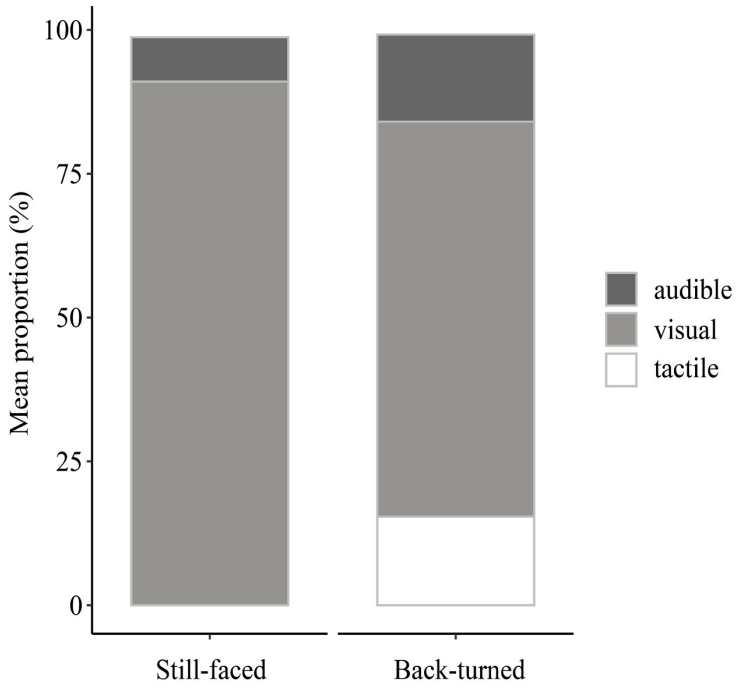


Figure S3.3.1 Mean percentage proportion of behavioral/signal perception modalities (audible, tactile and visual) for condition 1 (still-faced) and condition 2 (back-turned) in adult chimpanzees.

To what extent do adult bonobos adapt the signal or behavioral modality to the condition (i.e., attentive state of the partner)?

In both conditions almost equally, the bonobos most often produced visual behaviors/signals (mean proportion 94.46% in the still-faced and 99.18% in the back-turned condition), see fig. S3.3.2; they produced a very small amount of audible behaviors/signals in the two conditions (3.73% in the still-faced and 0.82% in the back-turned condition), and very few tactile behaviors/signals in the still-faced condition (1.81%).

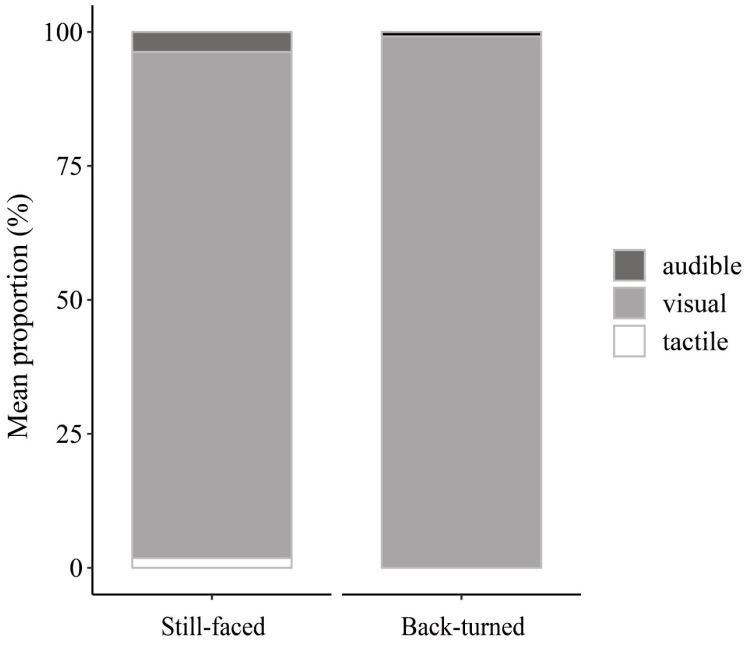


Figure S3.3.2 Mean percentage proportion of behavioral/signal perception modalities (audible, tactile and visual) for condition 1 (still-faced) and condition 2 (back-turned) in adult bonobos.

To what extent do infant bonobos adapt the signal or behavioral modality to the condition (i.e., attentive state of the partner)?

In the still-faced condition, the infant bonobos most often produced visual behaviors/signals (mean proportion 77.28%), second tactile behaviors/signals (19.58%), and third audible behaviors/signals (3.13%), see fig. S3.3.3. In the back-turned condition, the infants produced less visual behaviors/signals as compared with the still-faced condition (50%), but produced more tactile (45.45%) and audible (4.55%) behaviors/signals (fig. S4.3).

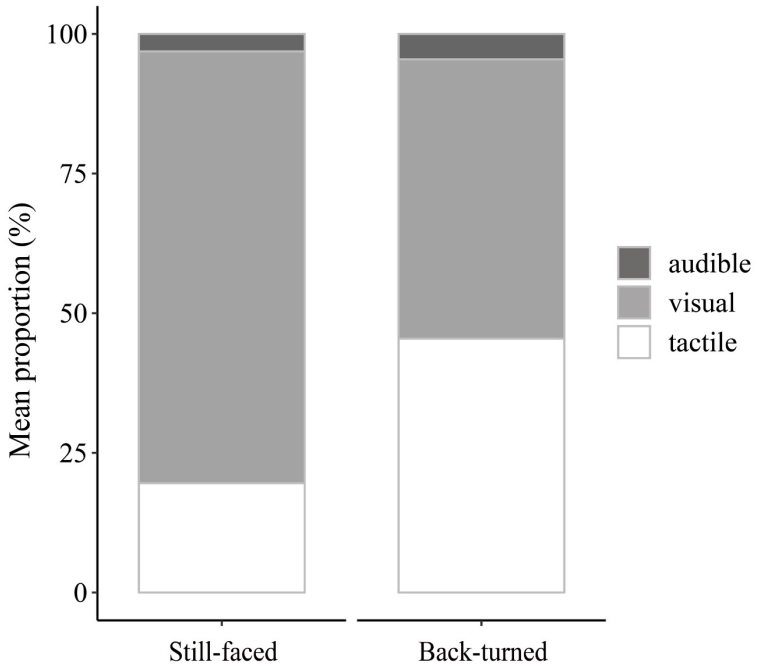


Figure S3.3.3 Mean percentage proportion of behavioral/signal perception modalities (audible, tactile and visual) for condition 1 (still-faced) and condition 2 (back-turned) in infant bonobos.

Brief discussion

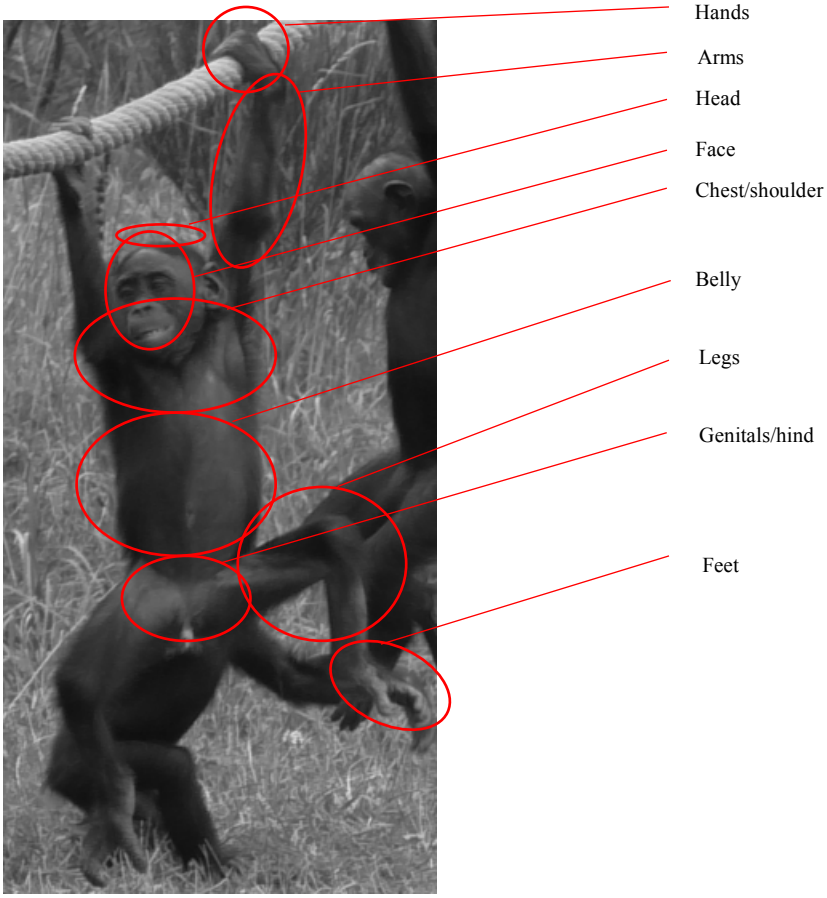
The chimpanzees and the infant bonobos adjusted the modality of their behaviors/signals to the attentional state of the human partner by using less visual and more tactile/audible behaviors and signals when the partner was facing away, relative to when the partner was facing the subject – a result congruent with previous research on this matter (Leavens, Hostetter, Wesley, & Hopkins, 2004; Theall & Povinelli, 1999, Genty et al., 2009; Hobaiter and Byrne 2011). Adult bonobos did not seem to change the use of the behavioral/signal modality like chimpanzees and bonobo infants, according to the condition. Independent of the condition, they seemed to rely mainly on visual behaviors/signals – as opposed to audible or tactile behaviors/signals. One could argue that this difference may be explained by the difficulty of identifying the modality of the behaviors and signals in this study. For instance, "hand back hose" and "simulate game action", or any other behavior or gesture that involves the garden hose, made a slight noise but was potentially not meant to be audible but rather visual. Since the material of the mesh that constructed the cage of the bonobos and chimpanzees was not the same, the sounds produced when handing back the hose or simulating the game action differed depending on the site, making it again difficult to identify by

