Emotions through the eyes of our closest living relatives: exploring attentional and behavioral mechanisms

Berlo, E. van

Citation

Berlo, E. van. (2022, May 19). *Emotions through the eyes of our closest living relatives: exploring attentional and behavioral mechanisms*. Retrieved from https://hdl.handle.net/1887/3304204

Emotions through the eyes of our closest living relatives: Exploring attentional and behavioral mechanisms

Evy van Berlo

Emotions through the eyes of our closest living relatives: Exploring attentional and behavioral mechanisms

Evy van Berlo

ISBN:	978-90-9035984-7
Design/Lay-out:	Ferdinand van Nispen tot Pannerden, *my*-thesis.nl
Cover design:	Evy van Berlo
Print:	Ovimex, Deventer NL

Emotions through the eyes of our closest living relatives: Exploring attentional and behavioral mechanisms

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof.dr.ir. H. Bijl,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 19 mei 2022
klokke 11.15 uur

door

Evy van Berlo
geboren op 26 maart 1990
te Maastricht

Promotor: Prof.dr. Mariska E. Kret
Co-promotor: Dr. Yena Kim

Promotiecommissie: Prof.dr. Sander T. Nieuwenhuis (Wetenschappelijk Directeur Instituut Psychologie / voorzitter)
Prof.dr. Karline R. L. Janmaat
Prof.dr. Elisabeth H. M. Sterck, Universiteit Utrecht
Prof.dr. Frans B. M. de Waal, Emory University
Dr. Anne E. Urai
Dr. Katharina Riebel

The research included in this dissertation was supported by grants from Stichting Elise Mathilde Fonds (Leiden University Fund, awarded to Mariska E. Kret and Evy van Berlo) and the Dr. J. L. Dobberke Stichting voor Vergelijkende Psychologie (Koninklijke Nederlandse Akademie van Wetenschappen, awarded to Evy van Berlo)

Acknowledgments

My Ph.D. journey would not have been made possible without the support of many dear friends and colleagues. Regardless of whether you find your name down below, I want to express my sincere gratitude for your support. I want to especially thank the following people:

Mariska Kret, thank you for being a great teacher, for giving me opportunities that I thought to be unreachable. I admire your perseverance, ambition, and your unyielding support of women in science. Thank you, Jorg Massen, for your kindness and support during my Master's and my Ph.D. Your critical and creative thinking has been an inspiration to me and helped me elevate my work to a higher level. Daan Laméris, thank you for your emotional and intellectual support, and your open-minded attitude. Your help in testing the orangutans across multiple studies has been invaluable. I am very proud of our work together. Alejandra Díaz-Loyo and Oscar Juárez-Mora, I am grateful for your fantastic collaboration on the orangutan yawning project. Your help has been crucial to the success of the project and thanks to you and Jorg, the project was able to continue during my sick leave.

My paranymphs Tom Roth and Iliana Samara; your wittiness and humor always make my day. I thank both of you for your immense support, your kindness, creativity, and your pragmatism. I learned so much from you, and you really helped me improve my writing and analysis skills. Thank you, Friederike Behrens-Scholten, for your never-ending enthusiasm, your keen eye for detail, and your help during difficult times. You light up the room like no other! Joyce Snijdewint, thank you for being such a kind-hearted person and a great friend. Thank you for always listening and offering help where you can. Your perseverance is inspiring to me. Yena Kim and Julia Folz, I highly enjoyed our karaoke nights to wind down from work. Julia, I loved your daily tea quotes on the whiteboard pre-corona times, and Yena, your cooking skills are simply amazing. Your down-to-earth attitude and insightful comments have been very helpful in my work. Tonko Zijlstra, thank you for your humor and your warmhearted spirit that kept things fun and light during work in the pandemic. Linda Jaasma, your training of the bonobos has been crucial to the success of several of my studies. Thank you for building the foundation that was necessary to succeed.

I also sincerely thank the Apenheul caretakers; Jacqueline Ruijs, Rudy Berends, Keetie de Koeier, Iris Schapelhouman, Grietje Grootenhuis, Carolijn de Jong, Roan de Boer, Bianca Klein, Frank Rijsmus, Cora Schout, Tijs Swennenhuis, and Martine Verheij. Without your time, support, and patience, these studies would not have

been possible. Thank you, Thomas Bionda, for facilitating the great ape studies at Apenheul, and Wendy Cieremans as well as Apenheul's volunteers for facilitating the research with visitors. I am grateful for having had the chance to work in such an exciting and beautiful environment.

I thank my participants, ape and human alike, for taking part in my studies, and the many students who helped with collecting all the data. Thank you to the technical support staff: Maureen Meekel, Evert Dekker, Elío Sjak-Shie, and Iris Spruit, for building custom-made devices and scripts to help with data acquisition and analyses. I also thank the doctorate committee for investing their valuable time in reading and evaluating my work, and for providing helpful feedback.

Thank you to my dear friends, Leonie Groeneveld, Sander Geuke, and Samantha Lems, for your never-ending support. You have been there with me all the way, and are truly the best! Jeannot van Berlo, thank you for being my creative, free-spirited brother. Jos van Riet and Thea van Riet-Sprengers, my parents-in-law, thank you for your guidance when I needed it the most. My aunt Mieke Spaetjens-Beerens and uncle Gerard Spaetjens: thank you for reading my Dutch summary and helping me improve it. Thank you to my sweet parents, Corry van Berlo-Beerens and Ger van Berlo, for always believing in me. Thank you for having taught me to never stop asking questions, and to persevere in the face of hardships. Without you, I would not have been where I am now. Finally, I thank you, Martijn van Riet, my fantastic boyfriend. Thank you for your unconditional support, your guidance, love, and trust. You are my beacon and my anchor. I love you.

Evy van Berlo
November 19th, 2021, The Hague

Table of Contents

Chapter 1

A general introduction to emotion perception and its underlying mechanisms

What are emotions? To many people, emotions are subjectively felt affective states. There is no doubt in our minds that we experience emotions, and that others do, too. Nevertheless, describing what exactly an emotion *is* has proved to be a major scientific challenge, and definitions are still highly contested (Adolphs et al., 2019; James, 1884; LeDoux, 2021; Russell & Barrett, 1999). For a long time, the lack of a clear definition for emotions stood in the way of understanding what role emotions play in not only our lives but the lives of other animals as well. Moreover, it made unraveling the evolution of emotions very difficult (Paul & Mendl, 2018). When Nikolaas Tinbergen published his seminal work on the four questions that scientists can ask to find proximate ("how") and ultimate ("why") explanations for animal behavior (Tinbergen, 1963), the topic of emotions in animals other than humans was still highly controversial. This was mainly due to the subjective nature of emotions, as animals cannot tell us what they feel or how they experience things. Since then, the interdisciplinary field that is concerned with emotions in humans and other animals has grown steadily and confidently and is now known by the name *comparative affective science* (Williams et al., 2020). With the rise of this new scientific field, attempts have been made to find a definition of emotion that makes it accessible to scientific inquiry. For instance, a broad definition was provided by Frans de Waal in his work *What is an animal emotion?*:

> *"Emotions [are] mental and bodily states that potentiate behavior appropriate to the environmental challenges"* (De Waal, 2011).

These emotional states are adaptive; they are shaped through the process of natural selection to prepare individuals for the most appropriate and optimal response. Emotional states are caused by external, biologically relevant stimuli, bringing about a range of parallel changes in an organism (e.g., behavioral, psychophysiological, cognitive, and somatic). The kind of emotion that is triggered, is closely tied to the context an individual finds itself in (De Waal, 2011). Moreover, although we typically assign subjective experiences ("feelings") to emotional states, these are arguably hard to probe in animals. However, even if we do not fully grasp the emotional experiences of animals, this understanding is not necessary to move forward (see for discussions e.g., Berridge, 2018; LeDoux, 2021; Mendl et al., 2010). Building from the premise that emotions are states that the body can be in, we can examine these states in more detail by looking at their cognitive and behavioral markers (Anderson & Adolphs, 2014; De Waal, 2011; Paul et al., 2005). These markers serve as proxies to the ″inner world″ of animals as well as of humans, and understanding them will allow scientists to determine whether the behaviors that we intuitively deem indicators of emotions in animals are indeed mediated by the same or homologous mechanisms shared between humans and other animals (Panksepp, 2011). Additionally, this will also allow us to reconstruct how our human ancestors may have behaved and felt.

The focus of this dissertation lies on how emotions are *perceived* by studying three markers of emotion perception: *attention*, *spontaneous mimicry*, and *implicit associations* (Figure 1i-iii). In this dissertation, emotional modulation of attention and mimicry is compared between humans, bonobos, and orangutans, and an adaptation to an existing paradigm to study *implicit associations* is validated for potential use in comparative science in the future. As will become clear, attention, mimicry, and implicit associations are building blocks that are fundamental to social cognition. Moreover, they provide an opportunity to study what emotions mean to other animals and make direct comparisons between species possible. The goal of the dissertation is twofold, namely to better understand the evolutionary continuity of emotion perception across hominids, but also to study the uniquely derived differences in emotion perception in the three species (Figure 1a). To do this, the

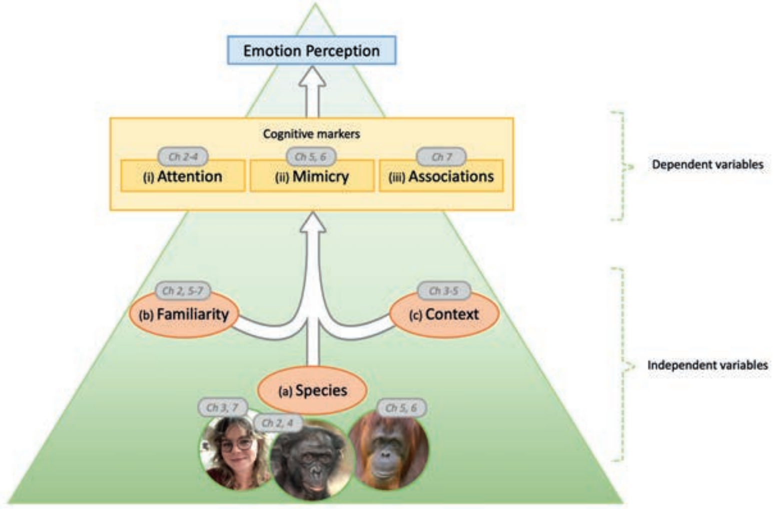

Figure 1. Schematic overview of the research topic. Emotion perception is a multifaceted phenomenon that is governed by many different cognitive mechanisms. Often, these mechanisms operate on an implicit level; automatically and unconsciously. To study emotion perception across species, I investigate its underlying implicit mechanisms or *cognitive markers*. The focus of this dissertation lies on *(i) attention, (ii) mimicry,* and *(iii) implicit associations*. Moreover, I investigate the effects of *species (a), familiarity (b),* and *context (c)* on these markers across six chapters (grey circles) in this dissertation.

effects of *species*, *familiarity*, and *context* (Figure 1b, 1c) on emotion perception are examined. Here, *familiarity* is defined as the familial and/or social relationship between the expressor of emotions and the observer. *Context* refers to the situation the observer and/or expressor of emotional expressions find themselves in (e.g., tense or relaxed situations).

A window into our evolutionary past

What are the evolutionary roots of emotions? In search of answers to this cardinal question, one logical step we can take is studying emotions in animals that are evolutionarily close to us: the great apes. Studying the behavior and cognition of extant (i.e., still living) great apes is extremely relevant for reconstructing the social and emotional characteristics of the last common ancestor of the *Hominidae,* as well as early *hominins* (ancestral humans) (Duda & Zrzavý, 2013). The *Hominidae* consist of the African apes: gorillas, chimpanzees, bonobos, humans, and orangutans (the only Asian great ape) (Figure 2). The last common ancestor of the African apes and orangutans likely lived about 14 million years ago (Ma) (Goodman et al., 1998), with the most recent common ancestor of chimpanzees, bonobos, and humans living at least 7 to 8 Ma (Langergraber et al., 2012). It is important to remember that evolution has continued for both the great apes as well as humans. Indeed, each of the hominids

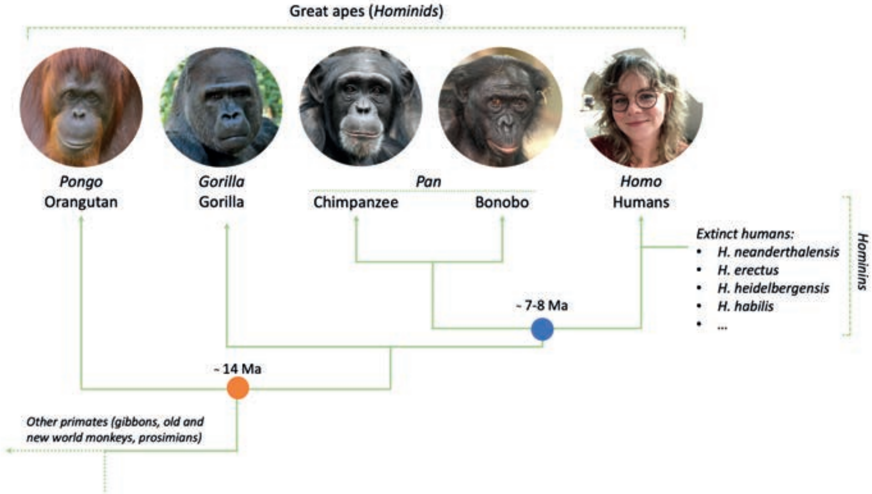

Figure 2. Great ape family tree. Bonobos and chimpanzees are humans' nearest living relatives, followed by gorillas, and then orangutans. *Pan* and *Homo* share a common ancestor roughly 7-8 million years ago (Ma). At one point in time, other human species roamed the earth together with or close in time to *Homo sapiens*. For instance, Neanderthals co-existed with *sapiens* until Neanderthals went extinct around 40.000 years ago (Higham et al., 2014).

has undergone unique changes: adaptations that made survival in their separate social and physical environments possible. Nevertheless, comparisons between species make it possible to get some sense of what our common ancestor may have lived and behaved like. If a specific characteristic is found in all the species that are compared, this may suggest a common evolutionary origin (Wilson, 2021). There are likely a lot of commonalities to be found in the behavior and cognition of great apes and humans. In this dissertation, the focus lies on two great ape species that can provide us with unique insights into the evolution of emotions, and specifically on how emotions are perceived: bonobos (*Pan paniscus*) and orangutans (*Pongo pygmaeus*).

Bonobos are known for their relatively tolerant nature, as well as their usage of sexual behaviors to reduce tension in their group, strengthen bonds with other females, and form new relationships with unfamiliar individuals (De Waal, 1988; Furuichi, 2011). Endemic to a small area in the Democratic Republic of the Congo, bonobos live in fission-fusion groups where males have life-long residency in the group and females disperse to other groups when they reach sexual maturity (Hohmann et al., 1999). Like all apes, bonobos are endangered, with a minimum estimated wild population of about 15.000-20.000, and around 225 individuals living in zoos across the world (Fruth et al., 2016). Though current research efforts into the psychology of bonobos are growing, we still understand relatively little about their social cognition. In stark contrast with their more territorial cousins, chimpanzees (*Pan troglodytes*) (Wilson et al., 2014), as well as ancestral humans (Bowles, 2009; Wrangham & Glowacki, 2012), bonobos show remarkable xenophilic tendencies. Females of different groups show high social tolerance and tend to affiliate with each other rather than fight (Furuichi & Thompson, 2007). Even within their social groups, bonobos show high levels of affiliation and cooperation, and low levels of aggression. Moreover, neuroscientific studies on the brains of bonobos have found that the brain structures involved in emotion processing and regulation are bigger in volume and have more dense connections in bonobos compared to other great apes, thus making them an important referential model for the evolution of social cognition in ancestral humans (Issa et al., 2019; Rilling et al., 2012; Stimpson et al., 2016).

Native to Borneo and Sumatra, orangutans are arboreal apes that lead a semi-solitary existence that is highly unusual among the great apes (Delgado & Van Schaik, 2000; Galdikas, 1985; Mitra Setia et al., 2009). Occasional social associations among individuals do occur, but not frequently (Singleton et al., 2009; Van Schaik, 1999). It is during these temporary formations of small groups that orangutans have the opportunity for socializing, playing, and mating. The formation of these parties, as well as the general tendency to live on their own, likely has a close link to the availability of fruit, with fruit scarcity leading to less association, and fruit abundance to more social interactions (Roth et al., 2020). Cognitive research on orangutans has gained more traction in the last decade (Damerius et al., 2019; Hopper, 2017), but almost no work has looked at how orangutans perceive emotional expressions, and how their socio-cognitive abilities compare to the other apes and humans. Their unique semi-solitary nature may therefore provide interesting insights into the development of these capacities over evolutionary time.

Enhanced attention for emotions

One of the earliest processes involved in emotion processing is *attention* (Figure 1i). Attention is the gatekeeper that selectively filters relevant from irrelevant information coming from the environment (James, 1890). This process is crucial, as the brain cannot attend to all information at once. Remarkably, already at the very early stages of processing information from the environment, attention automatically and efficiently tunes to emotionally salient signals (Whalen, 1998). Indeed, a large body of evidence stemming from human studies has shown that emotions are so fundamental to our species, that our brains evolved sensory mechanisms that preferentially process emotional information over other, more neutral signals (Phelps & LeDoux, 2005). Especially negative emotions (for instance fear or anger) appear to strongly capture our attention. From an evolutionary perspective, this makes a lot of sense. Perceiving anger or fear in others could mean that there is imminent danger, requiring immediate action from the observer. Early studies on this so-called *implicit attentional bias* for threatening signals, showed that humans automatically attend to threatening stimuli such as snakes, spiders, or angry faces (Öhman et al., 2001a), and suggested the brain and especially its emotion center (including subcortical structures such as the amygdala) is "hard-wired" to detect such threats in the service of evolutionary goals (Öhman et al., 2007).

New research suggests that parts of the brain's emotion centers are not just hard-wired threat detectors, but are highly sensitive to *motivationally* relevant emotional signals (Cunningham et al., 2008). Indeed, enhanced attention to negatively- or positively-valenced emotions appears to differ between individuals, across developmental trajectories and the age spectrum (Todd et al., 2012), and is affected by an individual's current affective state (Mendl et al., 2009). This also may explain why highly anxious individuals show a particularly strong bias towards angry or fearful faces (Bar-Haim et al., 2007). Moreover, allocation of attention to emotional stimuli could be driven by highly salient, low-level perceptual characteristics of stimuli (for instance the visibility of teeth, direct eye gaze, or simply their novelty) and therefore reflect a more bottom-up process (Öhman et al., 2001a). Attention allocation can also be driven by a top-down process that takes into account the motivations of the observers, as well as the context the observers find themselves in (Victeur et al., 2020). Most likely, however, attention for emotions involves an interplay between bottom-up and top-down processes (Pessoa et al., 2006). Notwithstanding this debate, it is undisputed that expressions of emotion are automatically and swiftly attended to by humans. But is this attentional mechanism uniquely human, or can we find similar mechanisms in other animals?

Evidence for emotion-biased attention in animals is much more limited, and also mixed. Most of the work that exists is conducted with primates. One research task that can measure (implicit) selective attention towards emotions is the dot-probe task, which measures how fast individuals can touch a dot (i.e., the probe) on a screen after being presented with emotional and neutral stimuli (i.e., signals in the form of for instance a picture). Typically, individuals are faster at touching a dot when it is preceded by an emotional stimulus because emotions draw attention. Using the dot-probe task, Japanese macaque monkeys (*Macaca fuscata*) were found to have a threat-specific attentional bias (Masataka et al., 2018) that also extended to threatening facial expressions (Lacreuse et al., 2013; Parr et al., 2013). However, enhanced attention to threatening signals was not found in chimpanzees using the same paradigm (Kret et al., 2018; Wilson & Tomonaga, 2018). In contrast, bonobos showed biased attention towards more positively valenced emotional scenes. Finally, studies measuring attentional biases with eye-tracking have shown that chimpanzees and orangutans look longer at threatening stimuli such as fearful faces or aggressive interactions (Kano & Tomonaga, 2010a; Pritsch et al., 2017).