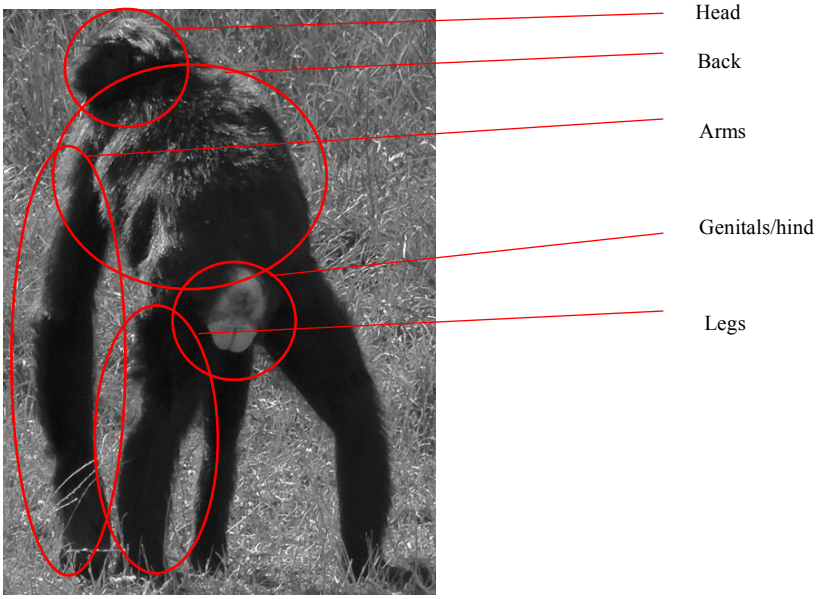
which modality a behavior/signal was intentionally meant to be expressed by the subject. Questions related to the use of the sensory modality type shall therefore be repeated in a paradigm that eliminates the influence of cage mesh on the modality type used within the signal produced.

Appendix 4

ESM for Chapter 5

S4.1 Description of body parts to determine whether the groomed body part is identical or different following interruptions. First photo shows front and second photo shows back of the individual. Photo credit: Raphaela Heesen.





S4.2 Summary of Bayesian generalized mixed models

Model 1 (Dependent variable: Reengagement, *n*=1013)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.88	0.22	[1.46;2.32]	17144	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.85	0.28	[-1.41;-0.30]	22107	1.00	Student t (3,0,10)
Severity of perturbation	-1.65	0.25	[-2.15;-1.19]	18731	1.00	Student t (3,0,10)
Activity type						
Grooming	*Reference*					
Play	-0.03	0.29	[-0.59;0.54]	22024	1.00	Student t (3,0,10)
Activity type * Species	0.31	0.48	[-0.66;1.26]	21381	1.00	Student t (3,0,10)
DSI	0.15	0.16	[-0.15;0.46]	24113	1.00	Student t (3,0,10)
Rank difference	0.13	0.18	[-0.23;0.50]	21909	1.00	Student t (3,0,10)
DSI * Species	-0.13	0.25	[-0.63;0.35]	21638	1.00	Student t (3,0,10)
Rank difference * Species	-0.33	0.26	[-0.85;0.17]	20930	1.00	Student t (3,0,10)
Random effects						
Main body ID, *n*=514	0.79	0.28	[0.17;1.31]	2889	1.00	Student t (3,0,10)
Dyad ID, *n*=144	0.57	0.20	[0.11;0.93]	4675	1.00	Student t (3,0,10)

Model 2 (Dependent variable : Continuation of same action, *n*=729)

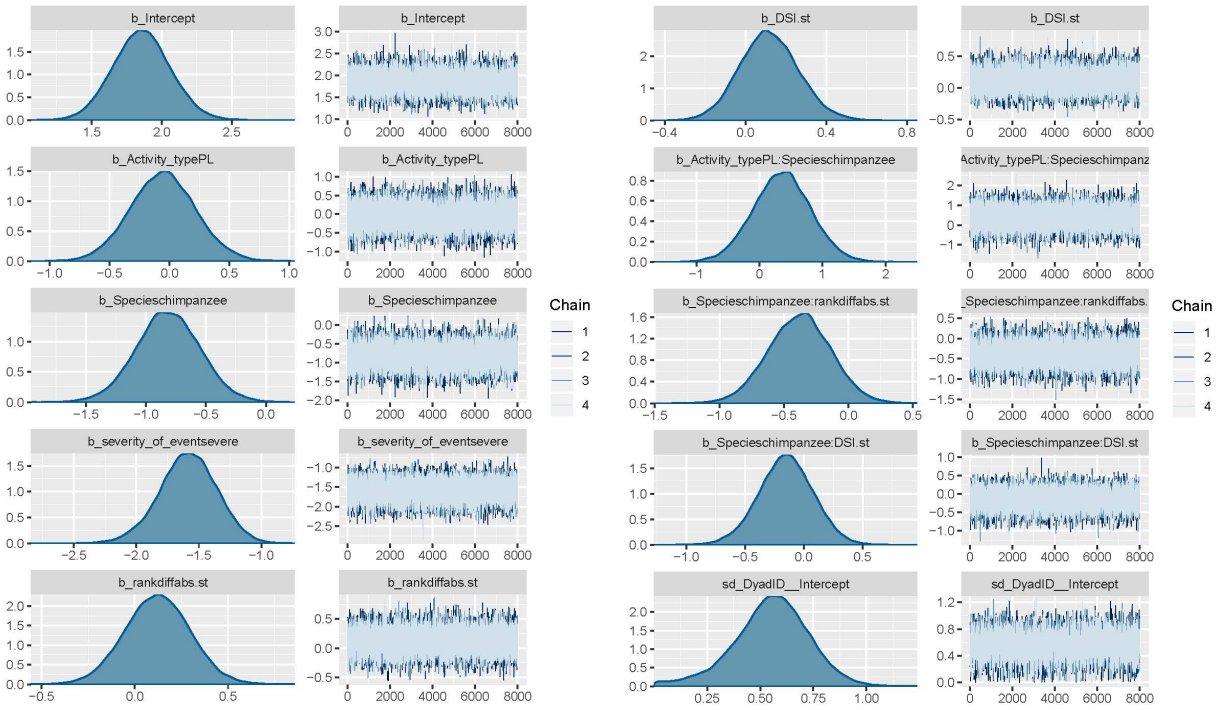
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.66	0.27	[1.14;2.22]	14588	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-1.44	0.38	[-2.19;-0.70]	14151	1.00	Student t (3,0,10)
Severity of perturbation	-1.04	0.35	[-1.73;-0.37]	23163	1.00	Student t (3,0,10)
Activity type						
Grooming	*Reference*					
Play	0.97	0.40	[0.22;1.78]	15698	1.00	Student t (3,0,10)
Activity type * Species	1.28	0.73	[-0.11;2.78]	16311	1.00	Student t (3,0,10)
DSI	0.19	0.21	[-0.22;0.62]	16904	1.00	Student t (3,0,10)
Rank difference	-0.04	0.24	[-0.52;0.43]	15326	1.00	Student t (3,0,10)
DSI * Species	-0.12	0.34	[-0.80;0.55]	15614	1.00	Student t (3,0,10)
Rank difference * Species	0.41	0.36	[-0.28;1.11]	14415	1.00	Student t (3,0,10)