The mixed results show that more research is needed to close the knowledge gap on how expressions of emotion affect attention in other animals, especially given the importance of implicit measures of emotion processing for comparative research. The first part of this dissertation (Chapters 2 to 4) will therefore examine emotion-biased attention in humans and one great ape species, bonobos, in more detail (Figure 2a).

Spontaneous mimicry and emotion contagion

Attention allows for studying to what extent emotions can capture and hold interest, thus to a certain extent it elucidates whether emotions are meaningful to individuals. However, it does not inform us about whether emotions are meaningful for interactions between individuals. Here, there is a role for *spontaneous mimicry*, which in the field of psychology is defined as the automatic and unconscious imitation of facial or bodily expressions of other individuals (Figure 1ii) (Chartrand & Bargh, 1999). Examples of spontaneous mimicry (henceforth: mimicry) are contagious laughter or crying, but also contagious yawning and self-scratching. Quintessentially a social phenomenon, mimicry is thought to facilitate the perception of emotions in others, as well as the transmission of emotional states between individuals (emotion contagion) (Preston & De Waal, 2002; Prochazkova & Kret, 2017).

Recently, a step-wise evolutionary development of emotion contagion was proposed, starting at multiple individuals showing a similar facial display that is automatically generated by a physiological internal state in response to external events (for instance a fearful face in response to aggressors or predators, or an expression of pain in response to bodily harm) (Palagi et al., 2020). Next, these independent facial displays may have acquired a communicative function by automatically triggering a similar response in others when observed. The function of this system may have been to synchronize activities within the group, for instance by rapidly spreading fear among group members to escape from predators. Importantly, it may have resulted in automatic emotion contagion when the displays reflected emotional states, leading individuals to feel – to a certain extent – what others are feeling. In this way, emotion contagion can help resonate with others, facilitating effective communication and social cohesion (Decety et al., 2012; Preston & De Waal, 2002). For example, spontaneous mimicry of expressions can provide a crucial fitness advantage in parental care, allowing parents to respond to the needs of their offspring (Decety et al., 2012; Roth et al., 2021).

It is important to emphasize that mimicry does not necessarily equate to uninhibited, full-blown copying of behaviors and expressions, as this would hinder individuals from appropriately responding to social situations (e.g., parents would not be able to adequately respond to their infant's needs if they would always be overwhelmed by sadness when their infant is crying (Mafessoni & Lachmann, 2019)). Rather, mimicry can be subtle, triggering the activation of a corresponding emotional state in oneself, thereby allowing the identification of what someone else is feeling and formation of an appropriate response (Preston & De Waal, 2002). As such, mimicry is considered to be a major building block for more complex socio-cognitive capacities such as empathy (De Waal & Preston, 2017; Koski & Sterck, 2010) and cooperation (e.g., to establish trust between cooperating individuals); two facets that are extensively developed in our species (Tomasello et al., 2012). Thus, through studying mimicry, we are provided with a window into the basic capabilities of humans and other animals to process emotions.

In the last decade, interest has grown in studying contagious yawning as a form of motor mimicry (Massen et al., 2012; Miller et al., 2012b; Norscia & Palagi, 2011a). Yawning itself is a widespread behavior among vertebrates, yet it is contagious only in a limited number of social species. In addition, it has been suggested that yawn contagion is linked to emotional contagion, as contagiousness is stronger between kin and friends (Palagi et al., 2020b). At the same time, the purported link between contagious yawning and emotion contagion is heavily debated, with some researchers suggesting that the effect of social closeness on contagiousness is established not because of a fundamental emotion-sharing mechanism, but because of for instance an attentional bias to familiar individuals (Gallup, 2021; Massen & Gallup, 2017).

There are a growing number of studies showing that yawning is more contagious between strongly bonded individuals (Campbell & de Waal, 2011; Joly-Mascheroni et al., 2008; Norscia et al., 2020, 2021; Palagi et al., 2014; Romero et al., 2014), but other studies report no link with closeness (Madsen et al., 2013; Madsen & Persson, 2013; Massen et al., 2012; Neilands et al., 2020; O'Hara & Reeve, 2011), making it unclear what characteristics of social relationships (if any) modulate contagious yawning. In addition to yawning, non-facial forms of mimicry (e.g., postures, gestures) have been extensively studied in humans (see Lakin et al. (2003) for a review), but much less in animals. Self-scratch contagion – one example of non-facial mimicry – has been proposed as a candidate behavior to study the link between mimicry and emotional contagion, but work on self-scratch contagion is still limited (Feneran et al., 2013; Nakayama, 2004; Schut et al., 2015). Self-scratching can be an indicator of stress or arousal in both humans and other animals (Maestripieri et al., 1992), thus its contagiousness could potentially reveal a link with *negative* emotional contagion. Continued efforts to study the possible link between contagious yawning and self-scratching and emotion contagion remain important, especially for the advancement of our understanding of how animals process emotions. The second part of this dissertation (Chapters 5 and 6) will therefore center around contagious yawning and self-scratching.

Emotional modulation of implicit associations

Sensitivity to emotions can also be measured indirectly via *implicit associations* (Figure 1iii). To navigate the complexities of the social world, it is beneficial to form simple heuristics rather than making new social evaluations with every new situation. These simple heuristics can come in the form of categorizing for instance other individuals into pleasant or unpleasant, and familiar or unfamiliar, and are driven by a cognitive system that automatically evaluates the environment (Greenwald & Banaji, 1995). Indeed, in the human literature, these unconscious evaluations are called *implicit associations*, and they guide our daily behavior. Emotions regulate these implicit social evaluations, making them stronger or weaker based on an individual's current state (Greenwald & Banaji, 1995).

Most of the work on implicit associations has been done in social psychology and has examined in- and outgroup implicit associations using the *Implicit Association Test* (IAT), which measures unconscious associations between certain concepts (e.g., objects, individuals) and evaluations (e.g., "good" or "bad", "positive" or "negative") (Greenwald et al., 1998). Implicit attitudes can for instance entail an unconscious preference for one's ingroup over an outgroup, as well as having negative associations with outgroups. Notwithstanding the importance of uncovering the cognitive mechanisms underlying intergroup biases in humans, to date, relatively little research has looked at where these processes come from in the first place. The pervasiveness of negative implicit attitudes towards outgroups (paving the way for phenomena such as prejudice and discrimination) suggests an evolutionarily old origin.

Kurzban, Tooby and Cosmides (2001) hypothesized that implicit negative attitudes towards for instance individuals of another ethnicity may be a byproduct of adaptations that once evolved to help detect coalitions and alliances in our hunter-gatherer ancestors. For quick and efficient processing of the social world, the cognitive mechanism underlying this alliance-detection becomes sensitive to otherwise meaningless markers such as physical traits to create social categories. Recently, this *alliance hypothesis of race* was further strengthened by direct experimental evidence (Pietraszewski, 2021). Given that other animals show sensitivity to intergroup biases in attention and mimicry, we may find that certain animals also implicitly evaluate other individuals.

Primates have group structures that closely resemble ours, including intergroup conflicts, hierarchies, and social categories (i.e., based on biological traits such as sex, kinship, age, etc.). Yet, we do not fully grasp the emotional and cognitive processes that drive intra- and intergroup social interactions, and evidence for intergroup biases in judgments and associations is currently limited to humans. Finding ways to probe implicit associations in animals might prove useful in progressing our understanding of the link between emotions and implicit attitudes. Having a method that probes implicit attitudes in animals can not only help us study a negative bias towards outgroups, but would also allow us to examine how animals judge emotional categories themselves.

Currently, we design experiments based on our notion of what entails an emotionally-salient cue to animals. These notions are of course grounded in the knowledge that we have gathered on emotions in animals, but it remains impossible for us to truly know the meaning of emotions to animals because they cannot use language to convey this information to us. We make a first step towards developing a way to probe all kinds of implicit associations in animals that are capable of categorizing images and working on a touchscreen in Chapter 7. Here, we validate a pictorial adaptation to the IAT in humans and adult children, with the hope that this pictorial version may one day be tested in animals. In essence, the IAT is a matching-to-sample task in which an image has to be matched to its appropriate category. Great apes have previously been shown to be capable of performing matching-to-sample tasks in which they categorized bodies in different configurations (Gao & Tomonaga, 2020), sexes (De Waal & Pokorny, 2011), genital regions (Kret & Tomonaga, 2016), familiar and unfamiliar faces (Parr et al., 2000; Pokorny & De Waal, 2009; Talbot et al., 2015; Vonk & Hamilton, 2014), facial expressions (Parr et al., 2008), and emotions (Parr, 2001). It is therefore plausible that great apes could also perform in a pictorial IAT.

Dissertation outline

This dissertation is based on six empirical research articles focusing on the unconscious and automatic cognitive and behavioral markers of emotion perception, as these markers offer a strong basis from which we can study emotions across species. Specifically, in **Chapter 2**, the role of *implicit, immediate attention* in perceiving emotions of familiar and unfamiliar conspecifics (i.e., other individuals of the same species) as well as heterospecifics (i.e., individuals of another species) is investigated in bonobos and humans (Figure 1i). Central to this chapter is i) replicating earlier findings on an emotion bias in bonobos and humans (Figure 1a), and ii) tackling how familiarity modulates attention for emotions, which to date has not yet been examined (Figure 1b).

Chapter 3 zooms in on emotion-biased attention in humans across all age categories, using emotional *scenes* as cues rather than isolated facial expressions, as they provide more contextual information to the observer (Figure 1a, c). Moreover, how children and adults perceive the emotions of other humans and bonobos is examined to assess differences and similarities in the perception of emotional expressions across closely related species.

Chapter 4 builds on the works in Chapters 2 and 3 by examining attention for emotions in bonobos and humans in more detail using an eye-tracking paradigm. Here, the goal was to examine the similarities and differences between the two species in *focusing* attention on emotional versus neutral scenes (Figure 1a, c).

In **Chapter 5**, the dissertation moves on to study *spontaneous mimicry* (Figure 1ii). Through behavioral observations, the existence of contagious self-scratching is examined in orangutans. Moreover, the link between contagious self-scratching, context, and social closeness is studied (Figure 1a-c).

Chapter 6 follows up on this work with an experimental investigation of yawn contagion in orangutans to move towards a better understanding of the function of contagious yawning, as well as its relation with familiarity (Figure 1a, b).

Next, **Chapter 7** focuses on the validation of a touchscreen-based, pictorial adaptation to the classic *Implicit Association Test* to examine *implicit associations* (Figure 1iii, a, b). Though a validation in great apes is outside the scope of this dissertation, the chapter makes a first step towards finding ways to probe implicit associations in animals other than humans.

Finally, in **Chapter 8**, the main findings are summarized and discussed to highlight the most crucial similarities and differences in the cognitive and behavioral mechanisms that underlie emotion processing in hominids.

I would like to emphasize that all the work disseminated in this dissertation is the result of intense collaborations with colleagues. I am the first author of the works described in Chapters 2, 4, 6, and 7, and I am the second author of the works described in Chapters 3 and 5. Chapter 3 is written by Daan W. Laméris, MSc., and Chapter 5 by principal investigator Dr. Mariska E. Kret. Both of these works are important to the dissertation's overarching theme, and although most of the work can be ascribed to Drs. Daan Laméris and Dr. Mariska Kret (as evidenced by their first-authorship), I have made substantial contributions to both studies. With the permission of Daan Laméris, MSc., and Dr. Mariska Kret, Chapters 3 and 5 have therefore become a part of this dissertation.

Part I:

Attention

Chapter 2

Attention towards emotions is modulated by familiarity with the expressor. A comparison between bonobos and humans

Abstract

Why can humans be intolerant of, yet also be empathic towards strangers? This cardinal question can be tackled by studying it in our closest living relatives, bonobos. Their striking xenophilic tendencies make them an interesting model for reconstructing the socio-emotional capacities of the last common ancestor of hominids. Within a series of dot-probe experiments, we compared bonobos' and humans' attention towards scenes depicting familiar (close associates and kin) or unfamiliar individuals with emotional or neutral expressions. Results show that the attention of bonobos is biased towards emotional scenes depicting unfamiliar bonobos, but not by emotional groupmates (Experiment 1) or expressions of humans, irrespective of familiarity (Experiment 2). Using a large community sample, Experiment 3 shows that human attention is biased towards emotional rather than neutral expressions of family and friends. On the one hand, our results show that an attentional bias towards emotions is a shared phenomenon between humans and bonobos, but on the other, that both species have their own unique evolutionarily informed bias. These findings support previously proposed adaptive explanations for xenophilia in bonobos which potentially biases them towards emotional expressions of unfamiliar conspecifics, and parochialism in humans, which makes them sensitive to the emotional expressions of close others.

Based on:
Van Berlo, E., Bionda, T., & Kret. M. E. (2020). Attention towards emotions is modulated by familiarity with the expressor. A comparison between bonobos and humans. *Manuscript submitted for publication*.

Introduction

Emotional expressions are a major force in navigating the social world; they provide valuable insights into the emotional states of others and help to predict others' behaviors (Cosmides & Tooby, 2000). The expression of emotions is not uniquely human and is shared with other animals (Darwin, 1872; Zych & Gogolla, 2021). Yet, we still understand little about how animals perceive and understand others' emotions (Nieuwburg et al., 2021; Paul & Mendl, 2018). Here, taking a comparative perspective will be crucial in elucidating how socio-emotional capacities evolved over time, in ancestral humans as well as other animals. One way to move forward is to compare the emotional processing capacities of humans with those of closely related species.

Within the primate order, bonobos (*Pan paniscus*) are humans' closest living relatives, together with chimpanzees (*Pan troglodytes*). Compared to chimpanzees and other apes, bonobos have strongly developed emotional pathways in the brain (Issa et al., 2019; Stimpson et al., 2016). Behaviorally, bonobos are more tolerant of others and show reduced aggression (Furuichi, 2011; Gruber & Clay, 2016; Hare et al., 2012; Tan & Hare, 2017; Tokuyama et al., 2021). Because of their xenophilic tendencies, bonobos form an interesting comparison species for gaining evolutionary insights into humans' emotional capacities (Gruber & Clay, 2016; Kret et al., 2016; Stimpson et al., 2016). We currently have limited knowledge about how bonobos perceive emotional expressions, and this is a pressing issue given that they are an endangered species (Fruth et al., 2016). Scientific progress is further hampered by bonobos being rare in zoos and sanctuaries (the worldwide zoo-managed population consists of only 225 individuals, managed by the EAZA in Europe and the SPP in the US; Stevens, 2020). Thus, to elucidate the socio-emotional capacities of our shared common ancestor, more comparative studies are needed that include bonobos. We make a step in this direction by investigating selective visual attention for emotions in a comparative framework including bonobos and humans. Specifically, we test whether the identity of the expressor (i.e., a familiar or unfamiliar conspecific) modulates early attention for emotions.

Expressions of emotions facilitate the communication of emotions and intentions between individuals, and are therefore integral to social animals (Prochazkova & Kret, 2017; Zych & Gogolla, 2021). The importance of emotional expressions is reflected in the fact that, over evolutionary time, selective pressures gave rise to brains that are able to quickly attend to and understand emotional expressions (LeDoux, 1998). Research in humans has demonstrated that already during the earliest stages of visual perception, attention is attuned to emotional expressions (Öhman et al., 2001b; Vuilleumier, 2005). Specifically, both threatening and positive signals in the environment can rapidly capture attention (Pool et al., 2016), and this attentional attunement is driven by both arousal-eliciting characteristics of the signal as well as its significance to the observer (Brosch et al., 2008; Frijda, 2017). Interestingly, a similar capacity has been observed in bonobos (Kret et al., 2016). In an experimental setting, bonobos showed an attentional bias towards emotional scenes depicting unfamiliar conspecifics, especially when these scenes were emotionally intense. Moreover, a recent study showed that emotional expressions interfere with attention allocation in bonobos in an emotional Stroop task (Laméris et al., 2022). These findings suggest that the attentional mechanisms that guide social perception have an evolutionarily old foundation, and were likely already present in the last common ancestor of *Pan* and *Homo*.

Aside from being attuned to emotional expressions, the brain systems that facilitate the social bond between individuals have also evolved to prioritize the processing of familiar, socially close others. Human studies have shown that faces of friends and family are detected faster than faces of strangers (Ramon & Gobbini, 2018), and that these familiar faces recruit a broader network of brain areas involved in face, emotional, and social processing (Gobbini et al., 2004). Similarly, a recent study with chimpanzees and bonobos showed that they gaze longer at familiar faces than at unfamiliar faces (Lewis et al., 2021). Familiarity can also affect the expressions of emotions. For example, work on the automatic mimicry of emotional expressions shows that individuals are more likely to mimic expressions of familiar others compared to strangers (Palagi et al., 2020b; Prochazkova & Kret, 2017). As attention gates which signals from the environment are preferentially processed, it is therefore plausible that evolution fine-tuned this mechanism to quickly differentiate not only between emotional and neutral cues, but also between expressions of familiar, socially close group members and unfamiliar others.

Compared to the other great apes and humans, bonobos are strikingly xenophilic. Intergroup encounters in the wild proceed relatively peacefully, and neighboring groups have been observed foraging together (Fruth & Hohmann, 2018). Remarkably, two wild female bonobos have recently been observed adopting an infant from a different social group (Tokuyama et al., 2021). Furthermore, in experimental settings bonobos show a prosocial preference for unfamiliar individuals rather than group members (Tan & Hare, 2013). In contrast, humans tend to prioritize their own group members over unfamiliar individuals when it comes to sharing resources (Fehr et al., 2008). Likely, the environments that both species evolved in contributed to how they interact with others. For bonobos, intergroup tolerance may have resulted from specific ecological conditions, as they live and evolved in a demarcated area in the Democratic Republic of the Congo. Here, reduced feeding competition and environmental stability lead to the formation of stable social parties that prevent extreme territorial encounters with other groups (Hare et al., 2012; Wrangham, 1999). The picture for human evolution is different: ancestral humans migrated great distances across the globe as a result of the extraordinarily volatile climate that caused scarcities in resources for substantial periods of time. This paved the way for intergroup conflicts among our hunter-gatherer ancestors (Ember & Ember, 1992). In turn, these aggressive interactions have fostered a strong focus on the in-group (e.g., family and friends) on the one hand, and xenophobia on the other (Bowles, 2009). Therefore, although humans and bonobos are both highly social animals, their different other-regarding tendencies warrant a closer look at how the two species process emotions of family, friends, and strangers. Specifically, we ask how familiarity impacts early attentional mechanisms that help distinguish between emotionally relevant signals from group members or other, unfamiliar individuals.

To make inter-species comparisons of selective attention for emotions possible, the emotional dot-probe paradigm has been proven useful (MacLeod et al., 1986; Van Rooijen et al., 2017). In the task, individuals have to press a central dot, followed by a short presentation of an emotional and a neutral stimulus. Another dot (the probe) then replaces either the emotional or neutral stimulus. Individuals are generally faster at tapping the probe that replaces the stimulus that biased their attention towards it (usually the emotional stimulus) compared to a probe replacing the other stimulus (the neutral stimulus. See e.g., Belopolsky et al., 2011; Koster et al., 2004 for in-depth discussions on the dot-probe and attentional capture or disengagement). As such, the emotional dot-probe task provides an easy way to tap into the underlying attentional mechanisms that guide emotion perception.

In the current study, we investigate how bonobos and humans attend to expressions of emotion of familiar and unfamiliar individuals. Here, we define familiarity by the social and familial relationship between the observer and the expressor of emotions on the one hand, and unfamiliar others on the other. Further, there is an ongoing debate on the definition of emotions and their expressions (Adolphs et al., 2019; Crivelli & Fridlund, 2018; James, 1884; LeDoux, 2021; Russell & Barrett, 1999; Waller et al., 2020). We here define emotions as adaptive brain states that produce a range of behavioral patterns (expressions) (De Waal, 2011).

Additionally, we define *expressions* of emotions descriptively (Paul & Mendl, 2018) and broadly as visually observable facial and/or body expressions that often occur in social interactions, and that can differ in terms of valence and arousal (Russell, 1980). Based on these definitions, we investigate whether bonobos have an attentional bias towards emotional expressions of unfamiliar and familiar conspecifics (Experiment 1), followed by whether this bias extends to unfamiliar and familiar human expressions (Experiment 2). In Experiment 3, using a large community sample of zoo visitors, we investigate whether attention is attuned to emotional expressions of familiar (family and friends also visiting the zoo) or unfamiliar (other zoo visitors) people.

We hypothesize that bonobos, due to their xenophilic tendencies, will show an attentional bias towards emotions expressed by unfamiliar conspecifics (Kret et al., 2016) and that a similar bias will be dampened when seeing familiar conspecifics. Furthermore, since certain aspects of emotion processing are shared between humans and extant apes (Kret et al., 2020), we further predict that bonobos will show an attentional bias towards emotional expressions of humans. Whether this bias is modulated by the familiarity of the human expressor is an exploratory question. For humans, we hypothesize that an attentional bias towards emotions exists for expressions of unfamiliar individuals, in line with established findings (Van Rooijen et al., 2017). We also expect that this bias will be more pronounced for familiar individuals as compared to unfamiliar individuals, reflecting the more in-group focused, parochial tendencies of humans (Bowles, 2009).

Experiment 1: Bonobos' attentional bias towards emotions of familiar and unfamiliar conspecifics

Method

Participants

Four female bonobos (Besede, 11 years old; Monyama, 6 years old; Kumbuka, 17 years old; Yahimba, 7 years old and daughter of Kumbuka)[1] living in a social group of 12 individuals at Apenheul primate park in Apeldoorn, The Netherlands, took part in the study and were tested over a period of 4.5 months.

All participating females were born in captivity and had prior touchscreen experience through participating in the study by Kret et al. (2016). There were eight months of no testing between the two studies. At the time of testing, none of the individuals were pregnant nor on contraceptives. All individuals were housed in large in- and outdoor enclosures (2970 m^2 in total) containing several climbing structures, trees, bushes and ropes, puzzles from which they could acquire food, and small streams of water. To mimic natural fission-fusion behavior, bonobos were always housed in two separated groups that varied in composition regularly. All participants in this study were exposed to humans since birth and interacted with them on a daily basis. Daily diet consisted of a variety of fruits, vegetables, branches and leaves, and pellets enriched with necessary nutrients. The bonobos were fed four to five times a day, and water was available ad libitum. Furthermore, bonobos were never deprived of water or food at any stage of the experiment.

Testing took place in the presence of non-participating group members and during winter when the park was closed for visitors. Bonobos were tested three to four times per week in one of the indoor enclosures, and one test session lasted ~15-20 minutes per individual.

Tests with the bonobos were conducted adhering to the guidelines of the EAZA Ex situ Program (EEP), formulated by the European Association of Zoos and Aquaria (EAZA). Bonobos participated voluntarily and were never separated from their group

---

[1] We acknowledge that our sample size is limited compared to studies with humans. Nevertheless, it is in line with touchscreen-based experiments involving apes, which have an average sample size of four (Egelkamp & Ross, 2019). Despite this limitation, we argue that comparative studies such as ours have scientific merit and provide crucial insights into the cognitive abilities of animals. This is especially true for bonobos, as they are a critically endangered species and rarely kept in zoos and sanctuaries (Fruth et al., 2016). Access to and testing of bonobos is very limited. To partly compensate for the low sample size, we maximize the number of trials per individual and make individuals' data available for future work.

during testing. Only positive reinforcements (apple cubes) were used during training and testing, and each bonobo (including ones that were not tested) received a reward equivalent to the reward of the bonobo being tested. Non-participating bonobos were distracted by the animal keeper who conducted a body-part training task used for veterinary purposes.

Equipment

The experiment was conducted using *Presentation* (NeuroBehavioralSystems) on an Iiyama T1931SR-B1 touchscreen (19", 1280x1024 pixels, ISO 5ms) encased in a custom-made setup (Figure 1). To limit exposure to the experimenter, rewards for correct responses were automatically distributed using a custom-made auto-feeder apparatus that dropped apple cubes into a funnel that ended underneath the touchscreen for the bonobo to grab. A camera was placed outside the enclosure to film the bonobos while performing in the experiment.

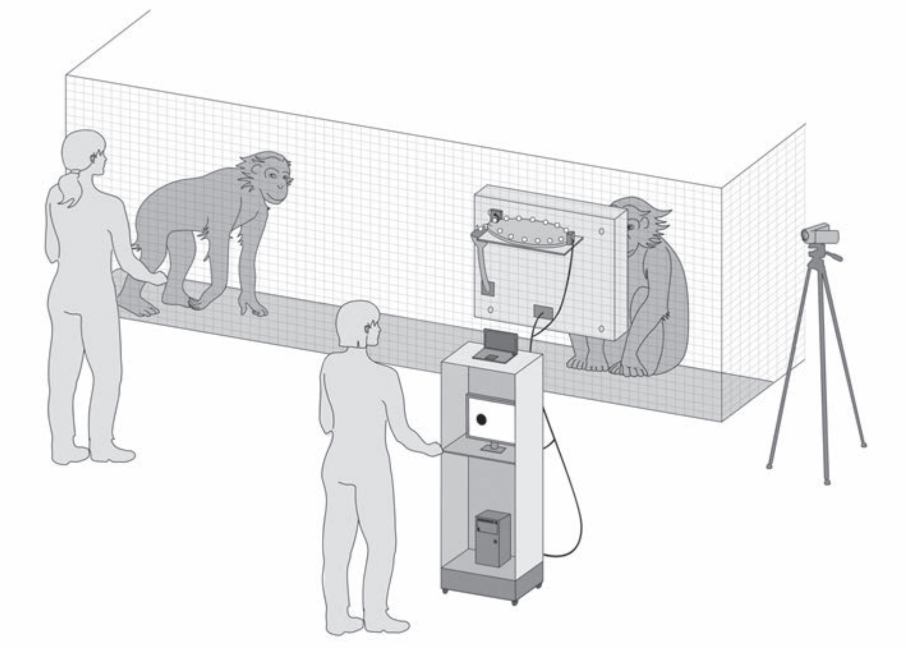

Figure 1. Abstract representation of the bonobo setup. The experimenter (right) controlled the experiment from behind the bonobo setup while a keeper (left) distracted the other bonobos. The experimenter was not visible and remained silent to the bonobos most of the time, but the experimenter would move to the side of the setup when an individual needed some encouragement to continue with the task. At the end of the task, the experimenter and caretaker would say "good job" to the participating bonobo to indicate the bonobo was done with the experiment for this day.

Stimuli and validation

Stimuli consisted of bonobo pictures collected in different zoos and from the internet. Stimuli of familiar individuals consisted of pictures of the group living in Apenheul, and unfamiliar stimuli depicted a small selection of individuals from five different zoos (Cologne, Planckendael, Twycross, Cincinatti, and San Diego Zoo). We only included clear pictures in our sample (i.e., no pixelations, adequate lighting). In total, the study included 656 novel and unique pictures (346 of familiar and 310 pictures of unfamiliar individuals). All pictures were resized to 330 x 400 pixels and showed either a neutral scene (i.e., individuals sitting or lying down or involved in a non-social activity, showing a neutral expression) or an emotional scene.

While we currently do not fully understand what bonobo emotions entail, we rely on existing observational work to establish relevant socio-emotional behaviors and expressions that may underlie emotional states. Here, the valence-arousal model by Russel (1980) can be used as a useful guideline. We use socio-emotional scenes of bonobos engaged in play, grooming, or sex (positively valenced), and bonobos showing distress or that were self-scratching (negatively valenced), or yawning (unclear valence) as proxies of emotional states (see Figure 2 and, in the supplements, Table S1). We used similar emotion categories as Kret et al. (2016) (but all novel images), with the exception that we included self-scratching as a new category and left out pant hoot and food, because these did not attract attention over neutral scenes in our previous study.

Play, grooming, and sex are important for establishing or maintaining social bonds (Moscovice et al., 2019; Palagi, 2008; Schroepfer-Walker et al., 2015), and may therefore reflect positively valenced behaviors (Furuichi, 2011). Play scenes involved playful interactions between two bonobos, or an individual playing with objects, and included the relaxed open-mouth (‘play face’) expression (Signe & Van Hooff, 2018). Grooming scenes involved grooming bouts between two or more individuals. Furthermore, sexual scenes displayed two or more individuals copulating, or showing an erection (males) or large genital swelling (females). Scenes showing distressed bonobos included one or more individuals displaying a fear grin that is typically produced by primates in distress (De Waal, 1988; Parr et al., 2007). Self-scratching scenes displayed one or two individuals scratching themselves on the head or body. Self-scratching is indicative of stress in both primates and humans (Troisi, 2002), and by incorporating it as an emotional stimulus, we increased the number of negatively valenced stimuli. Finally, yawning scenes showed one individual with an open mouth, with or without teeth exposure. It is unclear what emotional state may underlie yawning (e.g., boredom (Burn, 2017) or stress (Maestripieri et al., 1992; Paukner & Anderson, 2006)), but it is a highly contagious behavior that could be a proxy for empathy (but see Massen & Gallup, 2017). Moreover, bonobos responded faster to probes replacing yawning stimuli than other categories in the study by Kret et al. (2016), and therefore we included it in our study.

We matched emotional and neutral scenes on the number of individuals depicted (ranging from one to six), their identity, and by visual inspection of color and luminance. All 12 bonobos in the Apenheul group were present in the familiar stimulus set, and we estimate the presence of 30 unique individuals in the unfamiliar stimuli. Furthermore, the pictures were cropped in such a way that the bonobos' faces and/or bodies covered most of the stimulus area. Backgrounds of the stimuli either showed a bit of grass or part of a tree, or, when the stimuli were of individuals in their inside enclosure, of a white-grey wall and sometimes a beam (part of the inside construction). All pictures were rated on emotional valence and intensity (arousal) by three primate experts from Apenheul and three primate researchers, who showed high intraclass correlations ($ICC_{valence} = .82$, $ICC_{intensity} = .87$, supplements, Table S2).

Figure 2. Examples of stimuli of all emotional categories used in Experiment 1. An emotional picture was always paired with a neutral picture. The emotional and neutral pictures were of either familiar or unfamiliar individuals.

Procedure

The bonobos were already familiarized with the dot-probe procedure during a previous study (Kret et al., 2016), but did go through a short refamiliarization phase (about 7 sessions per ape spaced across a 2-month period). During this phase, bonobos performed a dot-probe task with pictures of black rabbits and goats. Only after all the apes were able to correctly pass 95% or more of the trials within one session, we moved on to the experiment. The criterion of a successful trial was to immediately press the probe while attending to the screen.

To start a training or experimental session, we called forth the highest-ranking participating individual of the subgroup that was present in the enclosure. In the experiment, a trial started with the presentation of the start dot in the middle, lower part of the screen (Figure 3). After the bonobo pressed the dot, a neutral and an emotional stimulus appeared on the left and right side of the screen for 300 ms (Kret et al., 2016, 2018; Petrova et al., 2013). Stimuli were always either of bonobos familiar to the participant or of unfamiliar individuals (thus, we never combined an emotional picture of a familiar with a neutral picture of an unfamiliar or vice versa). Stimuli were subsequently followed by another dot (the probe) replacing either the neutral or emotional stimulus. The probe remained on the screen until touched, after which an apple cube was provided through the auto-feeder system. After a delay of 2000 ms the next trial started. Each test session consisted of 25 trials in which the location of the stimuli on the screen (left/right) and the location of the probe (behind the emotional or neutral stimulus) were counterbalanced, and the order of stimulus presentation was randomized based on emotion category and familiarity. In each session, half of the trials consisted of emotional and neutral stimuli of familiar individuals, and half of emotional and neutral stimuli of unfamiliar individuals. If a trial was deemed unsuccessful, it was repeated at the end of the study. In total, each bonobo finished between 21 to 24 sessions and on average a total of 541 trials (*SD* = 28.76, Table S3).

Non-participating bonobos were distracted by the animal caretaker with a body-part training in which bonobos were instructed to present specific body parts to the animal caretaker, and were rewarded with an apple cube for each correct presentation, just like the participating bonobos when they completed a trial. Importantly, bonobos were never separated from their group members, thus sometimes leading to disruptions during the experiment. From the recorded videos, two experts coded unsuccessful trials by looking at the following events: bonobos were distracted by other bonobos or did not attend to the screen, another individual pressed the probe, hands were switched within a trial, or bonobos performed movements that interfered with the task (self-scratching or nose wiping). The experts showed high agreement in coding (ICC = .95, $p < .001$).

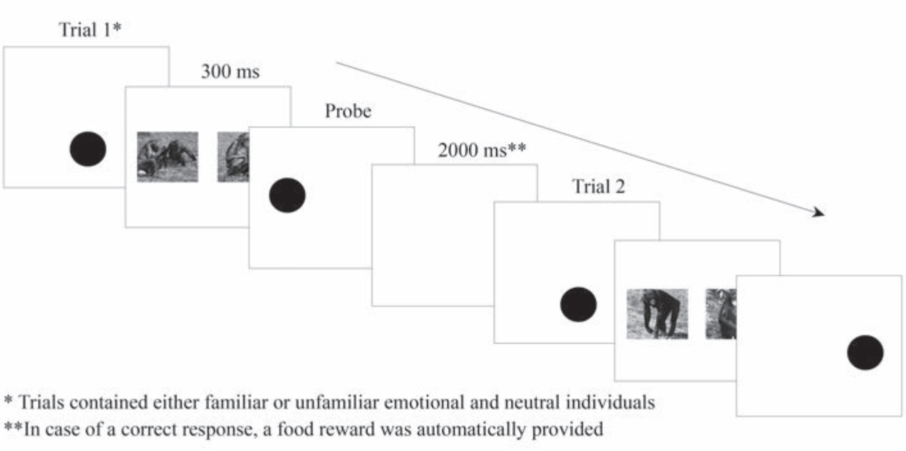

Figure 3. Trial outline of the bonobo dot-probe task.

Data filtering

Based on the coding of the two experts, erroneous trials were discarded. Moreover, extreme reaction times (RT < 250 ms and RT > 5000 ms) were filtered out. Finally, trials with RTs higher than the median RT per subject minus 2.5 * the median absolute deviation per subject (MAD) were excluded. Based on these criteria, 514 trials (23.8%) were removed from the analysis (The majority of these invalid trials (90%) were caused by bonobos being distracted or other individuals interfering in the task. See also Table S3). Therefore, we had a final number of 1650 datapoints (~413 per condition). This is less than has recently been recommended for performing mixed model analyses (Brysbaert & Stevens, 2018), but is in line with most other dot-probe studies (Van Rooijen et al., 2017).

Statistical analyses

We used a generalized linear mixed model (v1.4.1106, glmmTMB package, $\alpha = .05$ (Brooks et al., 2017; R Core Team, 2020)) for the analyses, with a nested structure defined by trials (25) nested within sessions (21-24) nested within participants (ID, 4). We included *Congruency* (the probe replaced an emotional [congruent] or neutral [incongruent] stimulus, sum coded) and *Familiarity* (familiar versus unfamiliar bonobos, sum coded), and their interaction terms as fixed factors, and used random intercepts per *ID* and *ID*Session*. Reaction time was used as the dependent variable. To determine which distribution family provided the closest fit to the observed data, we compared AIC statistics of models with a normal and gamma distribution (Lo &

Andrews, 2015). Model assumptions were checked by visually inspecting QQ plots and the residuals plotted against fitted values.

Results

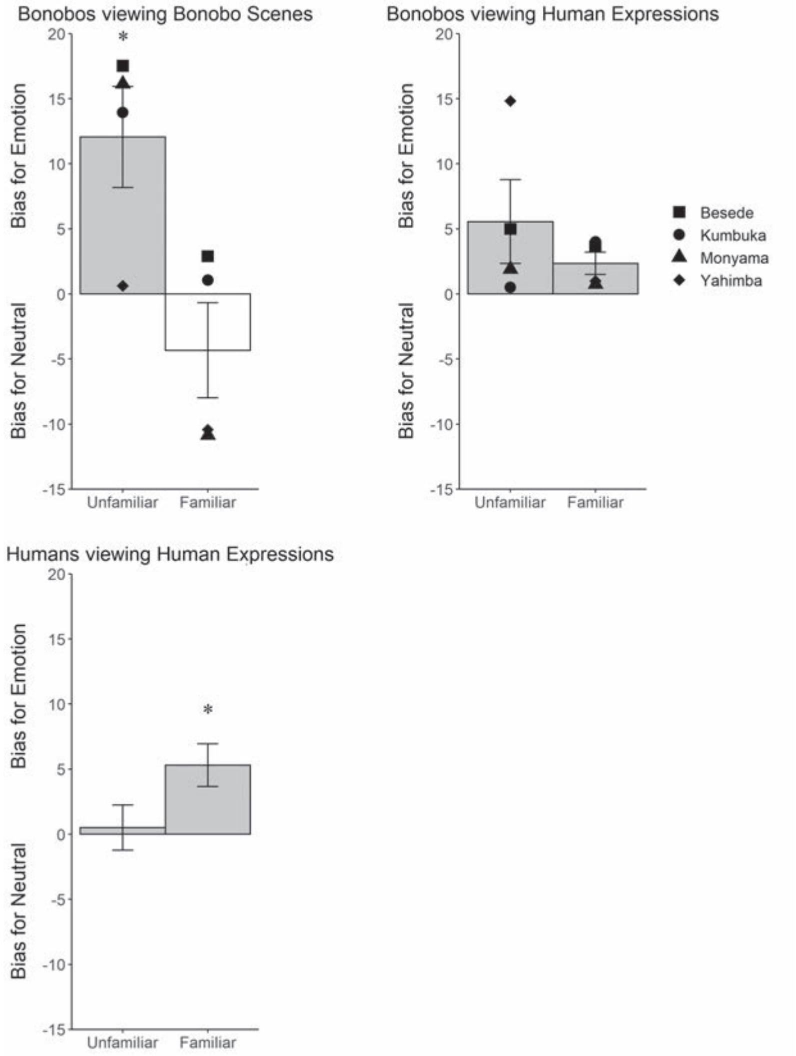

Figure 4. Experiment 1: Bonobos show an attentional bias towards emotions of unfamiliar, but not familiar conspecifics (top left). Experiment 2: Bonobos do not show an attentional bias towards emotions of familiar or unfamiliar humans (top right). Experiment 3: Humans have an attentional bias towards emotional expressions of familiar others (bottom left). To illustrate an attentional bias, we calculated the difference between mean reaction times (RTs) on neutral scenes minus mean RTs on emotional scenes per condition (Unfamiliar, Familiar). Bars in the positive direction indicate a bias towards emotional scenes or expressions rather than to neutral scenes or expressions. Error bars represent the SEM. * $p < .05$.