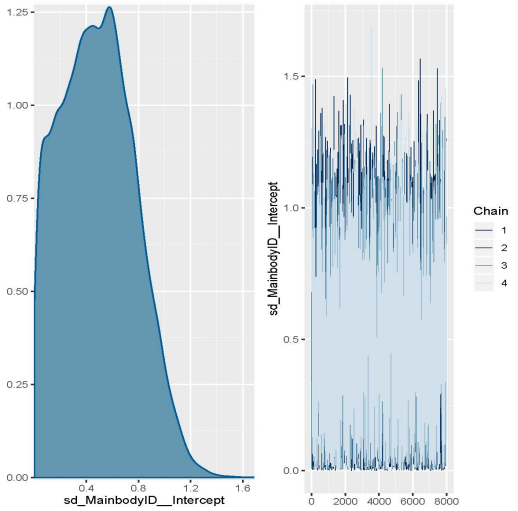
Random effects						
Main body ID, *n*=364	0.97	0.35	[0.23;1.68]	2792	1.00	Student t (3,0,10)
Dyad ID, *n*=133	0.75	0.28	[0.14;1.27]	4133	1.00	Student t (3,0,10)

Model 3 (Dependent variable : Continuation at same location, *n*=109)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.77	0.59	[0.68;3.00]	18362	1.00	Student t (3,0,10)
Species						
Bonobo	*Reference*					
Chimpanzee	-0.31	0.99	[-2.21;1.71]	16701	1.00	Student t (3,0,10)
Severity of perturbation	-1.32	0.64	[-2.60;-0.10]	23332	1.00	Student t (3,0,10)
Activity type						
Grooming	*Reference*					
Play	-0.72	0.62	[-1.96;0.45]	20676	1.00	Student t (3,0,10)
Activity type * Species	-0.79	1.25	[-3.30;1.59]	17238	1.00	Student t (3,0,10)
DSI	0.23	0.28	[-0.30;0.80]	24769	1.00	Student t (3,0,10)
Rank difference	0.37	0.38	[-0.35;1.13]	22443	1.00	Student t (3,0,10)
DSI * Species	-0.95	0.65	[-2.32;0.26]	21797	1.00	Student t (3,0,10)
Rank difference * Species	-0.49	0.57	[-1.63;0.62]	20090	1.00	Student t (3,0,10)

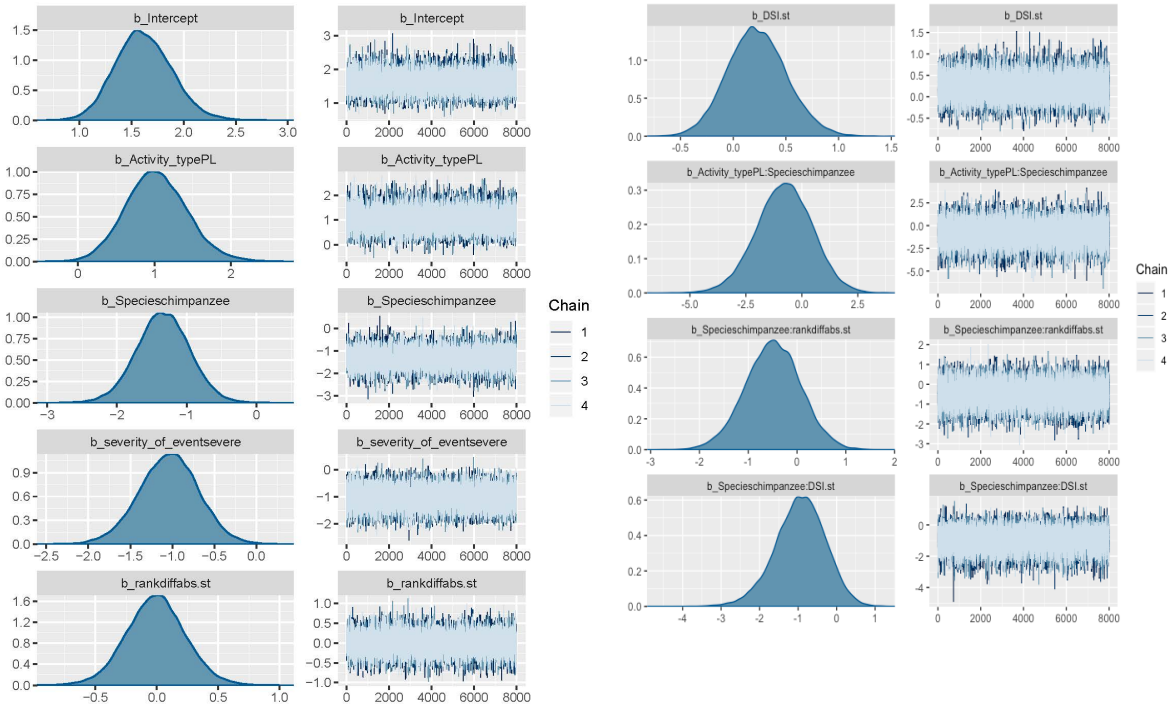
S4.3 Posterior distributions and MCMC trace-plots of Bayesian generalized linear mixed models

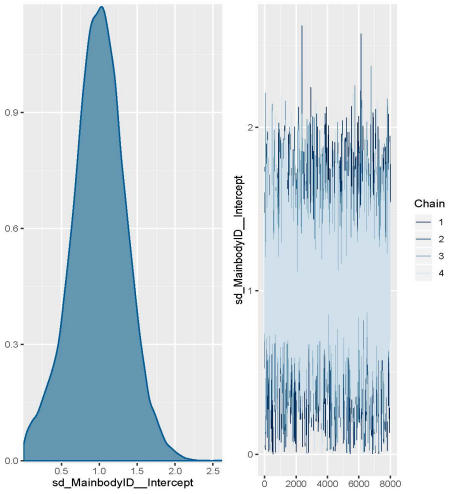
Model 1



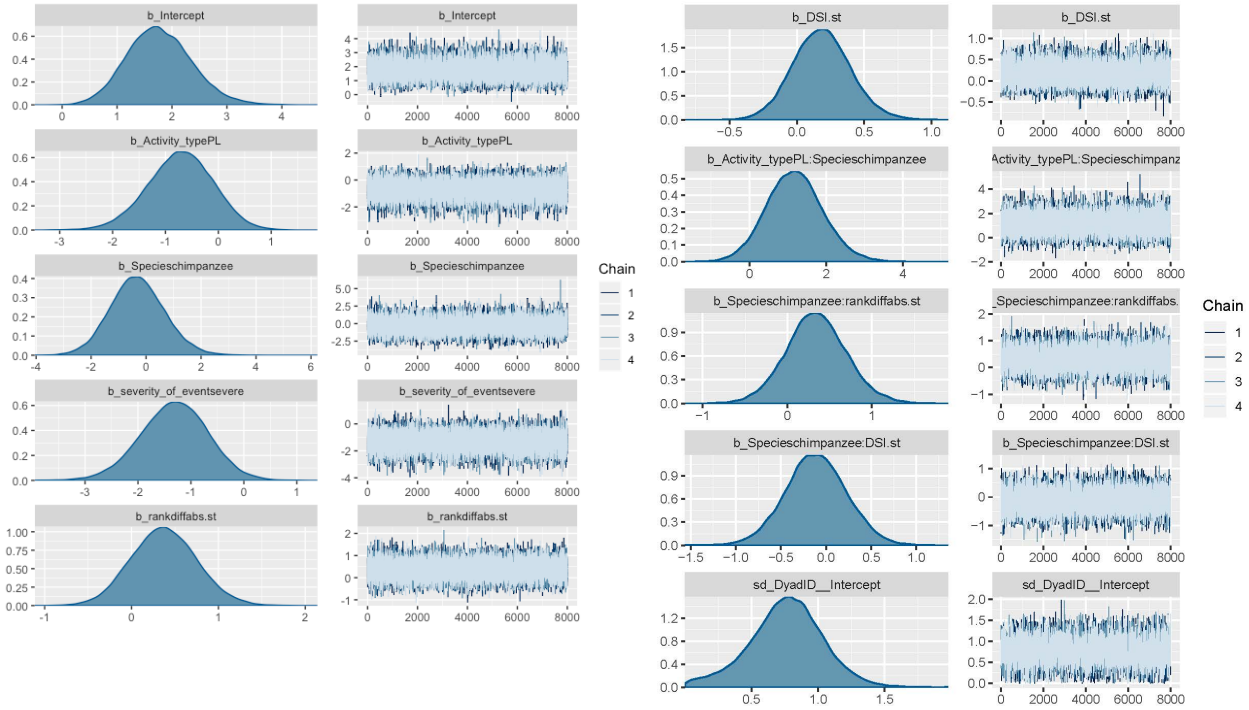

sd_MainbodyID__Intercept
sd_MainbodyID__Intercept
Chain
1
2
3
4

Model 2


b_Intercept
b_Activity_typePL
b_Specieschimpanzee
b_severity_of_eventsevere
b_rankdiffabs.st
b_DSI.st
b_Activity_typePL:Specieschimpanzee
b_Specieschimpanzee:rankdiffabs.st
b_Specieschimpanzee:DSI.st
Chain
1
2
3
4


sd_MainbodyID__Intercept
sd_MainbodyID__Intercept
Chain
1
2
3
4

Model 3


b_Intercept
b_Activity_typePL
b_Specieschimpanzee
b_severity_of_eventsevere
b_rankdiffabs.st
b_DSI.st
b_Activity_typePL:Specieschimpanzee
b_Specieschimpanzee:rankdiffabs.st
b_Specieschimpanzee:DSI.st
sd_DyadID__Intercept
Chain
1
2
3
4

S4.4. Leave-one-out cross-validation comparing the reduced model with models incl. added fixed effects for age, and sex differences, and for random effects. When the difference in the expected log predictive density (ELPD) is positive, the expected predictive accuracy for the second model is higher. A negative ELPD difference favors the reduced, fist model. Only ELPD differences twice as large as SE difference can be considered true differences in accuracy.