We aimed to replicate and extend previous findings by Kret et al. (2016) and tested for a possible interaction between familiarity and emotional attention in bonobos. When comparing the AIC statistics of a normal and gamma distribution ($AIC_{normal} = 18949$, $AIC_{gamma} = 18995$), the model with a normal distribution was found to be a better fit. We found a significant interaction effect between *Familiarity* and *Congruency* ($\chi^2(1) = 4.14$, $p = .042$); bonobos responded faster on probes replacing emotional ($M = 521.11$, $SD = 131.50$) rather than neutral scenes ($M = 529.84$, $SD = 127.13$) in the *Unfamiliar* condition ($\beta = -10.48$, $SE = 5.12$, $t(1641) = -2.05$, $p = .041$) but not in the *Familiar* condition ($\beta = 4.59$, $SE = 5.34$, $t(1641) = .86$, $p = .391$, see Figure 4 and Table S4.1 and S4.2 for individual averages and further model results). In short, familiarity with the expressor of an emotion significantly modulated attentional bias towards emotions, with responses to emotional scenes being faster than neutral scenes when they involved unfamiliar, but not familiar conspecifics.

Discussion

Previous research has shown that bonobos have heightened attention to the emotional expressions of unfamiliar conspecifics, especially when these were rated as emotionally intense by their keepers (Kret et al., 2016). The current study builds on this research. Specifically, by adding photographs of group mates to the stimulus materials, Experiment 1 showed that familiarity with the expressor has a moderating effect on an attentional bias towards emotions; early attention appears to be modulated mostly by emotional expressions of unfamiliar individuals, but not familiar individuals. From a human perspective, this finding may appear counter-intuitive. However, this novel finding largely confirms our à priori predictions which were based on previously conducted behavioral studies in bonobos highlighting their strong xenophilic tendencies and other-regarding preferences (Fruth & Hohmann, 2018; Tan et al., 2017; Tan & Hare, 2013; Tokuyama et al., 2021). Attention can be driven by the biological relevance of the emotional signal to the observer, for instance by the presence of dangerous animals such as snakes (Öhman et al., 2001a). It is thought that for bonobos, the benefits of bonding with new individuals outweigh the costs, thereby making socializing with unfamiliar conspecifics beneficial (Hare et al., 2012). In turn, these interactions may enhance survival by promoting cooperation among individuals (Tan et al., 2017). Though we have to be careful with generalizations given our sample size, our results appear to support this notion and suggest that the brains of bonobos developed to selectively attend to emotional signals from potentially interesting unfamiliar social partners.

At the same time, it is interesting that there is no effect of emotion in the familiar condition. A recent eye-tracking study by Lewis et al. (2021) showed that bonobos attended longer to familiar group members rather than unfamiliar bonobos, indicating that seeing familiar individuals somehow interests the bonobos. It is possible that when viewing familiar individuals, the effect of emotional expressions on attention is further affected by pre-existing knowledge about those individuals. Other research indeed suggests that social characteristics of the observer in relation to the observed individual(s) may play a role in how emotions are processed. For instance, attention has been shown to be modulated by e.g., sex (Schino et al., 2020), social bond (Kutsukake, 2006; Whitehouse et al., 2016), rank (Lewis et al., 2021; Micheletta et al., 2015; Schino & Sciarretta, 2016), and kinship (Schino & Sciarretta, 2016). The current study sample did not allow us to disentangle potential effects of social characteristics on an attentional bias towards emotions. However, inspection of the two bars representing the familiar and unfamiliar condition in the top left plot of Figure 4 suggests that the inter-individual variance was comparable between these two conditions. Another possibility for why an attentional bias towards the emotional expressions of familiar conspecifics was not observed may be related to the fact that familiar and unfamiliar conspecifics were shown within the same experiment (and not within the same trial). The emotional expressions of unfamiliar conspecifics may be of such high relevance for this species, that it rendered biases towards expressions of close associates and kin insignificant. We cannot test this in our data, but future work could try to zoom in on how attention to emotions is modulated by specific characteristics of familiar individuals (e.g., age, relationship, rank).

An alternative explanation for our findings is that results are driven by heightened novelty of the unfamiliar stimuli (Bradley, 2009). However, we could rule this out, because bonobos on average responded as fast to stimuli of unfamiliar (novel) as of familiar individuals. A worthwhile follow-up experiment is to directly compare familiar and unfamiliar individuals (emotional and neutral) within trials in order to disentangle effects of emotion and familiarity. In addition, studying an attentional bias towards emotions of familiar and unfamiliar individuals in chimpanzees could be a fruitful next step. While chimpanzees and bonobos are very closely related to each other and equally closely related to humans, differences in social organization (with females being dominant in bonobos, and males in chimpanzees), and social tolerance (chimpanzees are highly territorial) may also differentially affect where attention is allocated. Indeed, Lewis et al. (2021) showed that bonobos preferentially attend to familiar, high-ranking females rather than unfamiliar females, whereas chimpanzees attend to familiar high-ranking males. We believe the study by Lewis et al. (2021) and ours complement each other in showing that at least for bonobos, seeing familiar individuals brings along a range of potentially relevant social information such as rank and emotional expressions that in turn may modulate attention differently. Therefore, we believe an interesting next step would be to study more closely how familiarity with the expressor modulates attention for emotions across species.

Experiment 2: Bonobos' attentional bias towards emotions of familiar and unfamiliar humans

Method

Participants and equipment

The bonobos taking part in Experiment 2 were the same as in Experiment 1. Experiment 2 followed directly after Experiment 1, and used the same setup.

Stimuli and validation

Stimuli consisted of isolated emotional and neutral human faces that were either unfamiliar to the bonobos (NimStim Set of Facial Expressions (Tottenham et al., 2009)) or familiar (4 female bonobo animal caretakers that interact with the bonobos on a daily to weekly basis, and 2 female experimenters that trained and tested the bonobos in the past). Emotional expressions consisted of six basic human expressions (Ekman, 1999): anger, fear, happiness, sadness, surprise, and disgust. (Figure 5). Stimuli were in color and sized 330x400 pixels. In total we had 144 stimuli (72 of familiar and 72 of unfamiliar individuals).

While making the photos of the caretakers, the experimenter enacted the facial expressions as an example for the caretakers and instructed them to mimic her. Photos of emotional expressions were taken in the following order: anger, fear, happiness, sadness, surprise, disgust. Each photo of an emotional expression was followed by a photo of a neutral expression to ensure that neutral photos were slightly different from each other. If the experimenter thought a photo was not similar enough to the faces used from the NimStim database, the photo was retaken. Stimuli from familiar individuals were all unique. However, because the NimStim database sometimes contained only two different neutral expressions per model, for the unfamiliar stimuli only the emotional pictures were all unique.

To check the validity of our stimulus materials, we first asked an independent group of research assistants (*N* = 5) to rate the materials on emotion type (whether the stimulus is an emotional or neutral expression), arousal and authenticity. Results indicated the following intraclass correlations: intensity of the stimuli (ICC = .78), emotion (ICC = .66), and authenticity (ICC = .69; see Tables S5.1 and S5.2).

Figure 5. Examples of the stimuli of emotional and neutral expressions. The NimStim models (not depicted here) were females 1, 2, 3, and 7, 8, 9 (Tottenham et al., 2009). An emotional picture was always paired with a neutral picture, and these pictures were always either from the NimStim or Keeper stimulus set.

Procedure

The procedure for bonobos in Experiment 2 was similar to Experiment 1, except that the stimuli in Experiment 2 consisted of the six human basic emotional facial expressions (Figure 3). The reason for using facial expressions rather than scenes was based on two considerations. First, previous dot-probe studies in humans have mainly used facial expressions to examine attentional bias (Van Rooijen et al., 2017), thus allowing us to compare the results of Experiment 2 to existing findings. Second, it would not have been possible to get some of the emotional scenes (e.g., play, grooming, sex) from the familiar people. On average, each bonobo finished 345 trials (*SD* = 24.56), divided over 13-15 sessions per individual (Table S6).

Data Filtering

As in Experiment 1, two experts rated the videos in high agreement (ICC = .96, $p <$ .001). We used the same data filtering criteria as in Experiment 1, resulting in removal of 373 trials (27.1%, Table S6).

Statistical analyses

Similar to Experiment 1, we used a GLMM with a nested structure with trials (25) nested within sessions (13-15) nested within participants (ID, 4) and random intercepts per *ID* and per *ID*Session*. The dependent variable was reaction time (ms), and we included *Congruency* (sum coded), *Familiarity* (familiar versus unfamiliar human model, sum coded) and their interaction terms as fixed factors. We checked which distribution family (gaussian vs. gamma distribution) fit best the data based on the AIC statistic. We checked the assumptions of our models by visually inspecting QQ plots and the residuals plotted against fitted values.

Results

The AIC statistics revealed a better fit for a model with a normal distribution rather than a gamma distribution ($AIC_{normal} = 11701$, $AIC_{gamma} = 11756$). The model did not show a significant main effect for *Congruency* ($\chi^2(1) = .33$, $p = .567$), nor for *Familiarity* ($\chi^2(1) = .04$, $p = .840$), nor an interaction effect between *Congruency* and *Familiarity* ($\chi^2(1) = .16$, $p = .693$. See Figure 4, top right. Also see supplements Tables S7.1 and S7.2 for individual averages and further model output).

To substantiate our null-finding, we conducted a Bayesian control analysis. Bayesian analyses have been proposed as a reliable way to establish the strength of evidence against the null-hypothesis when frequentist analyses show non-significant results (Rouder et al., 2009). Specifically, calculating a Bayes Factor (BF) can assist in examining evidence for the null-hypothesis, which is not possible within the frequentist framework (Kass & Raftery, 1995). To do so, we fitted a Bayesian mixed (Gaussian) model using the brms package in R (Bürkner, 2017, 2018). In the model, *Congruency*, *Familiarity* and their interaction were defined as fixed factors, with reaction time as dependent variable. *Congruency* and *Familiarity* were sum-coded, and we included a nested random intercept (with *Session* nested within *Participant*). Priors consisted of a weakly informative Gaussian prior for the intercept (M = 500, SD = 100) and a more conservative Gaussian prior the fixed effect (M = 0, SD = 10). For the random effect and residual standard deviation, we used the default half Student-*t* priors (with 3 df). We also ran a null model that included the same parameters, excluding the fixed factors and their interactions. For each model, we ran four chains with 4000 iterations (of which 2000 iterations were warmups). Model validity was established by following the WAMBS checklist (Depaoli & van de Schoot, 2017), including trace plots, histograms of the posteriors, Gelman-Rubin diagnostics, and autocorrelation checks. We then calculated an average Bayes Factor_{01} using 1000 iterations, and found that the mean $BF_{01} = 61.07$ ($SD = 16.06$), indicating very strong evidence for the null-hypothesis (Jeffreys, 1961). Thus, given our data, we found no evidence that bonobos have an attentional bias towards human facial expressions of emotion.

To verify whether our null-findings could be due to any inherent qualities of the used stimuli, we conducted a control experiment with a new group of zoo visitors and using the same stimuli as the ones presented to the bonobos, here. Crucially, the visitors had no prior experience with the individuals on the stimuli. If we find an attentional bias towards emotions using this stimulus sample, then we can at least say that the stimuli are salient enough to elicit an attentional bias in humans. In total, we tested $N = 150$ zoo visitors (75 men and 75 women, 18-88 years old, $M_{age} = 39.79$, $SD = 14.98$). Note that for clarity, we stick to using 'familiar' and 'unfamiliar' to denote the two stimulus sets (Caretaker vs. NimStim), but keep in mind that for the participants, none of the stimuli were familiar. We created two versions of the task each containing 72 trials with 36 trials of 'familiar' humans and 36 'unfamiliar' humans. The only difference between the two versions was that the probe location was mirrored (i.e., if in version 1 it appeared behind one of the emotional pictures, it appeared behind the neutral stimulus in version 2, and vice versa). Per participant, every stimulus was only shown once. To filter our data, we first divided every participant into an age category as reaction times can be higher in older versus younger individuals (i.e., 18-20, 21-25, 26-30, etc.). Next, we filtered out extreme RTs (RT < 250 ms and RT >5000 ms) and then calculated the median absolute deviation for reaction times per age category. Finally, we used the following data filter: [RT < (Median RT + (2.5 * Mean Absolute Deviation))]. 606 Trials (5.61%) were subsequently removed for further analysis.

We performed a GLMM with *Congruency*, *Familiarity*, and their interaction, with random intercepts per ID, and using a gamma distribution with a log-link function (as $AIC_{normal} = 111557$, $AIC_{gamma} = 111624$). We found the expected main effect of *Congruency* ($\chi^2(1) = 4.00$, $p = .046$) and, importantly, neither an effect of *Familiarity* ($\chi^2(1) = .01$, $p = .931$) nor an interaction between the two ($\chi^2(1) = .03$, $p = .866$). As such, participants had a faster reaction time to a probe replacing an emotional stimulus ($M = 434.87$, $SD = 85.41$) than to a probe replacing a neutral stimulus ($M = 437.72$, $SD = 84.87$, $\beta = -0.007$ (log scale), $t = -2.00$, $p = .046$. See Table S8 for further model output), irrespective of the stimulus set. This is important because it shows that the expressions of the caretakers attracted as much attention as the ones from the NimStim set, thus, the null-result in bonobos is unlikely to be attributable to any qualitative characteristics of the stimuli used, at least to the human eye.

Discussion

Bonobos did not show an attentional bias towards human facial expressions of emotion. When repeating this experiment in human participants, an attentional bias towards emotional expressions was observed. The result for bonobos was unexpected for several reasons. First, great apes and humans show some continuity in facial expressions of emotions in both morphology and function, making it likely that modulation of attention by emotions stretches beyond conspecifics. Indeed, the facial musculature of chimpanzees and humans are remarkably similar (Burrows et al., 2006), paving the way for potential functional similarities in emotional expressions. The starkest examples of how expressions are shared is the human smile and the equivalent bared-teeth display in apes, as well as human laughter and the relaxed open mouth ('play face') in apes (Parr et al., 2007; Van Hooff, 1972). Similarly, the expression of anger in humans is suggested to be equivalent to the bulging-lip display in chimpanzees and the tense 'lip press' in bonobos (De Waal, 1988; Parr et al., 2007). Moreover, some facial configurations of emotions appear only in humans (fear), or have not yet been studied in detail in apes (surprise, sadness). Overall, given the continuity between (some) expressions, it may have been plausible for bonobos to show a bias towards emotional expressions of humans.

A second reason for why we believe the result is unexpected is that there is some evidence that apes view emotional expressions of humans similarly as those of conspecifics. A study in orangutans showed that when presented with isolated facial expressions of humans and orangutans, they generally looked longer to emotional expressions as compared to neutral expressions, regardless of whether the expressions were of humans or conspecifics (Pritsch et al., 2017). Moreover, one study showed that apes ($N = 32$) have some understanding of the directedness and valence of human emotional expressions and use these expressions to infer desires (Buttelmann et al., 2009). Furthermore, Kano and Tomonaga (2010) examined how chimpanzees and humans view isolated neutral and emotional faces, and found that both species show similar facial scanning patterns regardless of whether the stimuli were of conspecifics or non-conspecifics. Crucially, in chimpanzees, scanning patterns changed according to the emotional expressions that were being viewed, but patterns were relatively similar across faces of humans and chimpanzees. Moreover, there is experimental evidence that shows that great apes can understand humans' emotional facial expressions to some extent, for instance to infer desires (Buttelmann et al., 2009). These findings suggest that great apes are sensitive to emotional expressions of another, phylogenetically close species. Possibly, differences in methodology may explain why we did not find an emotion bias. The dot-probe has very short exposure times, likely tapping into an early attentional process (MacLeod et al., 1986), whereas longer stimulus exposure (such as in the studies by Kano & Tomonaga (2010b) and Pritsch et al. (2017)) can provide additional information to the observer and recruit prior knowledge and experiences to process what is seen.

Finally, although the expressions were salient enough for humans to reveal an attentional bias towards these expressions in our control experiment, they may simply not have been salient for bonobos. This is difficult to quantify; though the stimuli were rated on valence and intensity, they were of course rated by humans and not bonobos. Moreover, our data may not have had sufficient power to detect an effect. It is therefore difficult to draw a definitive conclusion on attentional biases for emotions of humans. A potential future direction could be to test immediate and sustained attention for familiar and unfamiliar humans, and subsequently test a possible interaction with emotions. Furthermore, a matching-to-sample task could be useful to study whether bonobos can distinguish between human expressions of emotion.

Experiment 3: Humans' attentional bias towards emotions of familiar and unfamiliar conspecifics

Method

Participants

We recruited pairs of individuals to be either taking part in the dot-probe task or to be on the photographs used for the *familiar* stimuli. Participants thus consisted of those partaking in the dot-probe task (*N* = 449, 253 women), or were the to-be photographed companions (*N* = 406, 208 men. For 43 companions, data on sex and age were missing due to a technical malfunction). Participants were adults and children (Dot probe participants: 262 adults, 187 children. Companions for the photos: 218 adults, 188 children) visiting Apenheul. Dot probe participants were between 3 and 84 years old (*M* = 24.9, *SD* = 16.43), and companions were between the ages 3 to 79 (*M* = 25.43, *SD* = 17.22). Apenheul allowed us to set up a research corner close to the bonobo enclosure where we could test the visitors (Figure 6). As bonobos were only exposed to group members (kin and friends) in *familiar* trials, our human participants were also selected based on their relationship with their companions (kin and friends or colleagues. See Table S9 for an overview on relationships). The experiment was approved by the ethics committee of Leiden University (adults: CEP17-0213/74, children: CEP17-0604/222).

Equipment

Participants performed the experiment on an Iiyama T1931SR-B1 touchscreen (19", 1280x1024 pixels, ISO 5ms) using E-Prime 2.0. The tests were conducted in an indoor compound in which visitors could see the bonobos. The touchscreen was placed on a table and participants were seated with their back against a wall to prevent others from distracting them (Figure 6).

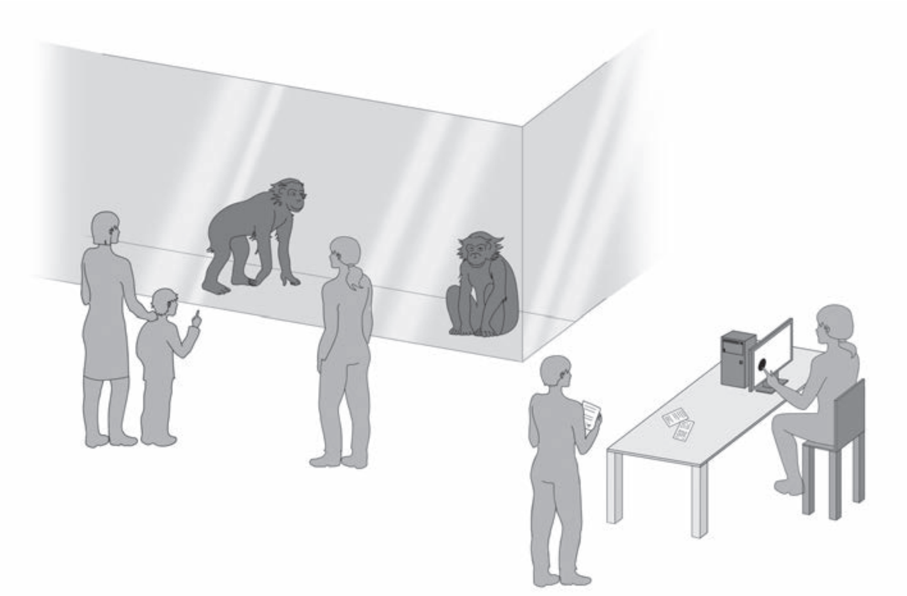

Figure 6. Abstract representation of the human setup near the bonobo enclosure.

Stimuli and validation

For our human participants, rather than emotional scenes, we used isolated facial expressions as stimuli. This was done for practical reasons, as we could not ask zoo visitors to enact specific social scenes like the bonobo scenes. For humans, there is evidence for an attentional bias towards emotions using only facial expressions (Van Rooijen et al., 2017), but also for emotional scenes (Kret & Van Berlo, 2021). Similarly, while a dot-probe study with bonobos found an attentional bias for emotional scenes (Kret et al., 2016), other studies found that isolated facial expressions can also modulate attention in apes (Laméris et al., 2022; Pritsch et al., 2017). As such, we did not expect the discrepancy between the bonobo stimuli and human stimuli to significantly impact outcomes of both experiments.

Stimuli consisted of pictures of the face showing either an emotional (angry, fearful, happy, sad) or neutral expression presented against a neutral background, similar to the expressions depicted in Figure 5. Each stimulus showed either a familiar companion (a family member, a close friend, or a colleague), or an unfamiliar individual (a companion of a previous, unfamiliar participant). For practical reasons, we only used four out of the six basic emotions (Ekman, 1999), as the task would become undesirably long given that our participants were voluntarily taking part in our study. Pictures were sized 400x300 pixels.

Each participant completed 40 trials. In 20 trials, the probe appeared behind an emotional stimulus, and in the other 20 trials it appeared behind a neutral stimulus. For each of these 20 trials, 10 trials depicted a familiar individual, and 10 trials an unfamiliar individual. Since we only had eight unique photos of a familiar other as well as eight unique photos of a stranger, we repeated two stimulus pairs within each condition to reach the maximum of 40 trials. The number of stimuli per emotional category was counter-balanced across participants (including the repetitions), and stimulus combinations (emotional plus neutral) were presented in a semi-randomized order. A total of 4040 pictures were split into three sets and rated on intensity, emotionality (whether a stimulus depicts an emotional or neutral expression), and authenticity by 18 university graduates and PhD candidates, and on average there was good agreement (ICCintensity = .80, ICCemotion = .80, and ICCauthenticity = .68, see Tables S10.1 and S10.2).

Procedure

Visitors passing by the bonobo enclosure with at least one other person were approached by test leaders. Visitors were told about the ongoing research with the bonobos, and were asked if they wanted to perform in a similar task. If they wanted to participate, the experimenter decided which participant was going to perform the dot-probe task (‘dot-probe participant’) and who was going to be on the photos that subsequently served as stimulus material (‘photo participant’). Individuals could only participate in the study once (and either as dot-probe or photo participant). After reading the information sheet and signing a consent form, photos were made of the photo participant and outside of the view of the dot-probe participant. Photos were taken on the same spot near the test location, around the corner and against a background with a brick wall. The participant was asked to express one of each of the four emotions (angry, fearful, happy, sad), based on an example from Model 1 from the NimStim database (Tottenham et al., 2009) printed on a sheet of paper. After each emotional expression, the participant was asked to show a neutral expression. This prevented the use of the same neutral photo across all trials, and ensured some variation in the neutral expressions. As such, eight unique photos were taken (four emotional, four neutral). Low quality photos were retaken on the spot.

Next, the pictures were loaded into the software (which automatically handled the resizing of stimuli) and the dot-probe participant was then seated behind the touchscreen. The experimenter entered personal data (age, handedness, sex of both the dot-probe and photo participant, the nature of their relationship, and how often they see each other (Table S9). The instructions for the task were kept to a minimum, as the bonobos could also not receive written or verbal instructions.

The experimenter told the participant that they would see a dot appear on the screen, and that they would have to touch the dot as soon as it appeared by using only their dominant hand. The task started with the pictures of the four bonobos participating in Experiment 1, with the text “Are you faster than the bonobos? Press anywhere on the screen to continue”. After three practice trials that included pictures of flowers, the participants saw a display asking them whether they understood the task and were reminded of using only one hand during the task. Participants could then move on to the experiment by pressing the red dot on the screen. The experimental procedure was similar to that in Experiments 1 and 2 with bonobos. A trial started with a start dot in the lower middle part of the screen. Upon touching the start dot, two stimuli were presented side by side for 300 ms. These stimuli were an emotional and neutral stimulus of either a familiar companion or an unfamiliar individual. A probe then replaced one of the stimuli. After touching the probe, a blank screen was shown for 2000 ms, after which the next trial would start. After completing all 40 trials, participants would see a screen depicting their average reaction time and how it compared to that of the bonobos. The location of the stimuli on the screen (left/right) and the location of the probe were counterbalanced, and stimuli were presented in a randomized order. The whole procedure took about 15 to 20 minutes to complete.

Data filtering

We filtered reaction times (RTs) with extreme values (i.e., RT < 250 and RT > 5000 ms). As our dataset contains a large age range, we also filtered reaction times per age category (0-5, 6-10, 11-15, …, 56-60, 61-85) and calculated the median absolute deviation for reaction times per age category. Finally, we used the following filter: [RT < (Median RT + (2.5 * Median Absolute Deviation))]. After applying this filter, five participants had less than 5% of data points left and were thus excluded from further analysis, leading to a final N of 444. In total, we excluded 15.25% of the data for further analysis.

Statistical analyses

Data were analyzed using a generalized linear mixed model in R studio (v1.4.1106, glmmTMB package, α =.05 (Brooks et al., 2017; R Core Team, 2020)). Experimental trials (40) were nested within participants (ID, 444). We used reaction time (ms) as the dependent variable, random intercepts for all *ID*s (subjects), and used *Congruency*, *Familiarity* (both sum coded), and their interaction terms as fixed factors. Moreover, we used the AIC statistic to determine which distribution (gaussian vs. gamma) fit our data best (Lo & Andrews, 2015). Model assumptions were checked by visually inspecting QQ plots and the residuals plotted against fitted values.

Results

A model with a gamma distribution appeared to fit our data best (AIC_{normal} = 200452, AIC_{gamma} = 198980). Testing whether humans have an attentional bias towards emotions of familiar and unfamiliar others, we found a significant interaction effect between *Congruency* and *Familiarity* ($\chi^2(1) = 3.47$, $p = .047$, Figure 4, bottom left). Planned comparisons showed that participants were significantly faster when a probe replaced an emotional stimulus (*M* = 563.78, *SD* = 116.89) versus a probe replacing a neutral stimulus (*M* = 568.89, *SD* = 121.66) in the *Familiar* condition ($\beta = -.01$, SE = .00, $t(16943) = -2.72$, $p = .007$) but not in the *Unfamiliar* condition ($\beta = .00$, SE = .00, $t(16943) = .08$, $p = .936$. See Table S11.1 for further model output. Also see Table S11.2 for an exploratory analysis of sex and familiarity effects on emotion bias). In short, while our control experiment in Experiment 2 showed that humans generally have a bias towards emotions, Experiment 3 shows that this bias is modulated by familiarity such that humans mainly have a bias towards emotional cues from familiar individuals.

Discussion

In Experiment 3 using a large community sample, we show that human attention is modulated by the emotional expressions of family members and friends. Traditionally, emotional attention is studied using stimuli that depict unfamiliar individuals only. For the first time, we show that familiarity with the expressor in terms of their social or familial relationship differentially affects immediate attention for emotions in humans. Humans have strong affinity with their own social group and often choose to associate with others who are similar to themselves in some respect (Brewer, 1979). This tendency to focus on one's ingroup is likely to be adaptive, as it bolsters cooperation between individuals within the same group, which subsequently provides protection from danger such as competitors (Bowles, 2009). As such, our results contribute to the existing literature by showing that intergroup bias likely already presents itself early on in social perception, and guiding attention to emotions of socially close others.

Interestingly, we did not find evidence for an attention bias towards emotions of unfamiliar individuals, even though this is commonly reported in the literature (Van Rooijen et al., 2017). Importantly, our control experiment as part of Experiment 2 showed that when all individuals on the stimuli are unfamiliar to the participants, an attentional bias towards emotions arises. Thus, it is possible that the presence of familiar individuals within the same experiment dampens the focus of attention on emotional expressions of unfamiliar others (and, for bonobos in Experiment 1, the reverse might be true). The social relevance of the stimuli may thus interact with detecting emotional expressions, prompting stronger activation of attentional and emotional brain mechanisms than when viewing emotions of less-relevant others. Indeed, according to appraisal theory (e.g., Lazarus, 2001), the social relevance of stimuli to the observer likely impacts attentional mechanisms (Wirth & Wentura, 2019). One example of this is that an attentional bias towards threatening stimuli such as angry faces is more pronounced in people with high anxiety compared to individuals with low anxiety, and sometimes not even observed in non-anxious people (Bar-Haim et al., 2007). As the relevance of stimuli can be determined by a range of factors including personal goals, values, and needs (Brosch et al., 2008; Pool et al., 2016), it could be interesting to explore these factors and how they affect attention in more detail in a comparative framework.

General discussion

Emotional expressions are pivotal to understanding the internal state of others and predicting their future behavior, and as such, receive privileged access to attention (Adolphs, 2008; LeDoux, 1998; Öhman et al., 2001b). Crucially, emotions can arise in social situations involving close others, yet are rarely studied in this context. In this study we investigated the potential link between emotional attention and familiarity with the expressor in two closely related species: humans and bonobos. In Experiment 1, attention of bonobos appeared to be attuned to emotional scenes depicting unfamiliar others, but not to emotional scenes depicting familiar others. This emotion bias did not extend to emotional facial expressions of familiar and unfamiliar humans (Experiment 2). For our zoo-visitor sample in Experiment 3, we found that emotional expressions of familiar companions (family, friends, or colleagues), but not unfamiliar others, grab attention. Below we discuss these results within a comparative framework, and consider the study's limitations.

Humans and bonobos seem to share an immediate bias for emotional scenes and expressions (Kret et al., 2016; Kret & Van Berlo, 2021) and we here show that this bias is modulated by familiarity. At least for bonobos, this modulation occurs only when viewing conspecifics, not humans. Interestingly, studies with chimpanzees (Kret et al., 2018; Wilson & Tomonaga, 2018) and our own study with orangutans (*Pongo pygmaeus*) (Laméris et al., 2021, under review) did not find a general bias for emotions using the dot-probe paradigm. No data exist on gorillas (*Gorilla gorilla*). It is possible that, compared to other apes, bonobos are more sensitive to emotions of conspecifics, evidenced by their strongly connected brain pathways involved in emotion processing (Issa et al., 2019; Stimpson et al., 2016). However, three dot-probe studies involving monkeys did find an attentional bias towards threatening faces of conspecifics (King et al., 2012; Lacreuse et al., 2013; Parr et al., 2013). Moreover, looking time paradigms have shown that, chimpanzees, orangutans, and rhesus macaques (*Macaca mulatta*) looked longer to (negative) emotional expressions than to neutral expressions (Bethell et al., 2012; Howarth et al., 2021; Kano & Tomonaga, 2010a; Pritsch et al., 2017). Overall, these findings suggest that an attentional bias towards emotional signals is shared at least within the primate order (and potentially also in other animals, although results are mixed (Kremer et al., 2021; Lee et al., 2018; Luo et al., 2019)).

The brain is proficient at distinguishing between faces of familiar, socially close others and strangers, evidenced by the prioritized and highly optimized processing of familiar faces (Ramon & Gobbini, 2018). Here, we have shown that familiarity further interacts with the processing of emotional expressions at a very early stage of visual attention. Moreover, our findings also suggest that *how* familiarity interacts with the processing of emotional expressions can depend on species-specific characteristics, such as other-regarding tendencies. While humans and bonobos are both social species, humans tend to prefer the social in-group over the out-group (Fehr et al., 2008), whereas bonobos are known to preferentially share food with out-group members (Hare & Kwetuenda, 2010) and peacefully interact with them (Furuichi, 2011). How familiarity modulates emotional attention has not yet been studied in other species, but chimpanzees would be an interesting comparison species as they typically empathize with group members but not with unfamiliar chimpanzees (Campbell & De Waal, 2014; Wilson & Wrangham, 2003) and gaze longer at familiar rather than unfamiliar males (Lewis et al., 2021). As such, one hypothesis could be that chimpanzees have a stronger attentional bias towards emotional expressions of familiar conspecifics than of unfamiliar individuals.

Studies on emotion perception and attention in the other great apes – gorillas and orangutans – are, to the best of our knowledge, rare, but would further provide further evolutionary insights. Gorillas and orangutans have unique social systems, with gorillas living in harem-like societies with one adult male and multiple females and their offspring (Robbins et al., 2004) and orangutans living a semi-solitary life (Singleton et al., 2009). While we did not find an emotion bias in orangutans in another study (Laméris et al., 2021), previous work shows that they look longer at negative facial expressions (Pritsch et al., 2017) and automatically mimic facial expressions (Davila-Ross et al., 2008), but that mimicry is not necessarily affected by familiarity (Van Berlo et al., 2020b). Furthermore, gorillas are known to affiliate less frequently with conspecifics than for instance chimpanzees (Cordoni et al., 2018), but they do appear to mimic facial expressions, specifically the play-face that occurs during playful interactions (Bresciani et al., 2021). We currently do not have clear predictions on how familiarity might modulate emotional attention in these species, but given the existing evidence, an immediate bias towards emotional expressions may only be present in species that have high affiliative tendencies (like bonobos, chimpanzees [but see Kano & Tomonaga (2010) and Kret et al. (2018)], humans, and some monkey species), given that they continuously need to monitor behaviors of others in the group. Again, more research is needed to pin down differences between different species' attention allocation to emotional expressions, and how these interact with social factors such as familiarity.

The results of our experiments should be viewed in light of the study's limitations. The most pertinent one is the disparity between our bonobo and human sample size. Unfortunately, comparative studies often suffer from low sample sizes due to the limited access to individuals and the major efforts and resources that are necessary to conduct non-invasive experiments with animals. Indeed, a recent overview of touchscreen-based studies with great apes in zoos shows an average sample size of four (Egelkamp & Ross, 2019). That said, the findings are still valuable for understanding our own evolutionary roots and great apes' socio-cognitive competencies (similar to how findings on one or two patients with unique brain lesions have been crucial for understanding the neuroscientific foundations of emotion recognition (e.g., Adolphs et al., 1994)). In our study, we partly replicate earlier findings by Kret et al. (bonobos: 2016, chimpanzees: 2018), showing that this type of work is fruitful and can lead to reliable results. We also report individual means in the hopes that these data can eventually lead to combined datasets for future examinations of great ape (social) cognition.

A second limitation is that we were only able to test female bonobos. For ethical reasons, we did not separate individuals from the group while testing, thus it was difficult for the three potential male subjects to get tested (because the females were eager to participate and did not allow the males behind the screen). Nevertheless, this makes it difficult to generalize our findings. For instance, it is possible that there are sex differences in attentional biases for emotions, and in humans, there is some evidence for this idea (for a review, see Kret & De Gelder, 2012). Nevertheless, we did not find any effect of sex of the participant performing the dot-probe and sex of the individual on the stimuli in Experiment3 (Table S11.2). We recently also conducted a dot-probe study involving human emotional scenes (rather than faces), and found no sex differences in attentional bias towards emotional scenes (Kret & Van Berlo, 2021). Yet, two primate studies did show that sex can impact attention allocation (Lewis et al., 2021; Schino et al., 2020). It therefore remains possible there are sex differences in bonobos' attention for emotions (particularly in relation to familiarity) that we could not capture in our study.

Another limitation in our study involves the differences between the configuration of the stimuli used for bonobos (emotional and neutral scenes in Experiment 1, facial expressions in Experiment 2) and humans (also facial expressions in Experiment 3). Moreover, there were discrepancies between the emotional categories used in the experiments, as we used socio-emotional categories for our bonobo stimuli, and basic human emotional expressions for our human stimulus set. Finally, there were differences between the number of unique individuals in the stimulus sets. Bonobos saw multiple unique individuals in Experiments 1 and 2, and humans saw only one familiar and one unfamiliar individual in Experiment 3. These discrepancies are important to note, but still allow for a qualitative comparison of the results. To the best of our knowledge, this study is the first to look at the modulating effects of familiarity on attention for emotions. Thus, we opted to base our study designs on existing literature on attentional biases for emotional facial expressions in humans (see Van Rooijen et al., (2018) for a review) and emotional scenes in bonobos and chimpanzees (Kret et al., 2016, 2018).

Finally, while our results extend previous findings by Kret et al. (2016), the average difference between emotional and neutral trials in the unfamiliar condition is numerically smaller than the difference reported by Kret and colleagues (i.e,. about 12 ms in our study versus 34 ms in the original study). This is likely due to crucial methodological differences. First, the trials in the original study by Kret et al., (2016) paired emotional or neutral bonobos with control animals (sheep or rabbits). In order to directly test how emotional and neutral scenes compete for attention, in the current study, we chose to present emotional and neutral stimuli within the same trial. Second, Kret et al. (2016) used slightly different categories, i.e., eating and panthoot, but these categories did not elicit an attentional bias and thus were replaced by self-scratching in our study. Third, our design also included stimuli of familiar individuals, which likely attenuated the effect we found for unfamiliar individuals.

To conclude, our study contributes to the understanding of how evolution shaped other-regarding preferences of bonobos and humans by showing that they are deeply ingrained in early social perception and, crucially, are shared between the species. The results also demonstrate that *how* familiarity modulates emotional attention can differ between species. Importantly, differences in the environments of bonobos and humans may have helped shape the striking differences in how bonobos and humans attend to emotions of familiar and unfamiliar others. It could therefore be interesting for future work to examine the link between emotional attention and familiarity in a wider range of species, progressing our understanding of the origins of the social mind.

Chapter 3

Attentional bias in humans towards human and bonobo expressions of emotion

Abstract

Correctly recognizing and efficiently attending to emotional situations are highly valuable skills for social species such as humans and bonobos, humans' closest living relatives. In the current study, we investigated whether humans perceive a range of emotional situations differently when these involved other humans compared to bonobos. A large group of children and adults participated in an emotion perception task and rated scenes showing either bonobos or humans in situations depicting distressed or aggressive behavior, yawning, self-scratching, grooming, playing, sex scenes or neutral situations. A new group of people performed a dot-probe task to assess attentional biases toward these materials. The main finding is that humans perceive emotional scenes showing people similarly as emotional scenes of bonobos, a result likely reflecting a shared evolutionary origin of emotional expressions. Other results show that children interpreted bonobos' bared teeth displays as a positive signal. This signal is related to the human smile, but is frequently seen in distressed situations, as was the case in the current experiment. Children may still need to learn to use contextual cues when judging an ambiguous expression as positive or negative. Further, the sex scenes were rated very positively, especially by male participants. Even though they rated these more positively than women, their attention was captured similarly, surpassing all other emotion categories. Finally, humans' attention was captured more by human yawns than by bonobo yawns, which may be related to the highly contagious nature of yawns, especially when shown by close others. The current research adds to earlier work showing morphological, behavioral and genetic parallels between humans and bonobos by showing that their emotional expressions have a common origin too.