Model	LOO IC	SE	ELPD difference	SE difference	Pareto k diagnostic	Divergent chains?
"Model 1"	1043.2	36.0			Good (<0.5)	No
"Model 1 age": Added effect of age difference	1042.1	36.2	0.5	1.9	Ok (<0.7)	No
"Model 1 sex": Added effect of sex difference	1045.2	36.1	-1.0	0.3	Good (<0.5)	No
"Model 2"	741.3	30.6			Ok (<0.7)	No
"Model 2 age": Added effect of age difference	741.6	30.7	-0.1	1.1	Ok (<0.7)	No
"Model 2 sex": Added effect of sex difference	743.2	30.8	-0.9	0.5	Ok (<0.7)	No
"Model 3"	156.4	13.7			Good (<0.5)	No
"Model 3 age": Added effect of age difference	157.1	14.4	-0.4	1.7	Good (<0.5)	No
"Model 3 sex": Added effect of sex difference	158.7	15.0	-1.2	1.5	Good (<0.5)	No
"Model 3 random effect": Added random effects for dyad and activity ("main body") ID	Not reliable	Not reliable	Not reliable	Not reliable	***Very bad (38% > 1)***	Yes; 5798

Appendix 5

ESM for Chapter 6

S5.1 Group composition of bonobos at La Vallée des Singes at the time of study (May-August 2018). Names marked in bold indicate selected focal individual to be interrupted via Untargeted and Targeted interruptions, when interacting with strongly / weakly bonded partners. The selection was based on the criteria that each individual had at least one strongly and one weakly bonded partner. For information on strong vs. weakly bonded partners, see Table S5.2 below.

ID	***Birthdate***	***Sex***	***Kin relationships***
Daniela	17/06/1968	Female	Mother of David, Diwani
Diwani	11/08/1996	Male	Son of Daniela, brother of David
David	27/07/2001	Male	Son of Daniela, brother of Diwani
Khaya	19/10/2001	Female	Mother of Khalessi
Lingala	17/07/2003	Female	Mother of Swahili
Ukela	19/12/1985	Female	Mother of Moko and Kymia, sister of Ulindi
Kelele	22/07/2004	Male	*N.A.*
Kymia	12/04/2017	Female	Daughter of Ukela
Lucy	01/12/2003	Female	Mother of Yuli
Moko	04/08/2012	Male	Son of Ukela, brother of Kymia
Khalessi	12/12/2012	Female	Daughter of Khaya
Ulindi	10/10/1993	Female	Mother of Loto and Lokoro, sister of Ukela
Loto	02/09/2009	Male	Son of Ulindi, brother of Lokoro
Yuli	14/07/2014	Female	Daughter of Lucy
Swahili	21/09/2014	Female	Daughter of Lingala
Lokoro	22/05/2015	Male	Son of Ulindi, brother of Loto
Yahimba	07/08/2009	Female	*N.A.*

S5.2 Description of dyads and individuals that were focals in the experiment. The table informs about the dyads' interruption time range (ITR), gender, age, DSI (measure of social bond; index <1 indicates weak bond, index >1 indicates strong bond), rank differences, number of interruption trials, and number of times the dyad resumed the activity after interruptions across experimental conditions and solitary activities. Boldly marked names of partners within dyads indicate focal individuals.

Nature of activity	**Identification of dyad / individual**	**Interruption time range (min)[1]**	**Sex**	**Age (years)**	**DSI**	**Bond strength[2]**	**Absolute rank difference**	**Untargeted condition**		**Targeted condition**	
								N° reengagement	N° interruptions	N° reengagement	N° interruptions
Social	Daniela & **David**	7.41-14.81	F-M	51;18	4.05	Strong	162	5	6	3	4
Social	Daniela & **Diwani**	*	F-M	51;23	0.81	Weak	418	1	1	0	0
Social	**Daniela** & Kelele	3.43-6.86	F-M	51;15	1.14	Strong	791	2	2	0	0
Social	**Daniela** & Lingala	*	F-F	51;16	0.50	Weak	77	1	1	0	0
Social	**Daniela** & Ukela	5.27-10.54	F-F	51;34	1.92	Strong	268	2	2	3	3
Social	**Daniela** & Ulindi	7.88-15.75	F-F	51;26	1.55	Strong	49	3	3	1	2
Social	David & **Diwani**	*	M-M	18;23	0.56	Weak	580	1	2	1	2
Social	**David** & Loto	*	M-M	18;10	0.40	Weak	799	1	1	0	0
Social	David & **Lucy**	1.25-2.51	M-F	18;16	1.41	Strong	378	2	2	1	2
Social	David & **Ukela**	9.89-19.78	M-F	18;34	1.55	Strong	106	1	1	1	3
Social	**David** & Ulindi	9.94-19.89	M-F	18;26	0.87	Weak	211	0	1	1	1
Social	Diwani & **Kelele**	2.12-4.24	M-M	23;15	1.94	Strong	373	1	1	1	3
Social	**Diwani** & Ukela	5.62-11.24	M-F	23;34	4.12	Strong	686	3	3	2	3
Social	**Diwani** & Ulindi	4.7-9.4	M-F	23;26	0.16	Weak	369	1	2	1	1
Social	Kelele & **Khaya**	5.04-10.07	M-F	15;18	0.99	Weak	529	3	3	1	1
Social	**Kelele** & Lingala	5.33-10.66	M-F	15;16	0.41	Weak	714	1	1	0	0
Social	Kelele & **Lucy**	*	M-F	15;16	0.28	Weak	575	1	1	0	0
Social	**Kelele** & Ukela	*	M-F	15;34	1.05	Strong	1059	0	1	1	1
Social	Kelele & **Ulindi**	9.58-19.16	M-F	15;26	1.86	Strong	742	1	2	2	2
Social	**Kelele** & Yahimba	*	M-F	15;10	0.07	Weak	311	0	0	1	1
Social	Khalessi & **Khaya**	1.27-2.54	F-F	7;18	1.23	Strong	371	1	1	1	1
Social	**Khaya** & Loto	*	F-M	18;10	2.75	Strong	375	3	3	2	2
Social	Khaya & **Lucy**	1.68-3.40	F-F	18;16	1.04	Strong	46	1	1	0	0
Social	Khaya & **Ukela**	8.41-16.82	F-F	18;34	0.74	Weak	214	2	2	0	0
Social	**Khaya** & Ulindi	*	F-F	18;26	0.65	Weak	213	0	0	1	1
Social	Loto & **Ulindi**	4.75-9.50	M-F	10;26	1.20	Strong	588	1	1	0	0
Social	Lucy & **Ukela**	2.7-5.39	F-F	16;34	0.67	Weak	484	1	1	1	1
Social	**Lucy** & Yahimba	4.46-8.91	F-F	16;10	0.83	Weak	264	1	1	0	0
Social	Ukela & **Ulindi**	*	F-F	34;26	1.37	Strong	317	1	1	1	1

Nature of activity	Identification of dyad / individual	Interruption time range (min)[1]	Sex	Age (years)	DSI	Bond strength[2]	Absolute rank difference	Untargeted condition		Targeted condition	
Social	**Ulindi** & Yahimba	14.58-29.16	F-F	26;10	1.04	Strong	431	0	1	1	2
Solitary	Daniela	*N.A.*	F	51	*N.A.*	*N.A.*	*N.A.*	1	1	0	0
Solitary	Khalessi	*N.A.*	F	7	*N.A.*	*N.A.*	*N.A.*	1	5	0	0
Solitary	Khaya	*N.A.*	F	18	*N.A.*	*N.A.*	*N.A.*	1	1	0	1
Solitary	Lingala	*N.A.*	F	16	*N.A.*	*N.A.*	*N.A.*	1	1	0	0
Solitary	Lokoro	*N.A.*	M	4	*N.A.*	*N.A.*	*N.A.*	1	1	0	4
Solitary	Loto	*N.A.*	M	10	*N.A.*	*N.A.*	*N.A.*	1	1	0	0
Solitary	Lucy	*N.A.*	F	16	*N.A.*	*N.A.*	*N.A.*	0	2	0	1
Solitary	Swahili	*N.A.*	F	5	*N.A.*	*N.A.*	*N.A.*	2	4	1	1
Solitary	Ukela	*N.A.*	F	34	*N.A.*	*N.A.*	*N.A.*	1	1	0	0
Solitary	Yahimba	*N.A.*	F	10	*N.A.*	*N.A.*	*N.A.*	0	1	0	0
Solitary	Yuli	*N.A.*	F	5	*N.A.*	*N.A.*	*N.A.*	0	0	1	1

Note. *For these dyads, there were either not sufficient or no grooming interactions to be analyzed from the previous data collection period in 2017. This meant, we could not pre-determine an interruption point for these dyads based on their median grooming time. Therefore, we interrupted these dyads within a standard time window between at least 3 and maximum 18 minutes (i.e., considered the interruption point ranges from predetermined dyads above, this seemed an appropriate time window for dyads with unknown main body lengths).

[1]Interuption time ranges were computed as time period in which dyads' grooming bouts were interrupted by either Untargeted or Targeted stimuli. The begin point of the range corresponds to the mid-time-point between the start of grooming and the median duration of the dyads' grooming bouts. The end point of the range corresponds to the median duration of the dyad's grooming bouts.

[2]Dyad's bond strengths were computed by dyadic sociality indeces (Silk et al., 2006, 2013) with the "socialindices" package in R (Neumann, 2017, accessed via https://github.com/gobbios/socialindices).

S5.3 Definition and frequency of signals used by all individuals to communicate about suspension and reengagement. Each signal followed intentionality criteria (Genty et al., 2009). Definition taken and partly adapted from Byrne et al., 2017; Clay, 2011; Clay, Archbold, & Zuberbühler, 2015; Clay & Zuberbühler, 2009; de Waal, 1988.

Signal type	**Description**	**Untargeted condition**		**Targeted condition**	
		Frequency at suspension	Frequency at resumption	Frequency at suspension	Frequency at resumption
Gestures					
Directional hand on	Touching other's body part (usually back) with palm of hand and maintaining contact for more than 2 s to orient partner in desired direction				
Directional push	Gentle push on another's body part with hand(s), arm(s), feet or head, to orient partner's body in desired direction		1		
Grab	Grabbing gently another's body part with closed hand(s)	1			4
Grab-pull	Grabbing gently another 's body part with closed hand(s) and pulling towards self		3		1
Hand on	Touching head (or other body part) of another with palm(s) of hand(s) and maintaining contact for more than 2 s	2	1		2
Head nod	Nodding head up and down on vertical axis	1			1
Push	Pushing away gently another with hand(s) or arm(s)			1	1
Reach hand	Holding a hand toward another by extending the arm and hand				1
Reach leg	Holding a foot toward another by extending the leg and foot		2		2
Shake head	Shaking head from side to side on horizontal axis	1	4	1	4
Stretch over	Stretching and raising arm till about head level with the palm facing downwards, like embracing another's body without touching	1			

Signal type	Description	Untargeted condition		Targeted condition	
Touch	Touching gently another individual's body part with palm of hand, for under 2s	8	3	3	3
Body signals					
Present rump	Standing quadrupedally in front of another with dorsal side up to expose hindquarters, while looking back at recipient				2
Present body part	Presenting body part by stretching body to deliberately expose an area of own body to another's attention, usually to be groomed or invitation to climb-on		10		5
Present concave back	Sitting in front of another with arched back to expose genitals with legs spread apart				1
Facial expressions					
Bared teeth	Retraction of lips; partial or complete exposure of the teeth and gums; mouth practically closed; can be combined with vocalization	3			
Pout face	Lips pursed forward and curled outward in front (circular opening); lips pressed together at mouth corners	1			
Tightened lips*	Lips horizontally tensed; drawn inward or slightly baring teeth	3	2		
Vocalizations					
Peep	High-frequency, closed mouth vocalization; short in duration and flat, simple acoustic structure	9	1	2	
Unknown	A vocalization that could not be classified without acoustic analysis	2			

Note. *The tightened lips facial expression was always accompanied by other signals.

S5.4 Summary of Bayesian generalized mixed models

Model 1 (Dependent variable: Resumption individual/dyad yes/no, *n*=111)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.41	0.27	[0.89;1.96]	15152	1.00	Student t (3,0,10)
Nature of activity						
Solitary	*Reference*					
Social	-1.73	0.49	[-2.71;-0.77]	16611	1.00	Student t (3,0,10)
Random effects						
	N.M.*					

Model 2 (Dependent variable : Resumption dyad yes/no, *n*=85)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.91	0.44	[1.12;2.83]	19622	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	-0.85	0.58	[-2.02;0.26]	19622	1.00	Student t (3,0,10)
DSI	0.15	0.31	[-0.42;0.78]	25765	1.00	Student t (3,0,10)
Absolute rank difference	-0.01	0.29	[-0.59;0.57]	25652	1.00	Student t (3,0,10)
Random effects						
	N.M.*					

Model 3 (Dependent variable : Resumption individual yes/no, *n*=158						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-0.72	0.52	[-1.76;0.30]	22860	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.98	0.45	[-1.91;-0.13]	20686	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	0.30	0.42	[-0.52;1.11]	25311	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.16	0.38	[0.43;1.92]	31787	1.00	Student t (3,0,10)
Random effects						
Individual ID	0.67	0.44	[0.05;1.72]	7381	1.00	Student t (3,0,10)

Model 4 (Dependent variable : Responsible individuals for suspension communicate to suspend, yes/no, *n*=79)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.06	0.84	[-2.86;0.54]	12663	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	-1.46	0.72	[-2.99;-0.11]	23911	1.00	Student t (3,0,10)
DSI	-0.94	0.46	[-1.93;-0.13]	17026	1.00	Student t (3,0,10)
Rank difference responsible for suspension & partner	-1.24	0.72	[-2.82;-0.03]	9245	1.00	Student t (3,0,10)
Random effects						
Responsible for suspension ID	1.93	1.10	[0.33;4.63]	7040	1.00	Student t (3,0,10)

Model 5 (Dependent variable : Resumers communicate to resume, yes/no**,** *n*=66)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	0.08	0.36	[-0.63;0.80]	43486	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	1.93	0.73	[0.59;3.47]	28220	1.00	Student t (3,0,10)
DSI	0.21	0.31	[-0.39;0.82]	34083	1.00	Student t (3,0,10)
Rank difference resumer & partner	1.00	0.33	[0.38;1.68]	21994	1.00	Student t (3,0,10)
Random effects						
	N.M.*					

Model 6 (Dependent variable : Communication at time of suspension, yes/no**,** *n*=158)

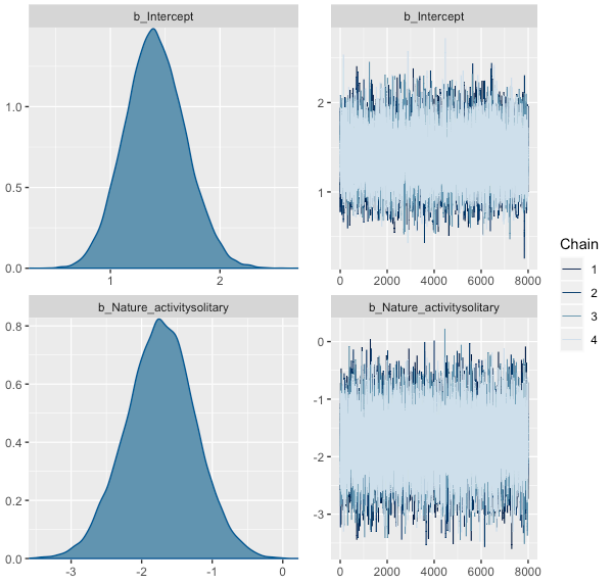
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.57	0.61	[-2.84;-0.43]	23972	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.83	0.53	[-1.90;0.17]	29595	1.00	Student t (3,0,10)
Responsible-non responsible	*Reference* 0.45	0.52	[-0.57;1.48]	30973	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	0.05	0.44	[-0.83;0.92]	41967	1.00	Student t (3,0,10)
Random effects						
Dyad ID	0.89	0.45	[0.09;1.87]	7167	1.00	Student t (3,0,10)

Model 7 (Dependent variable : Communication of individual at moment of resumption, yes/no, *n*=126)						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.39	0.66	[-2.72;-0.12]	14907	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.12	0.55	[-1.21;0.94]	15938	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	1.06	0.49	[0.11;2.05]	23103	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	0.97	0.44	[0.12;1.83]	27572	1.00	Student t (3,0,10)
Random effects						
Individual ID	0.99	0.55	[0.15;2.29]	7992	1.00	Student t (3,0,10)

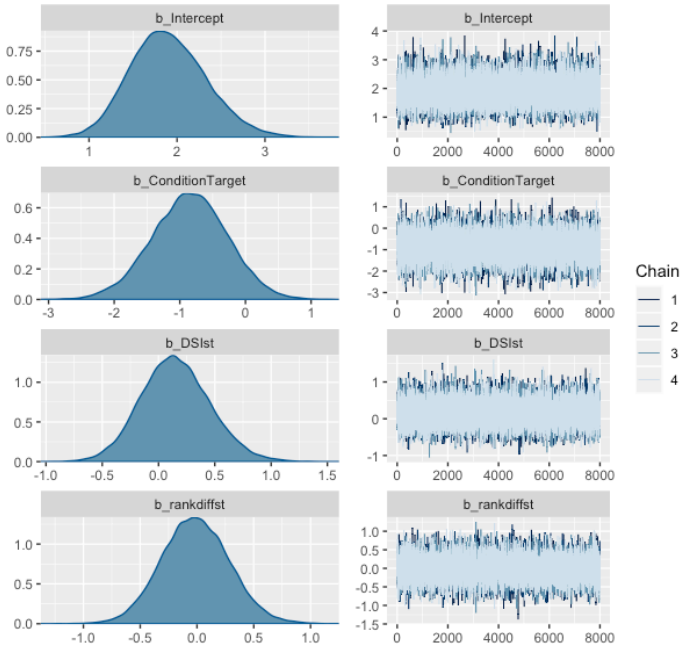
*N.M. = not modeled as variance of random effects was nearly 0, and because model without random effects had best predictive accuracy (see LOOIC comparisons). A GLM was fit instead.