Based on:
Kret, M. E. & Van Berlo, E. (2021). Attentional bias in humans toward human and bonobo expressions of emotion. *Evolutionary Psychology, 19*(3). https://doi.org/10.1177/14747049211032816

Introduction

Social animals, like humans and great apes, spend a considerable amount of their time among conspecifics. In interactions with them, they produce, observe and exchange emotional expressions. Because emotional expressions provide relevant information and inform subsequent actions, they are efficiently processed; they readily attract the attention of observers and are recognized easily (e.g., Ekman, 1984; Frijda et al., 1989). The underlying mechanisms of producing and recognizing emotional expressions are deeply grounded in our evolutionary past and it is therefore not surprising that important parallels in emotion processing exist between humans and other great apes (Darwin, 1872; Kret et al., 2020). The majority of studies investigating emotion recognition and attentional biases toward emotions in humans has made use of isolated facial expressions as stimulus materials. However, emotions and facial expressions of emotion alike, are embedded in complex social scenes involving multiple individuals. Moreover, the literature is troubled with inconsistencies regarding gender (Kret & De Gelder, 2012a) and age differences (Van Rooijen et al., 2017). The goal of the current study was to assess the effect of naturalistic emotional scenes on the perception of emotional expressions. To address the aforementioned inconsistencies, we opted for a large community sample consisting of males and females and children as well as adults. Crucially, we compared the processing of scenes showing humans and apes in similar emotional scenarios to address the question of whether effects are human-emotion specific, or more generally linked to emotions and similar when observing emotions expressed by apes. Congruent findings across human and ape scenes would further support their shared evolutionary basis.

There is discussion in the human emotion literature about whether or not emotions and their expressions can be put into categories. Research showing that humans across the world can assign emotion labels to prototypical facial expressions suggests that such distinct categories exist (Ekman, 1984). However, contrasting literature shows that the same emotional expression can be interpreted differently based on context (Kret & De Gelder, 2010, 2012a, 2013; Kret & Straffon, 2018). Crucially, emotional expressions are perceived differently when posed by actors as compared to when real-life stimuli are employed (e.g., McLellan et al., 2010; Motley & Camden, 1988). Despite this evidence, the majority of studies has been using posed expressions, ignoring the fact that these prototypical expressions are not commonly observed in real life. For example, how often do we see the facial expression of fear? The smile, an expression that is common in real life, often gets the simplified label "happiness" in emotion research. However, the smile is a communicative signal with meanings ranging from greeting another person, an expression of love, to a contemptuous smile or a sign of nervousness (Kret, 2015; Kret et al., 2020). Given that emotions can be expressed in multiple ways, the use of isolated stimuli in research leads to perceptual confounds, and therefore muddies the interpretation of results. Take for instance the observed attentional bias toward smiling faces. The question is whether we can be sure this attentional bias is related to the smile or whether it can be attributed to low-level characteristics of the included stimulus materials such as the exposure of the white teeth with the smile (Blanco et al., 2017). Further discussion about these debated topics is beyond the scope of this article. What we aim to do here is to take a step aside and study how emotional expressions are perceived from a large number of complex, naturalistic scenes depicting real life, authentic emotional situations, where emotional expressions are embedded within rich contextual features, thereby partly circumventing the above problems. This approach is ecologically more valid since in the natural world emotional expressions are always embedded in a scene (Kret & De Gelder, 2010).

The experience and expression of emotions is heavily reliant on contextual cues (Hess et al., 2016). The interpretation of facial expressions relies on contextual information, such as body language (Kret et al., 2013a, 2013b; Kret & De Gelder, 2013), outgroup cues (Kret & De Gelder, 2012b; Kret & Fischer, 2018; Liedtke et al., 2018) and the global processing of a social scene (Ngo & Isaacowitz, 2015; Righart & de Gelder, 2008; Van den stock et al., 2014). Likewise, the perception of emotional body language is influenced by the facial expression (Kret & De Gelder, 2012a) and the social context (Kret & De Gelder, 2010). In a study where participants were asked to explicitly label a person's emotional expression, contextual influences on emotion perception varied based on the type of contextual cue, cue relevance, and the perceiver's age (Ngo & Isaacowitz, 2015). EEG studies have demonstrated that the integration of these different pieces of information occurs early in the processing stream, further underscoring its relevance (De Gelder et al., 2006; Righart & de Gelder, 2008; Righart & De Gelder, 2008). Moreover, the perceived valence and arousal from such scenes is modulated by gender (for a review, see Kret & De Gelder, 2012a) and age (e.g., Backs et al., 2005). The dot-probe paradigm is widely used to measure attentional biases toward certain stimulus categories such as expressions of emotion. An advantage is its implicitness and simplicity, so that it can be administered in young children or nonhuman primates (Kret et al., 2016). A recent meta-analysis (of 38 articles including 4,221 children) confirmed that children show a significantly greater bias to threat-related stimuli than to neutral stimuli and this effect was stronger in children with high anxiety, a difference that further increased with age (Dudeney et al., 2015).

In typical emotion perception tasks it is impossible to pull apart effects of emotion as an isolated construct from effects of emotions expressed by a specific species. The latter is commonly assumed, but whether effects are specific to the emotional expressions of other humans can only be investigated by directly comparing biases toward emotional expressions of other species. On the one hand, one could argue that it is easier to embody the emotional expressions from conspecifics than from other species with very different bodies. On the other hand, as Darwin (1872) already noted, the similarities in the expression of emotion across species are high and may render the former argument trivial. An earlier preliminary study gives some insight into this question (Kret et al., 2018). In the study, participants performed a dot-probe task with threatening or neutral expressions expressed only by adult male humans and chimpanzees. The results showed a significant bias toward threat, but no interaction between threat and species; suggesting that processing emotional expressions may not rely on the species of the expressor. This finding further supports the evolutionary argument.

An alternative explanation is that attentional capture from threat is also functional: humans benefit from quickly detecting and responding to threatening stimuli irrespective of whether it is displayed from a human or chimpanzee. Nevertheless, whether the same principle holds for positive expressions remains uncertain. Positive emotions, child or female models were not included so the generalizability of the findings is uncertain. For instance, humans would likely perceive expressions of sexual arousal as more relevant or interesting when expressed by a human compared to a chimpanzee. In contrast, the image of two playing apes expressing joy might attract human attention. Indeed, it is common in zoos for apes to end up tangled in a play of tag or peek-a-boo with human children. In the current study, we aim to disentangle the effect of sender species on emotional attentional biases. To that extent, the bonobo provides an excellent model. Not only is it our closest living relative (together with the chimpanzee we shared a common ancestor that lived roughly 6 million years ago) and shows very similar musculature of the face and body (Diogo, 2018), it also is a very rare species that people do not get to see often. There are only two zoos in the Netherlands that house a group of bonobos, so except for frequent visitors of these zoos, people in general have had few or no learning experiences with these animals.

An important factor that might account for discrepancies in attentional biases toward emotions is individual differences, such as age and gender. As humans age, certain situations might become more prevalent or important. For instance, older and more experienced individuals might quickly see opportunities that children do not, like recognizing that someone is flirting. Similarly, children might detect other opportunities that adults fail to detect, such as recognizing playful intentions in potential play mates. Gender also influences the ability to detect emotional expressions. Specifically, biases toward threat tend to be larger in males than females, especially when they experienced violent environments (e.g., Kret & De Gelder, 2013; Shechner et al., 2012). In our earlier described study, gender or age differences did not modulate the observed attentional bias toward emotions (Kret et al., 2018). However, that study only included greyscale fearful and aggressive expressions of males and therefore this question needs to be investigated in an experiment including multiple emotions expressed by both genders. Whether attention is differentially captured depending on age or gender is part of what will be investigated in the current study.

The current study investigates how people perceive emotions from a large number of naturalistic scenes showing humans or bonobos. With rating scales, we gain insight into how participants perceive the observed images in terms of valence and arousal. Using dot-probe tasks, we address the question of which types of emotion scene capture attention most. Furthermore, using a large community sample allows us to unveil possible effects of age and gender.

Method

Participants

Participants consisted of a large group of visitors and employees of a Dutch zoo (Apenheul Primate Park, Apeldoorn, the Netherlands). The sample size was the result of a fixed number of days of testing agreed with the zoo. One part of the participants participated in a task assessing the perceived emotional valence and arousal of a series of stimuli, the other part in a dot-probe experiment. There were two reasons for deciding à priori to create separate groups for adults and children. First, children and adults took part in slightly different versions of the task: children did get trials with pictures showing bullying behavior, but no overt aggressive scenes and no sex scenes. Second, based on our experience with testing visitors in the zoo, we knew in advance that there would be relatively few 14–18 years old as most families visit with younger children and that age would not be normally distributed. Table 1 summarizes all demographic information. After indicating interest in participating in one of the experiments, participants gave informed consent. The study was approved by the Ethics Committee of Leiden University (CEP17-0213/74 for adults, CEP17-0604/222 for children). All participants were debriefed after the completion of the study.

Table 1. Demographic information of the participants.

Type of task	Species on stimulus	Participants	Gender		Age			
			Female	Male	Mean	*SD*	Min	Max
Emotion perception	Human	Adult	53	47	36.28	15.61	18	75
		Child	13	13	8.41	2	5	12
	Bonobo	Bonobo keeper	4	1	37.88	9.62	26	50
		Adult	201	157	30.5	13.54	18	84
		Child	48	53	9.93	3.05	4	17
Dot-probe	Human	Adult	84	68	33.94	13.08	18	74
		Child	69	80	9.36	2.67	3	17
	Bonobo	Adult	153	135	37.45	16.13	18	78
		Child	51	69	9.28	2.9	4	17

Stimuli

The stimulus materials have been used in two previous studies with bonobo subjects. In the first study, Kret et al. (2016) showed bonobos emotional pictures including sex, grooming, yawning, panthoot, interactions with food, play, and distress (Figure 1). The authors observed that bonobos' attention was immediately captured by emotional scenes. In a subsequent study, Van Berlo et al. (2021) used all new, but similar stimulus materials. In their study, the scenes showing food and panthoots were dropped as these showed hardly any effect in the first study. In addition, based on behavioral observations that were made during the first study, they decided to include scenes where individuals were self-scratching, which occurs under arousal (Maestripieri et al., 1992). The findings of van Berlo replicated the effect observed in Kret et al. (2016), suggesting that bonobos recognize the emotional expressions of conspecifics and that these are thus readable and distinctive from the neutral scenes.

Figure 1. Examples per stimulus category depicting bonobos or humans.
** Note: aggression was only used in adults, and replaced by bullying in the child version of the task. Sex was removed from the child version.*

In the current study, we investigated how humans perceive scenes which are relevant for bonobos and thus kept these scene categories consistent with our previous work. To examine the differences in attentional bias when expressions are produced by bonobos and humans, we created a human dataset equivalent to the bonobo one (Van Berlo et al., 2020a) (Figure 1). The stimulus set included people yawning or self-scratching, two highly contagious behaviors (Campbell & de Waal, 2011; Palagi et al., 2014; Swithenbank et al., 2016). Play scenes showed adults and/or children immersed in playful interactions. For the category grooming we used images of humans embracing each other or combing or braiding each other's hair. The category sex showed a man and a woman in an intense, erotic embrace (without depicting their genitalia). The neutral category showed people walking or sitting. The distress category depicted crying individuals. The category aggression differed for adults and children; adults viewed violent scenes (e.g., fist-fights) and children saw images of bullying scenes. Adults and children further viewed the same stimulus set, excluding the category sex for children. In the bonobo version, all scenes showing distress showed the bared teeth display, a facial expression that is frequently used to signal submission and shows commonalities with the human fearful facial expression and to the expression of the smile (Kret & Straffon, 2018). The distressing images had a negative valence, but did not include overt aggression. Omission of overt aggression was due to the fact that bonobos are a relatively non-aggressive species, therefore there was no sufficient amount of aggressive images available. All stimuli were sized 400 * 330 pixels.

Procedure

The task was displayed on a Dell S2240 Tb touchscreen monitor (21.5 inch, 1920 * 1080 resolution, 60 Hz refresh rate) and ran on a Dell OPTIPLEX 990 desktop computer using E-prime 2.0. The experimental set-up was placed in a quiet, public corner of the indoor bonobo enclosure and facing the wall to minimize distractions (Figure 2).

Figure 2. Picture of a child performing one of the dot-probe tasks in the zoo.

Emotion perception task

Adult participants were asked to rate the valence and arousal of the human ($n = 126$) and bonobo ($n = 467$) pictures on 1–7 rating scales (1 being negative or low arousing; 7 being positive or high arousing). Children used the Self-assessment manikin (line drawing of a cartoonish human figure with a facial expression ranging from very negative to very positive Lang, 1980).

Emotional dot-probe task

Based on the scores of the emotion perception task, we selected 10 scenes per emotional category that were a) best recognized regarding valence and b) most emotionally arousing. The selected pictures were used as stimuli in the human-scene dot-probe task. For the bonobo-scene dot-probe task, stimuli were selected using the validation data of experts gathered during previous studies with bonobos (Kret et al., 2016; Van Berlo et al., 2020a).

During the dot-probe task, a trial started with the presentation of a black dot centered horizontally in the lower quadrant of the display that remained visible until response (i.e., tapped) (see Figure 3). Following a response, two images (emotional vs. neutral) were simultaneously presented on the left and right quadrants of the display (50% chance) for 300 ms and replaced by a black dot either on the left or right location of the display (50% chance). Every trial ended with an inter-trial interval (ITI) of 2,000 ms. The location of the dot and stimuli were balanced. Trials were presented in a pseudo-randomized manner. There were two different versions of the bonobo scene dot-probe task. Version 1 included 45 trials and Version 2 (which had one

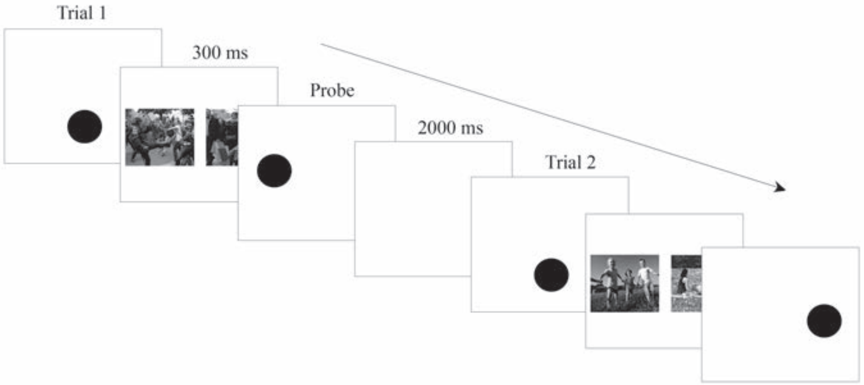

Figure 3. Trial outline of the dot-probe.

additional emotional category, i.e., "self-scratching") included 60 trials. The human scene dot-probe task included two threat categories, namely distress and aggression. For children there was an equivalent category of pictures for aggression showing bullying behavior.

Statistical analyses

All data were analyzed using generalized mixed modeling with experimental trials nested within subjects (SPSS; version 25). A random intercept of participant was included in all analyses. Significant interactions and main effects of Emotion Category were followed-up using *t*-tests. In order to reduce type I errors, the *p*-value was set to 0.01. Follow-up simple contrasts are Bonferroni-corrected. Only the significant effects of major theoretical interest are visualized. We start with reporting main effects, followed by two-way interactions, followed by higher order interactions, in the order of significance (most significant effects first).

Rating scales

The dependent variables in this emotion perception task were Emotional Valence and Intensity (1–7 scaled ratings). We used a linear distribution function to model the data, as they were normally distributed. Three separate models were conducted with minor differences. All models included the fixed factors of Age Group (Child/Adult), Gender (Male/Female), Emotion Category (Sex [adults only]/Groom/Play/Self-scratch/Distress/Aggression [adults and human scenes only]/Yawn/Neutral) and Species Scene (Human/Bonobo). In the first model, we examined the effect of age, gender, expressor species, and emotional display on intensity ratings. Therefore, we omitted the categories Sex and Aggression since a) children did not see these images and b) the aggression category was not included in the bonobo scenes. Therefore, this model included scenes depicting grooming, yawns, play, distress, and self-scratching. In the second model, we zoomed in on human adults and therefore added the emotion category Sex. In the third model, we included only human adults and human scenes, and added the emotion category Aggression.

Attentional bias

The reaction times were filtered with the following procedure. First, age categories were created using 5-year bins (e.g., 0–5 years old; 5–10 years old) and trials with extremely fast (<250 ms) or extremely slow responses (>5000 ms) were excluded. Next, the trials exceeding 2.5 mean absolute deviations (MAD) from the mean per age bin were excluded. The reason for this procedure is that very young children and older people tend to be a lot slower on this task than those with an average age. After applying these criteria, 10.14% of the data was excluded. The data were modelled using a Gamma distribution (log link function), given the typical skewness in RT data.

The statistical procedure was similar to the analysis of the rating scales, unless stated otherwise. In the model for the dot-probe task, two additional factors were included, a) Congruency (probe position replacing the emotional stimulus or not) and b) Distance of dominant hand to the screen (Short/Long). This latter factor is added as a control to account for the shorter distance for a left-handed person from the left hand to the left side of the screen and for a right-handed person using the right hand to the right of the screen compared to left to right and vice versa.

Results

Perceived valence of human and bonobo scenes

In a first analysis where the emotion categories Sex and Aggression were excluded, participants (adults and children), showed a main effect of Emotion Category ($F(1, 11.36) = 124.86, p < .001$). Play was most positively rated out of all categories, followed by groom, neutral, yawn, distress and self-scratch (the latter being least positively rated of all). All emotion categories differed from the neutral scenes in the expected direction (ps ≤ .001). A main effect of Species showed that in general, bonobo scenes were perceived more positively than human scenes ($F(1, 11.36) = 11.83, p = .001$).

An interaction between Emotion Category and Species ($F(5, 11.36) = 87.76, p < .001$) showed that human scenes received more extreme ratings than bonobo scenes. That is, positive emotions received more positive ratings and negative ones more negative ratings when expressed by humans rather than bonobos. This was significant for the categories Play ($t(12.54) = 6.37, p < .001$) and Distress ($t(11.36) = 20.63, p < .001$), with a trend toward significance in the same direction for self-scratching ($p = .031$). In contrast, another trend was observed for yawning behavior, which was rated somewhat more negatively in the bonobo compared to human scenes, possibly because of the visibility of the canines ($p = .024$). There were no differences for neutral ($p = .922$) or grooming scenes ($p = .551$).

Further, an interaction between Emotion Category and Age ($F(5, 11.36) = 7.81, p < .001$) showed that compared to adults, children differentiated the categories less based on valence. Specifically, they gave less negative evaluations of the negative scenes (this was significant for the category Distress *t*(11.36) = 3.64, *p* < .001, with trends in the same direction for self-scratch (*p* = .034) and yawn (*p* = .014)) and less positive ratings following positive scenes (which was significant for grooming *t*(11.36) = 3.44, *p* = .001 and showed a trend in play (*p* = .023).