S5.5 Posterior Distributions and MCMC Trace-plots of Bayesian generalized linear mixed models

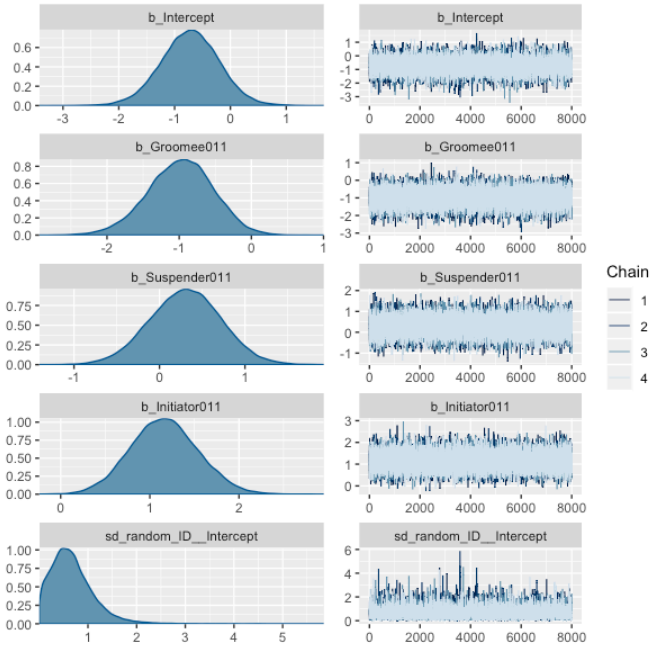
Model 1



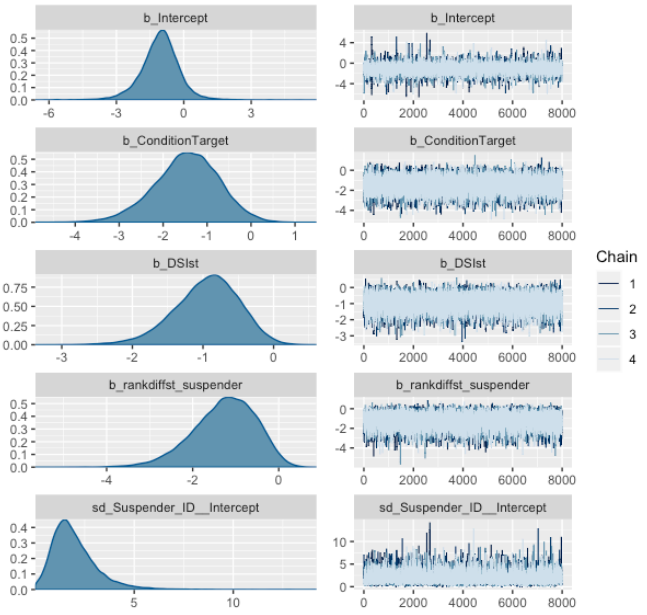
Model 2



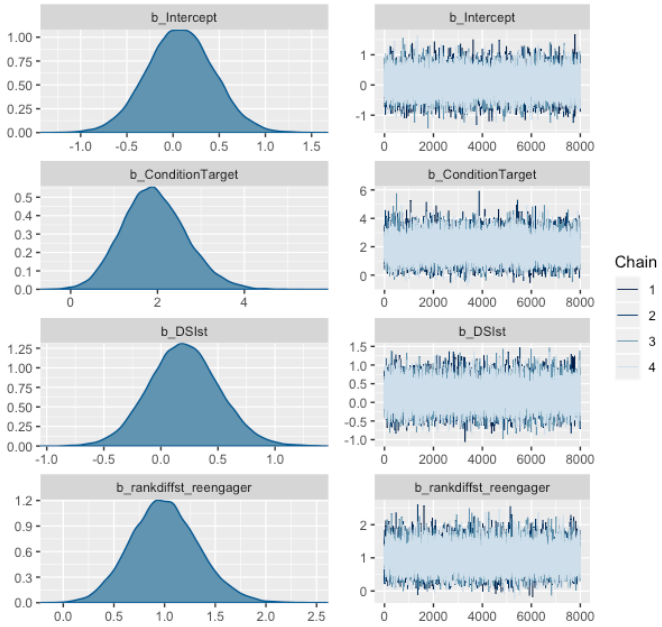
Model 3



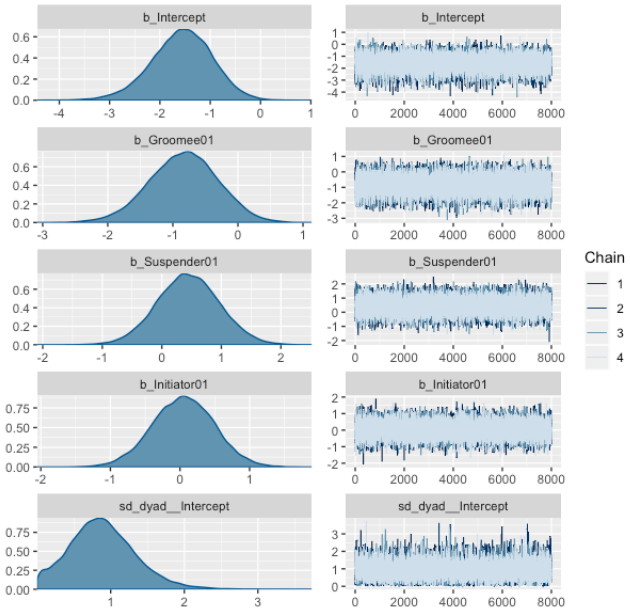
Model 4



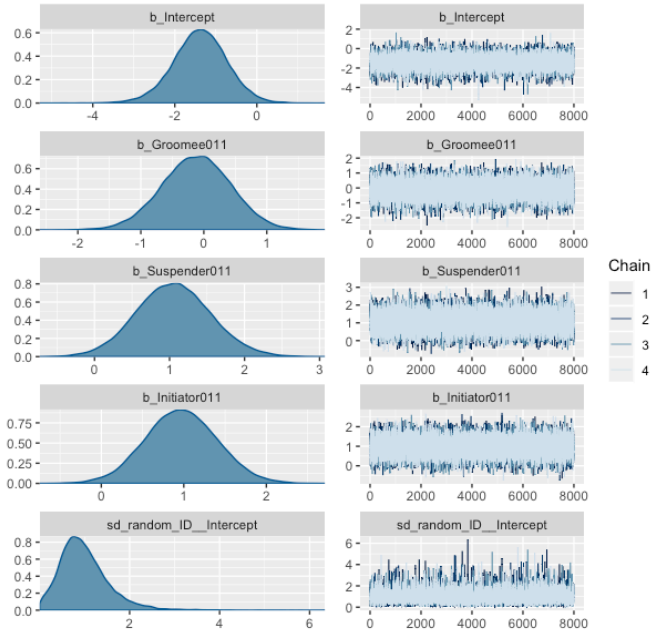
Model 5



Model 6



Model 7


b_Intercept
b_Intercept
b_Groomee011
b_Groomee011
b_Suspender011
b_Suspender011
b_Initiator011
b_Initiator011
sd_random_ID__Intercept
sd_random_ID__Intercept
Chain
1
2
3
4

S5.6. Leave-one-out cross-validation comparing model fit including or excluding random effects. The first listed model version is always compared against following model versions. When the difference in the expected log predictive density (ELPD) is positive, the expected predictive accuracy for the full model is higher. A negative ELPD difference favors the reduced model. The selected final model is marked with underscore. We also checked posterior distributions of random effects via graphical inspection (see Figure S3). If the random effects in the reduced models had a variance near 0 and lower LOOIC, the models were presented as GLM in the main paper, instead of GLMM, as the random effects modelling was not justified. For problematic pareto k values we set IC comparisons to RELOO=TRUE.

Model	*LOO IC*	*SE*	*ELPD difference*	*SE difference*	*Pareto k diagnostic*	*Divergent chains?*
Model 1 Full model (r*andom effects* dyad/individual ID)	127.1	11.3			Ok (<0.7)	No
Reduced model (no random effects)	124.8	10.9	-1.2	0.9	Good (<0.5)	No
Model 2 Full model (random effects dyad ID)	95.6	12.5			Ok (<0.7)	No
Reduced model (no random effects)	92.3	12.0	-1.7	0.8	Good (<0.5)	No
Model 3 Full model (random effects individual ID and dyad ID)	206.2	10.5			Good (<0.5)	No
Reduced model (random effect individual ID)	203.8	10.3	-1.2	0.2	Good (<0.5)	No

Reduced model (random effect dyad ID)	206.4	9.6	0.1	1.9	Good (<0.5)	No
Reduced model (no random effects)	204.0	9.4	-1.1	2.0	Good (<0.5)	No
Model 4 Full model (random effects responsible for suspension ID and partner ID)	93.3	12.1			Ok (<0.7)	No
Reduced model (random effect responsible for suspension ID)	89.8	11.5	-1.7	1.4	Ok (<0.7)	No
Reduced model (no random effect)	90.9	10.1	-1.2	3.9	Good (<0.5)	No
Model 5 Full model (random effects resumer ID and partner ID)	83.4	11.3			Ok (<0.7)	No
Reduced model (random effect resumer ID)	79.7	10.2	-1.8	1.2	Ok (<0.7)	No
Reduced model (no random effect)	76.6	9.7	-3.4	1.6	Good (<0.5)	No
Model 6 Full model (random effects individual ID and dyad ID)	161.2	16.0			Ok (<0.7)	No
Reduced model (random effect individual ID)	160.8	15.6	-0.2	2.2	Good (<0.5)	No
Reduced model (random effect dyad ID)	158.9	15.6	-1.2	1.0	Ok (<0.7)	No

Reduced model (no random effects)	159.4	15.3	-0.9	2.7	Good (<0.5)	No
Model 7						
Full model (random effects individual ID and dyad ID)	169.0	9.3			Ok (<0.7)	Yes
Reduced model (random effect individual ID)	166.5	8.9	-1.2	0.7	Ok (<0.7)	No
Reduced model (random effect dyad ID)	172.0	7.5	1.5	2.8	Good (<0.5)	No
Reduced model (no random effects)	169.6	7.3	0.3	3.0	Good (<0.5)	No

S5.7. Summary of all versions of Bayesian generalized mixed models (non-selected models).

Model 1 (Dependent variable: Resumption dyads yes/no, *n*=111) *Full model (r*andom effects *dyad/individual ID)*						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	1.54	0.36	[0.92;2.35]	18160	1.00	Student t (3,0,10)
Nature of activity						
Solitary	*Reference*					
Social	-1.78	0.60	[-2.99;-0.63]	31774	1.00	Student t (3,0,10)
Random effects						
Dyad/Individual ID	0.58	0.43	[0.02;1.60]	8342	1.00	Student t (3,0,10)

Model 2 (Dependent variable : Resumption dyad yes/no, *n*=85) *Full model (random effects dyad ID)*						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	2.04	0.52	[1.14;3.18]	18212	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	-0.80	0.61	[-2.02;0.36]	34306	1.00	Student t (3,0,10)
DSI	0.08	0.40	[-0.74;0.84]	27495	1.00	Student t (3,0,10)
Absolute rank difference	-0.02	0.35	[-0.71;0.67]	23358	1.00	Student t (3,0,10)
Random effects						
Dyad ID	0.67	0.54	[0.03;2.04]	7478	1.00	Student t (3,0,10)

Model 3 (Dependent variable : Resumption individual yes/no, *n*=158) *Full model (random effects individual ID and dyad ID)*						
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-0.74	0.54	[-1.81;0.31]	19719	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.99	0.46	[-1.91;-0.11]	22456	1.00	Student t (3,0,10)
Non-responsible	*Reference*					

Responsible	0.31	0.42	[-0.52;1.14]	19705	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.18	0.38	[0.44;1.95]	23208	1.00	Student t (3,0,10)
Random effects						
Individual ID	0.69	0.46	[0.05;1.77]	7437	1.00	Student t (3,0,10)
Dyad ID	0.22	0.18	[0.01;0.65]	22478	1.00	Student t (3,0,10)

Model 3 (Dependent variable : Resumption individual yes/no**,** *n*=158)
Reduced model (random effect dyad ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	***R̂***	***Default Prior***
Intercept	-0.73	0.43	[-1.59;0.10]	39660	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.82	0.40	[-1.61;-0.05]	39848	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	0.35	0.40	[-0.44;1.13]	42024	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.11	0.36	[0.42;1.81]	55071	1.00	Student t (3,0,10)
Random effects						
Dyad ID	0.22	0.17	[0.01;0.64]	23819	1.00	Student t (3,0,10)

Model 3 (Dependent variable : Resumption individual yes/no**,** *n*=158)
Reduced model (no random effects)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	***R̂***	***Default Prior***
Intercept	-0.73	0.42	[-1.56;0.10]	22860	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.81	0.40	[-1.60;-0.03]	22651	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	0.35	0.40	[-0.44;1.12]	19705	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.09	0.35	[0.41;1.79]	26800	1.00	Student t (3,0,10)

Random effects						
	N.M.*					

Model 4 (Dependent variable : Responsible for suspensions communicate to suspend yes/no**,** *n*=79)
Full model (random effects responsible for suspension ID and partner ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.40	1.25	[-4.30;0.84]	7787	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	-1.58	0.80	[-3.27;-0.12]	19074	1.00	Student t (3,0,10)
DSI	-1.24	0.65	[-2.75;-0.21]	6364	1.00	Student t (3,0,10)
Rank difference responsible for suspension & partner	-1.48	1.11	[-4.12;0.28]	6352	1.00	Student t (3,0,10)
Random effects						
Responsible for suspension ID	2.53	1.60	[0.45;6.59]	6760	1.00	Student t (3,0,10)
Partner ID	1.11	1.02	[0.04;3.79]	5343	1.00	Student t (3,0,10)

Model 4 (Dependent variable : Responsible for suspensions communicate to suspend yes/no**,** *n*=79)
Reduced model (no random effect)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-0.80	0.35	[-1.51;-0.14]	40123	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	-1.18	0.61	[-2.44;-0.02]	30701	1.00	Student t (3,0,10)
DSI	-0.49	0.32	[-1.16;0.11]	32863	1.00	Student t (3,0,10)
Rank difference responsible for suspension & partner	-0.22	0.28	[-0.77;0.32]	34382	1.00	Student t (3,0,10)
Random effects						
	N.M.*					

Model 5 (Dependent variable : Resumers communicate to resume , yes/no**,** *n*=66)
Full model (random effects resumer ID and partner ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***

Intercept	-0.01	0.65	[-1.39;1.23]	16527	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	2.09	0.82	[0.60;3.81]	33201	1.00	Student t (3,0,10)
DSI	0.15	0.38	[-0.58;0.89]	32033	1.00	Student t (3,0,10)
Rank difference resumer & partner	1.16	0.53	[0.19;2.32]	17021	1.00	Student t (3,0,10)
Random effects						
Resumer ID	0.82	0.78	[0.03;2.85]	9072	1.00	Student t (3,0,10)
Partner ID	0.86	0.71	[0.03;2.63]	9546	1.00	Student t (3,0,10)

Model 5 (Dependent variable : Resumers communicate to resume , yes/no**,** *n*=66)
Reduced model (random effects resumer ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	0.03	0.50	[-0.99;0.99]	17667	1.00	Student t (3,0,10)
Experimental condition						
Untargeted	Reference					
Targeted	2.00	0.76	[0.61;3.59]	30626	1.00	Student t (3,0,10)
DSI	0.19	0.33	[-0.46;0.85]	31677	1.00	Student t (3,0,10)
Rank difference resumer & partner	0.98	0.40	[0.20;1.78]	22436	1.00	Student t (3,0,10)
Random effects						
Resumer ID	0.69	0.63	[0.03;2.30]	9893	1.00	Student t (3,0,10)

Model 6 (Dependent variable : Communication at time of suspension yes/no**,** *n*=158)
Full model (random effects individual ID and dyad ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.60	0.66	[-2.98;-0.35]	23311	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.96	0.56	[-2.09;0.12]	27964	1.00	Student t (3,0,10)
Responsible-	*Reference*					

non responsible	0.42	0.53	[-0.62;1.48]	32358	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	0.09	0.46	[-0.80;1.00]	34559	1.00	Student t (3,0,10)
Random effects						
Individual ID	0.55	0.43	[0.02;1.59]	11243	1.00	Student t (3,0,10)
Dyad ID	0.87	0.46	[0.08;1.88]	8710	1.00	Student t (3,0,10)