A three-way interaction between Age, Species, and Emotion Category (*F*(5, 11.36) = 10.86, *p* < .001) demonstrated that this effect was driven by the bonobo scenes. There were no differences between adults and children in the perception of valence from human scenes (*ps* ≥ .023 [for distress]). In contrast, the bonobo scenes were perceived differently. Specifically, compared to adults, children perceived bonobo scenes showing grooming and play less positively (grooming: *t*(11.36) = 4.26, *p* < .001; play: *t*(11.36) = 3.20, *p* = .001) and scenes depicting individuals yawning less negatively (yawning: *t*(11.36) = 5.15, *p* < .001). Further, adults gave all positive emotional scenes more positive ratings than the neutral scenes and all

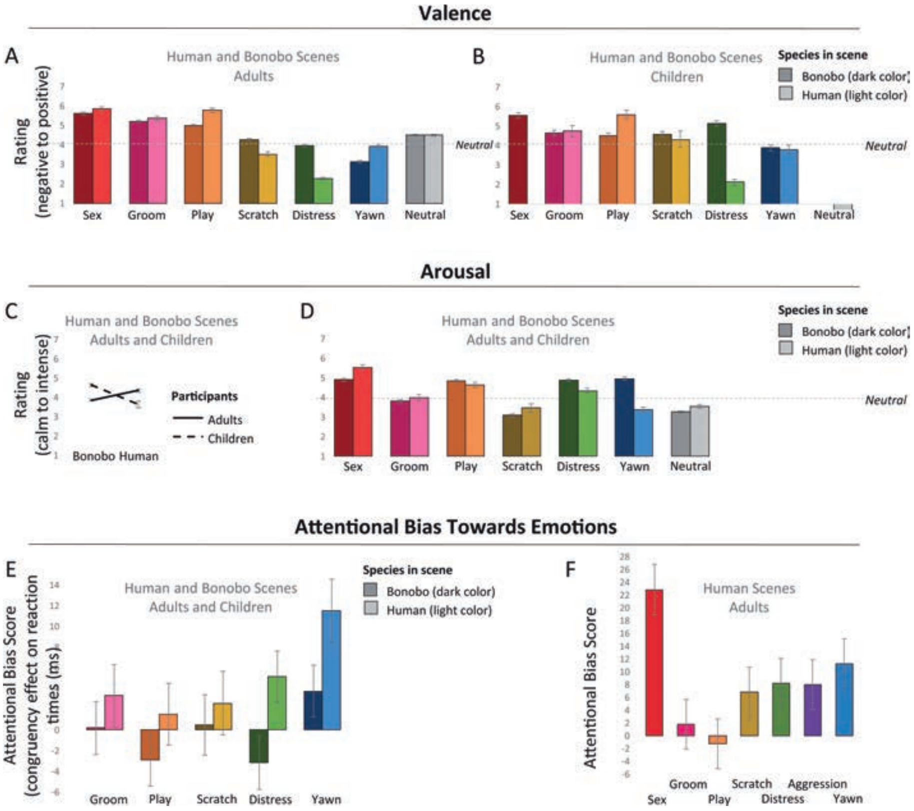

Figure 4. Strongest effects of valence (A: adults; B: children) and arousal ratings (C, D) and attentional bias (E, F: the neutral category is at zero).

negative scenes more negative ratings than the neutral ones. Children, in contrast, did often not differentiate the emotional scenes from the neutral ones in terms of valence. However, they did rate human yawn (*p* = .002) and distress scenes (*p* < .001) more negatively, and play (*p* < .001) more positively than neutral scenes. Regarding the bonobo scenes, they rated scenes depicting yawning bonobos more negatively than neutral (*p* < .001) and distress scenes more positively than neutral scenes (*p* < .001). See Figure 4.

In a second analysis we focused on human adults, adding the category Sex. Within the adult sample, the main effect of Emotion Category remained significant (*F*(1, 10.54) = 357.92, *p* < .001) with the ordering of the categories being similar as before, except that the category Sex received the most positive rating. Within human adults and with the category Sex included, an effect of Gender emerged, with male participants giving more positive ratings than females (*F*(1, 10.54) = 13.58, *p* < .001). The earlier observed effect of Species disappeared (*p* = .239), showing that adults perceive human and bonobo scenes to be equal in valence overall. The interaction between Emotion Category and Species was maintained (*F*(6, 10.54) = 64.88, *p* < .001), and the interpretation was not altered. Interestingly, an interaction between Emotion Category and Gender (*F*(7, 10.54) = 6.73, *p* < .001) showed a significant gender difference in several emotion categories, most prominently in the category Sex (*t*(10.54) = 4.39, *p* < .001), which men rated with almost half a point more positively than women. Further, distress and yawn images were perceived more negatively by women compared to men (distress: *t*(10.54) = 3.32, *p* < .001; yawn: *t*(10.54) = 2.75, *p* = .006).

In a third analysis within human adults, we zoomed in on the human scenes specifically, and added the category Aggression. The results showed a main effect of Emotion Category (*F*(1, 2.62) = 462.22, *p* < .001) and an interaction between Emotion Category and Sex (*F*(1, 2.62) = 3.587, *p* = .001). Of all categories, Aggression was perceived most negatively and Sex most positively. Further, the only gender difference that remained significant after having excluded the bonobo scenes from the analysis was the category Sex, which males rated as more positively than females (*t*(2.62) = 2.85, *p* = .004).

Perceived arousal of human and bonobo scenes

In a first analysis where the emotion categories Sex and Aggression were excluded, participants (adults and children), showed a main effect of Emotion Category (*F*(1, 11.03) = 111.39, $p < .001$). Ordering the categories from high to low based on the arousal ratings led to the following order: Play (most arousing out of all categories), Distress, Yawn, Groom, Neutral, Self-scratch (rated as least arousing of all). Apart from the category Self-scratch ($p = .110$), all emotion categories differed significantly from neutral (*p*s < .001). A gender difference showed that women gave higher intensity ratings than men ($F(5, 11.03) = 29.11$, $p < .001$).

An interaction between Emotion Category and Species ($F(5, 11.03) = 31.39$, $p < .001$) showed that human observers judged specific emotion categories differently when the scenes depicted bonobos compared to humans. Most strikingly, while yawning was considered the least arousing emotion category of the human scenes, it was the most arousing one of the bonobo scenes. For the rest of the expressions, the overall pattern was similar although the bonobo scenes received higher arousal ratings than human scenes for the categories Yawning ($t(11.03) = 8.71$, $p < .001$) and Distress ($t(11.03) = 4.03$, $p < .001$), but lower arousal ratings than human scenes for the category Neutral ($t(11.03) = 6.84$, $p < .001$ [with a similar trend for self-scratching ($p = .05$)]).

Further, an interaction between Species and Age Group ($F(1, 11.03) = 110.02$, $p < .001$) showed that while children rated bonobo scenes as more intense than human scenes ($t(11.03) = 6.14$, $p < .001$), the opposite was true for adults, who rated human scenes as more intense ($t(11.03) = 10.90$, $p < .001$). On similar lines, children rated bonobo scenes as more intense compared to adults ($t(11.03) = 14.72$, $p < .001$) while the opposite was true regarding the human scenes ($t(11.03) = 4.59$, $p < .001$).

An interaction between Age Group and Emotion Category ($F(5, 11.03) = 12.45$, $p < .001$) showed that compared to adults, children gave higher intensity ratings to scenes depicting play ($t(11.03) = 4.00$, $p < .001$) or neutral actions ($t(11.03) = 9.37$, $p < .001$), but lower arousal ratings to distress scenes ($t(11.03) = 2.61$, $p = .009$).

The two-way interactions were further qualified by two three-way interactions. First, there was an interaction between Age Group, Species and Emotion Category ($F(5, 11.03) = 21.87$, $p < .001$). Follow-up tests revealed that compared to adults, children gave higher ratings following most bonobo scenes (all categories *p*s < .001, except self-scratching; $p = .18$, although numerically in the same direction). Zooming in on the human scenes showed that compared to adults, children perceived the distress scenes as less intense ($t(11.03) = 8.89$, $p < .001$) and the neutral scenes as more intense ($t(11.03) = 4.74$, $p < .001$). Further, while adults perceived all but the yawn scenes as more intense when showing humans compared to bonobos (*p*s < .001, but with an opposite effect for yawning [$p < .001$]), children perceived play (*t*(11.03) = 3.76, *p* < .001), distress (*t*(11.03) = 8.24, *p* < .001) and yawning (*t*(11.03) = 7.68, *p* < .001) more intensely from scenes showing bonobos rather than humans. A final set of comparisons showed that almost all emotion categories were perceived as more intense than neutral (*p*s < .003) except self-scratching bonobos (perceived by adults *p* = .77 with an opposite effect in children *p* = .003), and in children, self-scratching humans (*p* = .085) or grooming humans (*p* = .933).

Finally, there was a three-way interaction between Sex, Emotion Category and Species (*F*(5, 11.03) = 5.69, *p* < .001). Follow-up tests revealed that compared to men, women gave higher arousal ratings for the following bonobo scenes: grooming (*t*(11.03) = 6.47, *p* < .001), yawning (*t*(11.03) = 4.09, *p* < .001), play (*t*(11.03) = 3.01, *p* = .003), neutral (*t*(11.03) = 3.75, *p* < .001) and the following human scenes: distress (*t*(11.03) = 4.57, *p* = .002), yawning (t(11.03) = 2.92, *p* = .004 and neutral *t*(11.03) = 2.91, *p* < .001). Also, while both males and females gave higher intensity ratings following yawning bonobos versus humans (males *t*(11.03) = 7.14, *p* < .001; females *t*(11.03) = 6.57, *p* < .001) and lower ratings following neutral bonobos versus humans (males: t(11.03) = 4.72, *p* < .001; females: *t*(11.031) = 6.10, *p* < .001), males also gave higher ratings following distressed bonobos versus humans (*t*(11.03) = 4.81, *p* < .001) and lower ratings following grooming *t*(11.03) = 3.79, *p* < .001 and playing *t*(11.03) = 2.33, *p* < .001) bonobos versus humans. Although both males and females generally gave higher arousal ratings to the emotional scenes compared to the neutral ones, this pattern was a bit stronger in men. In men, this was significant for all categories (*p*s ≤ .001) except self-scratch (bonobo scenes: *p* = .045; human scenes *p* = .466) and human yawns (*p* = .082). In women, apart from generally perceived higher intensity from the emotional compared to neutral scenes (*p*s ≤ .001), no differences were found in the category self-scratch (bonobo scenes: *p* = .036; human scenes *p* = .490), and of the human scenes the categories grooming (*p* = .225) and yawn (*p* = .850).

In a second analysis, we focused on human adults, adding the category Sex, since children did not have this category. Within the adult sample, the main effect of Emotion Category remained significant (*F*(1, 10.13) = 289.23, *p* < .001) with the ordering of the categories almost being identical as before, except that the category Sex received the most positive rating of all. The earlier observed main effect of Species was maintained, showing higher ratings for the human scenes (*F*(1, 9.52) = 289.23, *p* < .001). Similarly, the gender difference was maintained as well (*F*(1, 10.13) = 22.71, *p* < .001). The interaction between Emotion Category and Species was also maintained (*F*(6, 10.13) = 28.50, *p* < .001), and the interpretation was the same.

An interaction between Emotion Category, Species and Sex (*F*(7, 10.13) = 4.25, *p* < .001) showed a significant gender difference in several emotion categories. Specifically, human distress, human and bonobo yawn images as well as neutral, grooming and playful bonobo scenes were perceived more intensely by women compared to men (human distress: *t*(10.13) = 3.74, *p* < .001; human yawn: *t*(10.13) = 2.84, *p* < .001; bonobo yawn: *t*(10.13) = 3.75, *p* < .001; neutral bonobo scenes: *t*(10.13) = 4.52, *p* < .001; bonobo grooming *t*(10.13) = 6.75, *p* < .001) and bonobo play *t*(10.13) = 3.71, *p* = .002).

In a third analysis in human adults, we zoomed in on the human scenes specifically, and added the category Aggression. The results showed a main effect of Emotion Category (*F*(1, 2.62) = 141.24, *p* < .001) and an interaction between Emotion Category and Sex (*F*(1, 2.62) = 2.22, *p* = .008). Of all categories, Aggression was perceived most intensely and Neutral least. Further, the only gender difference that remained significant after having excluded the bonobo scenes from the analysis was the category Distress, with females rating these scenes as more intense than males (*t*(2.62) = 2.78, *p* = .006).

Attentional bias toward human and bonobo emotional scenes

Reaction times were analyzed in a generalized mixed model with the fixed factors Age Group, Gender, Species Scene, Emotion Category, Congruency (i.e., the probe appearing behind the emotional scene or not), their interactions, and in addition Dominant Hand Distance (i.e., location of probe in relation to handedness of participant). Again by using a backward elimination method, we came to the final, most parsimonious and best-fitting model that included a subset of these factors. For brevity, only significant effects that include the factor Congruency are described in the text.

In a first analysis where the categories Sex and Aggression were excluded, participants (adults and children), showed a main effect of Congruency (*F*(1, 29.07) = 6.78, *p* = .009), demonstrating that reaction times were faster when the probe replaced an emotional compared to a neutral scene, indicating heightened attention for the emotional category. An interaction between Congruency and Species (*F*(1, 29.07) = 6.84, *p* = .002), showed that the Congruency effect was significant for human scenes (*t*(29.07) = 3.76, *p* < .001), but not for bonobo scenes (*p* = .698, see Figure 2E). There were no gender differences observed (*p*s = .441 in a pre-final model) and Age Group did not significantly modulate the congruency effect either (*p* = .053, trending toward a larger effect in adults).

In a second analysis, we focused on human adults, adding the category Sex. After adding this factor, the main effect of Congruency was maintained (*F*(1, 15.45) = 8.17, *p* = .004). In addition, without the child participants and with the extra category of Sex, the previously significant interaction between Species and Congruency was rendered insignificant (*p* = .051), which replicates our earlier findings, showing that adults attend to emotions quickly, regardless of the species that expresses them (Kret et al., 2018).

Third, we zoomed in on human adults further, and specifically on their attentional biases toward human scenes. Here we had an additional emotion category, being Aggression. This analysis showed a main effect of Congruency (*F*(1, 5.71) = 13.02, *p* < .001) and an interaction between Congruency and Emotion Category (*F*(1, 6.65) = 3.80, *p* = .001). Significant congruency effects were observed in the categories Sex (*t*(6.65) = 5.75, *p* < .001) and Yawn (*t*(6.65) = 2.89, *p* = .004), trends toward significance were observed for Distress (*p* = .037), Aggression (*p* = .041) and Self-scratch (*p* = .084) and no effects for Grooming (*p* = .641) and Play (*p* = .742, see Figure 2F).

Discussion

Emotional expressions are pivotal to our social life. Correctly recognizing expressions and quickly attending to them can have life-saving consequences and long-lasting effects on social relationships. Like humans, bonobos are social species and have a rich repertoire of expressions. The goal of the current study was to investigate whether human participants perceive emotional scenes showing people similarly or differently as matched emotional scenes of bonobos. Specifically, how do laypeople perceive human scenes compared to scenes depicting bonobos? Overall, the results show more similarities than differences between the perception of human compared to bonobo scenes, especially in adult observers. In general, participants were able to assign appropriate valence and arousal ratings to the emotional scenes and also showed an attentional bias toward them. Interestingly, they did not only do so for the human scenes, but also for the bonobo scenes. In addition to these overall findings, the perception of the scenes differed between adults and children and females compared to males, which can potentially be attributed to different levels of experience with certain expressions or to differences in which emotional expressions are most relevant for specific individuals. In the following section, we first discuss the results regarding the valence ratings, followed by the arousal ratings and last but not least, attentional biases, reflected in the results of the dot-probe tasks.

Overall, participants' valence ratings supported our hypothesis: positive scenes were given positive ratings and negative scenes negative ratings. Despite a similar pattern in valence ratings between human and bonobo scenes (See Figure 2A), this effect was amplified for the human scenes, particularly in child observers. This pattern demonstrates that expressions from conspecifics might indeed be easier to interpret than expressions from another species (see also Fugate et al., 2010; Kret et al., 2018). Interestingly, despite similarities in ratings between adults and children, we observed differences in the magnitude of this effect. Specifically, compared to adults' ratings, the valence pattern was less pronounced in children. This finding is in accordance with earlier literature, showing that children's understanding of emotional expressions is not fully developed yet (Widen, 2013). Interestingly, we found that children perceived the photographs showing bonobos in distress positively rather than negatively. The bared teeth display that was shown in scenes of the distress category is related to the smile; the latter is a ritualized version of the former (Van Hooff, 1972). Importantly, the meaning of the bared teeth display varies, and should be interpreted within the context in which it occurs. This expression can signal fear, subordinance or affiliation and the same is true in humans; a smile is not always a positive signal and can express nervousness or even contempt (Kret & Straffon, 2018). However, taking into account the context when it comes to interpreting these expressions may be something that most adults have learned but that children have not yet learned to incorporate. This interpretation remains speculative as in the current study, the human positive scenes showed smiles and laughter and these expressions never occurred in the negatively valenced scenes. Another possibly is that children's understanding of specific emotions begins not with static images and the expressions visible in them, but with the antecedents and behavioral consequences of the emotional situation.

Of all categories, self-scratch was perceived most negatively and play most positively. Within the adult sample we also included sex scenes and scenes showing aggression, which received extremely positive and negative ratings respectively. Male participants in particular evaluated the human sex scenes more positively than women. Previous literature has shown that gender differences in the pleasant dimension occur only for erotica, with more positive ratings for men than women (Bradley et al., 2001). Surprisingly, these differences were very minor in the current sample. For instance, in the study by Bradley et al. (2001), women rated images of "erotic couples" or "opposite sex erotica" roughly 65% positive and men 85%.

However, in our study, the percentages were about 81% for women and 87% for men. A possible explanation for this effect is that we included more women-friendly images (i.e., pictures of couples embracing, with no close-ups of their genitalia or explicit depiction of sexual acts). Indeed, previous research has shown that women report increased negative affect after viewing mainstream (i.e., male-centered) sexual material (Koukounas & McCabe, 1997) but increased positive affect after viewing women-friendly material (Laan et al., 1994; Mosher & Maclan, 1994, see also Heiman, 1977). The contextual cues provided in women-friendly sexual material facilitate their positive appraisal and might promote the detection of sexual feelings in women (Laan & Everaerd, 1995; Laan & Janssen, 2007). Given these findings, the observed gender differences in self-report of positive affect might be due to stimulus selection, even though men consistently tend to give higher ratings to sexual material (Janssen et al., 2003). Alternatively, social desirability factors were at play, as women might be more inclined to modulate self-reports of positive affect in response to sexual material (Morokoff, 1985; but see Brody et al., 2003). In contrast, men might overestimate their responses (Catania et al., 1990). The difference between the current and the previous study might reflect a cultural difference between the USA versus the Netherlands. For example, while naked breasts or explicit sex scenes are common in typical European movies, such scenes are less common in American movies. Although we cannot with certainty say what the different finding in the literature regarding the perception of sexual images in males as compared to females causes, a crucial consideration is that our sample consisted of people from the general public. This implies that our findings are more representative for the general population than previous laboratory-conducted sex research that typically included students.

Compared to male participants, females interpreted the human distress and yawn scenes relatively negatively. These findings are in line with earlier literature. The majority of previous work used images from the International Affective Picture System (IAPS; Lang et al., 2005), a collection of standardized and digitized color photographs that depict objects and scenes. Although considerable agreement has been found between men and women in their categorical labeling of these images to different emotions (Mikels et al., 2005), women typically assign unpleasant pictures a more negative valence rating than do men (Bradley et al., 2001; Gard & Kring, 2007).

Apart from valence ratings, we also asked participants how they perceived the scenes in terms of eliciting arousal. In general, they evaluated the emotional scenes as more arousing than the neutral scenes. Play scenes received the highest intensity scores and the self-scratch category the least. Interesting differences were observed in how people evaluated human compared to bonobo scenes. Specifically, while yawning was considered the least arousing emotion category of the human scenes, it was the most arousing one for the bonobo scenes. Possibly, the display of the relatively large canines of bonobos played a role here. When humans yawn, they typically do not show their teeth and the fact that they often cover their mouth with their hands, shows that it is not a socially well-accepted expression. The canines provide a likely interpretation given that a similar finding was observed for the distress category where the canines are visible as well. At the same time, such differences make it problematic for the interpretation of the current findings. Is the effect indeed due to the canines? Or is it that the opening of the mouth of a bonobo raises a greater risk of being bitten because it is more difficult to predict subsequent actions of another species that has a completely different body and that one has hardly any experience with? One species might be better at expressing a certain emotion thanks to certain physical characteristics. For example, bonobos have more bodily hair than humans and during dominance displays, these hairs my bristle, called pilorection, which makes them appear larger and perhaps get the message across better. Whether or to what extent physical differences in the face and body of different species translate into perceptual differences of their expressions is a research topic that has remained unexplored. This needs to be confirmed in a comparative study with bonobos and humans. Importantly, differences in facial musculature between bonobos and humans are negligible (Diogo, 2018).