Model 6 (Dependent variable : Communication at time of suspension yes/no**,** *n*=158)
Reduced model (random effects individual ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.45	0.58	[-2.66;-0.36]	22307	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.90	0.53	[-1.96;0.11]	22994	1.00	Student t (3,0,10)
Responsible- non responsible	0.38	0.50	[-0.59;1.38]	26155	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	0.12	0.44	[-0.74;0.99]	29720	1.00	Student t (3,0,10)
Random effects						
Individual ID	0.57	0.41	[0.03;1.58]	8718	1.00	Student t (3,0,10)

Model 6 (Dependent variable : Communication at time of suspension yes/no**,** *n*=158)
Reduced model (no random effects)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.40	0.50	[-2.41;-0.44]	21565	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.76	0.49	[-1.73;0.19]	21510	1.00	Student t (3,0,10)
Responsible- non responsible	0.41	0.48	[-0.53;1.36]	22515	1.00	Student t (3,0,10)

Receiver	*Reference*					
Initiator	0.07	0.42	[-0.74;0.89]	25379	1.00	Student t (3,0,10)
Random effects						
Individual ID	N.M.*					

Model 7 (Dependent variable : Communication at time of resumption yes/no, *n*=126)
Full model (random effects individual ID and dyad ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.46	0.71	[-2.92;-0.12]	20987	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	-0.13	0.56	[-1.24;0.95]	21826	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	1.09	0.50	[0.12;2.10]	29671	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.02	0.46	[0.14;1.93]	37595	1.00	Student t (3,0,10)
Random effects						
Individual ID	1.08	0.59	[0.17;2.50]	8318	1.00	Student t (3,0,10)
Dyad ID	0.41	0.32	[0.02;1.18]	11879	1.00	Student t (3,0,10)

Model 7 (Dependent variable : Communication at time of resumption yes/no, *n*=126)
Reduced model (random effects dyad ID)

Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.49	0.53	[-2.58;-0.50]	30280	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	0.32	0.46	[-0.57;1.23]	36981	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	1.17	0.47	[0.26;2.10]	35329	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	1.00	0.39	[0.24;1.79]	46009	1.00	Student t (3,0,10)
Random effects						

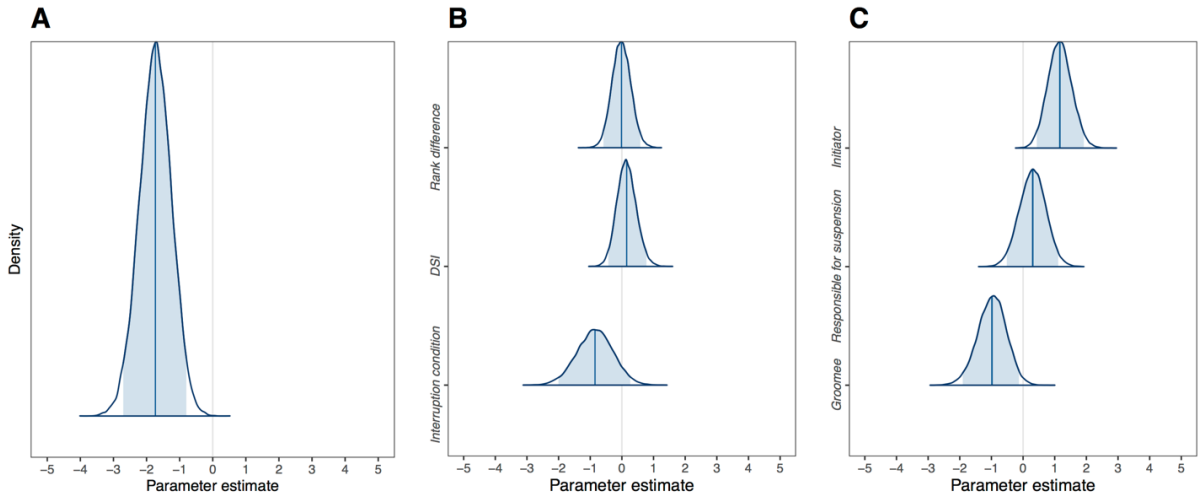
Dyad ID	0.33	0.25	[0.01;0.93]	13902	1.00	Student t (3,0,10)

Model 7 (Dependent variable : Communication at time of resumption yes/no, *n*=126)
Reduced model (no random effects)

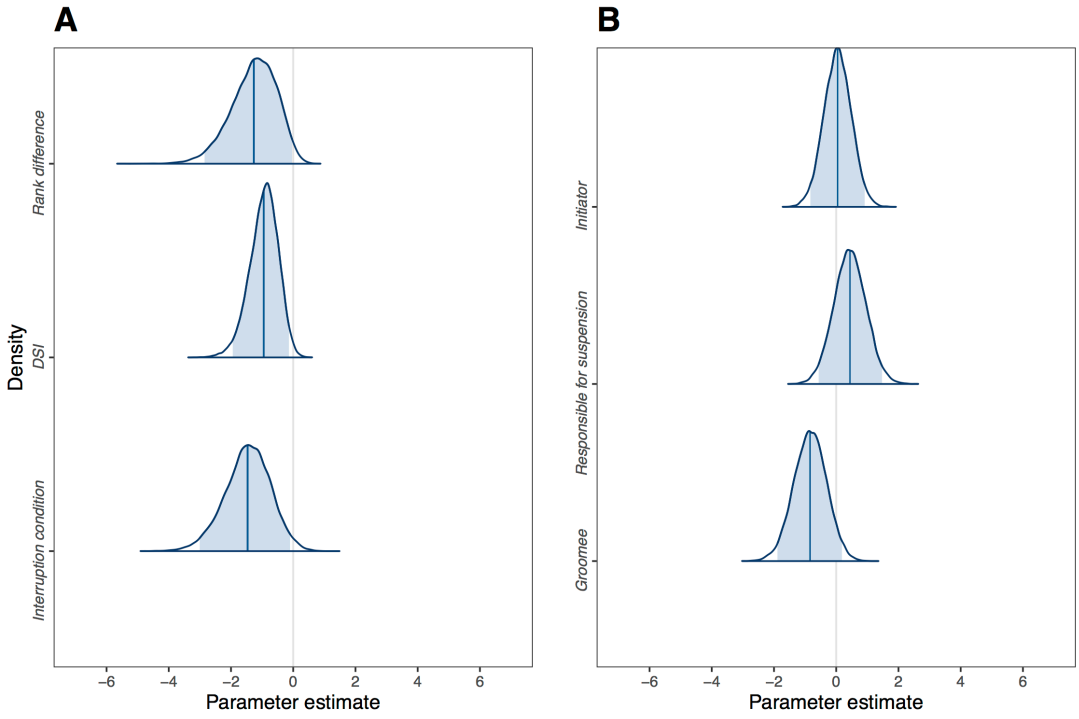
Fixed effects	***b***	***SE***	**95% CrI**	***Effective N***	$\hat{R}$	***Default Prior***
Intercept	-1.42	0.50	[-2.43;-0.47]	18920	1.00	Student t (3,0,10)
Role						
Groomer	*Reference*					
Groomee	0.31	0.45	[-0.56;1.21]	20652	1.00	Student t (3,0,10)
Non-responsible	*Reference*					
Responsible	1.13	0.45	[0.27;2.05]	21017	1.00	Student t (3,0,10)
Receiver	*Reference*					
Initiator	0.97	0.39	[0.22;1.74]	25363	1.00	Student t (3,0,10)
Random effects						
	N.M.*					

*N.M. = not modeled as variance of random effects was nearly 0, and because model without random effects had best predictive accuracy (see LOOIC comparisons above). A GLM was fit instead.

S5.8. Density plots representing the posterior distribution for the marginal effects nature of activity (A), interruption condition, DSI and rank difference (B), and roles "groomee", "responsible for suspension" and "initiator" (C) of the models 1, 2, and 3, respectively, with dependent variable being "resumption". The blue highlighted areas correspond to the 95% credible interval of the estimated effects. The value of 0 is highlighted by a vertical grey line, and the estimated posterior means are highlighted by a blue vertical line.



S5.9. Density plots representing the posterior distribution for the marginal effects interruption condition, DSI and rank difference (A), and roles "groomee", "responsible for suspension" and "initiator" (B) of the models 4 and 6, respectively, with dependent variable being "suspension communication". The blue highlighted areas correspond to the 95% credible interval of the estimated effects. The value of 0 is highlighted by a vertical grey line, and the estimated posterior means are highlighted by a blue vertical line.



S5.10. Density plots representing the posterior distribution for the marginal effects interruption condition, DSI and rank difference (A), and roles "groomee", "responsible for suspension" and "initiator" (B) of the models 5 and 7, respectively, with dependent variable being "resumption communication". The blue highlighted areas correspond to the 95% credible interval of the estimated effects. The value of 0 is highlighted by a vertical grey line, and the estimated posterior means are highlighted by a blue vertical line.