The literature is torn on the intensity between differentially valenced stimuli. Which one would be more intense: sex or aggression? Interestingly, aggressive and sexual images were, on average, perceived as most intense in our study. In contrast to earlier literature showing that men perceive sexual images as more intense compared to women (Bradley et al., 2001), we observed no such difference. However, in Bradley et al. (2001), the greatest difference was found in the valence ratings and the intensity ratings of the sex scenes differed only mildly. In the current study, a gender difference did occur in the human distress category. Females indicated to perceive these scenes as more arousing than males, which is also in line with earlier work (Bradley et al., 2001). Females also perceived some of the bonobo scenes as more arousing then men, including scenes showing yawns, grooming and play. However, a gender difference was also observed in the neutral bonobo scenes so this might also just reflect a more general gender difference in self-report. Indeed, in general, women gave higher intensity ratings than men, as demonstrated by a main effect (see also Bradley et al., 2001; for a review, see Kret & De Gelder, 2012a).

The analysis of the arousal ratings also yielded some interesting results when comparing adults and children. Intriguingly, while adults perceived the human scenes as more intense than the bonobo scenes, an opposite effect was observed in children. To what extent the zoo setting influenced these results is a factor that may be taken into consideration in a future study. However, the effects were not general, which speaks against such an explanation. More specifically, children gave higher intensity ratings to bonobo play scenes. Moreover, compared to children, adults gave higher intensity ratings for scenes depicting human distress, but lower ratings for bonobos in distress. As we will see in the following section, none of these effects were linked to specific age-related attentional biases.

The results of the dot-probe task show that in line with our earlier study, humans show a robust attentional bias toward emotions (Kret et al., 2018). The attentional bias was stronger when humans observed emotions that were expressed by other humans compared to expressions by bonobos. Interestingly, while humans' attention was immediately captured by images of yawns, this was not driven by the visibility of the larger canines of bonobos which could potentially pose a high threat. Instead, attention capture was most pronounced when seeing other humans yawn. Yawns are extremely contagious and yawn contagion seems to work particularly well between close others, possibly due to heightened attention (Massen & Gallup, 2017). At first sight, this increased attentional bias toward human emotions seems to be in contrast to the findings of our earlier study (Kret et al., 2018). However, the current study deviates from our previous work in several important ways. Most crucially, the included stimulus materials consist of naturalistic scenes instead of isolated and greyscale neutral or threatening body expressions of male apes and humans in our earlier study. More importantly, within the adult sample of our study, the interaction between species and congruency disappeared. This suggests that the perception of other species emotional expressions is not fully developed yet in children, which is also in line with their dampened valence effects. As a following, attentional biases toward emotions or other stimuli may be partly learned (Guo et al., 2019). Other research has indeed shown that learning effects might modulate the outcomes of a dot-probe task. For instance, previous research has shown that people suffering from alcohol dependency have greater biases toward photographs of alcoholic beverages than control subjects (Townshend & Duka, 2001). Even more strikingly, in non-dependent social drinkers an attentional bias toward alcohol-related stimuli increased after priming with a small (but not large) dose of alcohol

(Duka & Townshend, 2004), which shows that even a brief learning episode might already influence attention mechanisms.

When zooming in on conspecific scenes, the results show that humans' attention was mostly captured by scenes showing sex and yawns. This is partly in line with our earlier study in bonobos where strongest effects were observed in scenes depicting yawns, grooming, and sex (in that order) (Kret et al., 2016). Can we conclude from this that humans are even more attuned to sex than the hyper-sexual bonobo? Unfortunately, we cannot, due to a limitation of this study. That is, the human sexual scenes were standing out from the rest of the images as only that category showed half-naked people. We could not circumvent this problem in the current study as alternative approaches had other drawbacks. For instance, it is difficult to have completely neutral images showing nudes. Even more tricky is to find other emotional scenes showing emotional expressions (e.g., a group of aggressive nudes). That said, future studies should address this confound by using a stimulus category showing nudes in at least relatively neutral poses. The fact that out of all stimulus categories, human adults rated the category sex as most arousing and most positive makes such a follow-up even more appealing.

Conclusion

Based on participants' explicit valence and arousal ratings and on their attentional biases toward certain stimulus categories, we can conclude that overall, humans perceive emotional scenes showing people similarly as emotional scenes of bonobos. Especially because this finding was observed in lay people who rarely see bonobos, this effect cannot be explained by learning, but likely reflects a shared evolutionary origin in these expressions themselves.

Some expressions have more communicative potential than others. The smile or the bared teeth display are examples of clear signals, meant for conspecifics to be seen. These expressions can have multiple meanings which can be interpreted correctly within the specific context. We found that children may still need to learn to use these contextual cues when judging a situation as positive or negative.

The sex scenes were rated very positively, especially by male participants. Even though they rated these more positively than women, their attention was captured similarly, by far surpassing all other emotion categories. It is interesting that this sex difference which was observed on an explicit level, disappeared when measured implicitly.

An important consideration when studying emotions comparatively are the differences in species' physique. Although there are few anatomical differences between humans and bonobos (Diogo, 2018), other differences, such as the fact that bonobos have fur and humans walk bipedally and stand up straight, may make the transmission of certain emotions easier in one species than the other. Bonobos have much larger canines than humans. These canines stand out and could provide a threat signal. Interestingly, humans' attention was captured more by human yawns than by bonobo yawns, an effect that is more likely to be explained by the highly contagious nature of this stimulus, which is typically enhanced when shown by close others (Palagi et al., 2014). From that point of view, the difference in height between children and adults and their subsequent different viewing angle when reading adults facial expressions, may impact in how these are perceived and learned.

Chapter 4

Attentional selectivity for emotions: humans and bonobos compared

Abstract

Perceiving emotions in others is at the foundation of higher-order social cognition. Currently, we do not fully understand how evolution shaped the cognitive processes underlying emotion perception. Bonobos (*Pan paniscus*) are our closest relatives, and have more developed brain structures involved in emotion processing and exhibit stronger emotion regulation abilities compared to other apes. This makes bonobos an important animal model for understanding the evolutionary development of emotion perception. Here, we investigated how bonobos and humans attend to emotionally-laden scenes in a preferential looking task using eye-tracking. With Bayesian mixed modeling, we established that in both species attention is spontaneously sustained to emotional scenes of conspecifics rather than heterospecifics. Moreover, scenes displaying distress held attention longest compared to neutral scenes, consistent with studies finding an initial attentional bias towards potentially threatening signals. Additionally, bonobos and humans attended longer to sexual scenes compared to neutral scenes, in line with sex being highly rewarding in both species. Humans also attended longer to scenes involving grooming and embracing, as well as play. These findings suggest that emotional signals are relevant to bonobos and that eye-tracking can provide a unique window into apes' affective capacities.

Based on:
Van Berlo, E., Kim, Y., & Kret, M. E. (2021). Attentional selectivity for emotions: humans and bonobos compared. *Manuscript submitted for publication*.

Introduction

Emotional expressions are the conduit through which information about experiences, desires, and intentions are communicated to others. Perceiving emotions is therefore an adaptive process that is crucial to humans and other social animals (Ferretti & Papaleo, 2019; Kret et al., 2020; Nieuwburg et al., 2021). In humans, emotional information is so important that the brain prioritizes its processing even when attentional resources are limited. There is now some evidence that this emotion-biased attention is not only present in humans, but also in great apes (Kano et al., 2018; Kano & Tomonaga, 2010a; Kret et al., 2016; Pritsch et al., 2017; Van Berlo et al., 2020a). However, the manner in which great apes perceive others' expressions of emotions is not yet well understood. As emotions drive not only behavior, but also cognitive mechanisms such as memory, learning, attention, and decision-making (Dukes et al., 2021), examining how they are perceived and recognized by non-human animals can help us reconstruct the evolutionary history of (social) cognition in our species. Moreover, it will allow us to improve our understanding of affective states in animals. Through a comparative framework, in this paper, we investigate emotion-biased attention to emotionally salient scenes in humans as well as our closest relatives, bonobos (*Pan paniscus*).

The human brain is adept at selectively processing information about conspecifics (other members of the same species), and especially emotional expressions are an important source of information that can trigger selective attention (Treue, 2003). In humans, a robust body of evidence shows that emotionally salient information such as smiles or angry faces is preferentially remembered and attracts attention when attentional resources are limited (Petersen & Posner, 2012). Sensory systems are not only tuned to favor facial expressions but also whole body expressions of emotions (Kret et al., 2013a) as well as emotional scenes (Kret & Van Berlo, 2021). In general, the findings show that an attentional preference for emotionally salient information is closely tied to survival, punishment, and reward, thus likely rooted in evolutionarily old mechanisms (Öhman et al., 2001b), and likely shared with other species.

The importance of perceiving and recognizing emotional expressions is not uniquely human. In the last decade, most research efforts on the perception and recognition of emotional expressions have focused on the great apes; our closest extant relatives. Great apes express a large range of behaviors to communicate their desires and intentions to others, and primate brain circuits that are involved in the processing of social and emotional information are similar to that of humans (Hirata et al., 2013; Pinsk et al., 2009; Tsao et al., 2008). Great apes are known to automatically mimic facial expressions of others (Davila-Ross et al., 2008; Laméris et al., 2020; Palagi et al., 2019b; Van Berlo et al., 2020b), which is often linked to emotion contagion, or the convergence of emotional experiences (Pérez-Manrique & Gomila, 2022) (also see (Adriaense et al., 2020) for a critical review). Furthermore, great apes console conspecifics in distress (Clay et al., 2018). Work on physiological determinants of emotion perception indicates that in chimpanzees (*Pan troglodytes*), seeing or hearing conspecifics fight creates changes in cortisol level, heart rate variability, skin temperature (Dezecache et al., 2017; Kano et al., 2016), and temperature in the inner ear (Parr & Hopkins, 2000). Finally, there is some evidence that great apes can discriminate between emotional faces of conspecifics (Buttelmann et al., 2009; Parr, 2001), and that memory is enhanced for emotional stimuli (Kano et al., 2008). This converging evidence, therefore, suggests that great apes share our sensitivity to emotional cues.

Some work has looked into the continuity of emotional expressions and their perception and recognition across different species. All mammals likely share homologous brain structures underlying emotional networks (Panksepp, 2011), and already over a century ago, Darwin theorized that expressions of emotions are universally shared among certain animals. Indeed, within the primate lineage, there is some overlap between human expressions of emotions and that of other primates (Kret et al., 2020). Moreover, one study showed that orangutans (*Pongo pygmaeus*) and human children looked longer at fearful human expressions, and the silent bared-teeth display of orangutans (Pritsch et al., 2017). These results suggests that emotional faces that carry a similar meaning in the two species (i.e., fear) are relevant enough to attend to. While this work is promising, it is clear that more research is needed to understand the phylogenetic continuity of emotional expressions and their perception across species. There is still a great deal to explore in terms of the mechanisms underlying emotion perception in great apes, and specifically, very little work has examined the attentional processes underlying emotion perception in these animals.

Two studies looking into implicit attention using a dot-probe paradigm found that bonobos attend faster to emotionally-laden scenes of others compared to neutral scenes (Kret et al., 2016), and especially of unfamiliar conspecifics (Van Berlo et al., 2020a). This effect has not been found in chimpanzees (Kret et al., 2018; Wilson & Tomonaga, 2018), but it is not yet clear whether methodological considerations (e.g., ecological validity of stimuli (Kret et al., 2018), or stimulus presentation duration

(Wilson & Tomonaga, 2018)) contributed to the null-results. Moreover, two recent studies showed that in apes, emotional cues such as the play face (Laméris et al., 2022) or snakes and food items (Hopper et al., 2021a) impact reaction time on an emotional Stroop task. Finally, eye-tracking studies with chimpanzees and orangutans revealed spontaneous gazing at negatively valenced emotional signals (Kano & Tomonaga, 2010a; Pritsch et al., 2017). Combined, these findings suggest that like in humans, apes' attention is tuned to emotionally salient information. However, different methodologies may tap into different attentional processes (with e.g., Stroop tasks measuring interference in attention, and dot-probes and eye-tracking likely measuring bottom-up or top-down attention), and few studies have directly compared how humans and great apes view emotional expressions or emotionally salient scenes. The aim of the current study is to further examine how apes, specifically bonobos, compare to humans in their allocation of attention to emotionally valent stimuli using eye-tracking. Compared to other apes, bonobos show marked differences in brain areas involved in social cognition, with a higher degree of connectivity and volume in the amygdala (regulating emotions, attention, memory, and social decision-making) and subgenual anterior cingulate cortex (regulating positive affect and arousal) (Issa et al., 2019; Stimpson et al., 2016). This makes them an interesting referential model to reconstruct the evolution of emotional capacities (Gruber & Clay, 2016). At the same time, bonobos are underrepresented in socio-cognitive studies due to their rarity and zoos and their endangered conservation status (Fruth et al., 2016). As such, bonobos' unique socio-emotional characteristics warrant a closer look at how this species perceives emotions.

To this end, we used a preferential looking paradigm with eye-tracking to investigate whether attention of bonobos (experiment 1) and humans (experiment 2) is preferentially sustained to emotionally-laden scenes of conspecifics or heterospecifics (i.e., the other species). Previous findings show that emotionally salient signals modulate the early stages of processing social signals (Hopper et al., 2021a; Kret et al., 2016; Laméris et al., 2022; Van Berlo et al., 2020a). Building on this, we expect that if emotions hold relevance to bonobos beyond an initial attentional bias, they will show a longer looking duration to emotional compared to neutral scenes, similar to humans. Moreover, we expect that bonobos and humans also attend longer to emotional scenes of heterospecifics, as there is some continuity between emotional expressions of great apes and humans (Kret et al., 2018).

Experiment 1: Examining biased attention to emotions in bonobos

Method

Participants

Our sample included four bonobos (Besede [12 yo], Kumbuka [18 yo], Monyama [7 yo], and Zuani [~16 yo]; all female) that are part of a social group of 12 individuals housed in the primate park Apenheul, Apeldoorn, The Netherlands. Except for Zuani, all bonobos took part in two prior touchscreen studies (Kret et al., 2016). At time of testing, none of the bonobos were pregnant nor were on contraceptives. During winter time (from November to the end of March), the park is closed for visitors, allowing us to conduct experimental research. All but one individual (Zuani) were born and raised in captivity. During non-testing hours, the bonobos had access to a 2812m^2 outdoor and 158m^2 indoor enclosure, and testing took place in the indoor enclosure.

The zoo kept the bonobos separated into two groups that varied in group composition on a weekly basis to mimic naturalistic fission-fusion dynamics. During testing periods, only one group of the bonobos was given access to the test apparatus. For ethical reasons, the group was never split further. This meant that when one individual was tested, its group members were present nearby. Water was available ad libitum, and food (a variety of vegetables, fruits, and branches and leaves) was provided four to five times a day, as well as nutritionally balanced mash.

Tests with the bonobos followed the EAZA Ex situ Program (EEP) guidelines, formulated by the European Association of Zoos and Aquaria (EAZA). Bonobos voluntarily participated in the experiment and were never restrained or forced to take part. Furthermore, only positive reinforcement (juice) was used during training and testing, and juice was also offered to the bonobos that did not take part in the experiment. Data were collected between February 2017 – March 2017, and December 2017 – April 2018.

Equipment

Our setup is comparable to those in other research facilities (see Hopper et al., 2020), and involves one PC running Tobii Studio (v.3.4.8), two computer screens (one for the experimenter, one for the participant, 1280x1024 pixels), a webcam to record the bonobos while they were tested, and a Tobii X2-60 eye tracker mounted on one of the screens. One computer screen, together with the eye tracker and the webcam, was placed inside a wooden box inside the bonobos' enclosure (Figure 1).

The front of the wooden box was a 3 mm thick, scratch proof polycarbonate plate. At mouth's height, a drinking nozzle was attached to the panel. During the experiment, bonobos were rewarded with diluted juice (1 part syrup, 5 parts water) at short intervals (roughly every 5 seconds), and provided through the nozzle. To minimalize distractions, other bonobos present in the enclosure were rewarded with the same juice by the caretaker, after they performed a body-part training that is used for veterinarian purposes. Bonobos were familiar with drinking from the nozzle because their enclosures were also fitted with these nozzles for drinking water. The computer and the other screen for the experimenter were located outside of the enclosure. This second screen displayed Tobii Studio Pro's Live Viewer, enabling the experimenter to track where bonobos were looking in real time.

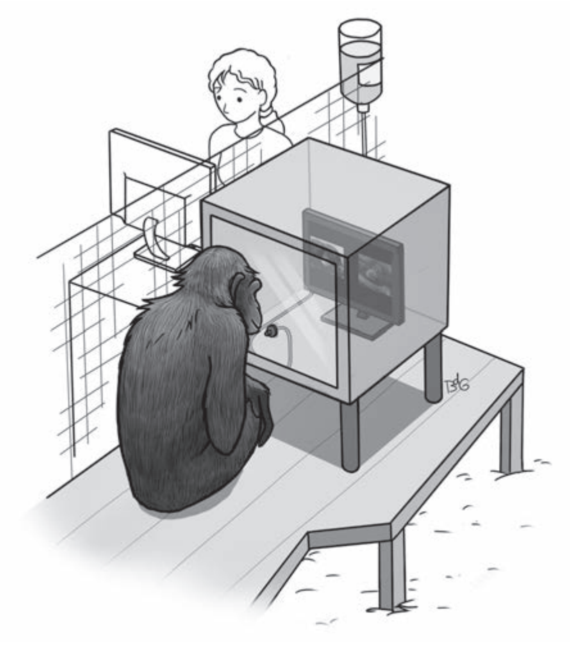

Figure 1. Drawing of the setup at primate park Apenheul. Illustration by Brenda de Groot.

Stimuli

Stimuli consisted of emotional and neutral scenes selected from previously validated sets (bonobos: Kret et al., 2016, humans: Kret & van Berlo, 2021; van Berlo et al., 2021). While it is common in psychological research to use isolated facial expressions of emotions (see .e.g., Adolphs, 2002), we used a combination of expressions as well as emotional scenes. Emotional scenes can convey more contextual information, as they contain whole-body expressions that can communicate emotions as well as action intentions (De Gelder et al., 2010). Furthermore, previous studies have shown that emotional scenes modulate attention in a similar way as facial expressions (e.g., Kret et al., 2016; Kret & van Berlo, 2021; van Berlo et al., 2021), therefore indicating that they provide sufficient emotional information to the participant.

Emotional scenes involved individuals engaged in socially relevant behavior and/or having an emotionally relevant facial or bodily expression. Though it can be argued that we do not exactly know what bonobo emotions are, we do know the social relevance of certain facial expressions (such as the fear-grin, the relaxed open-mouth play face and yawning) and socio-emotional behaviors (sex, grooming) (De Waal, 1988). The fear-grin is often expressed during stressful situations and agonistic interactions, while the relaxed open-mouth face (or play face) is expressed during playful interactions (De Waal, 1988). Yawning is a widespread behavior in vertebrates and it is highly contagious (Demuru & Palagi, 2012; Massen et al., 2015; Palagi et al., 2014; Van Berlo et al., 2020b). Its contagiousness is linked to social closeness, and yawning could therefore serve a social function (Casetta et al., 2021; Norscia et al., 2020). Furthermore, yawns capture immediate attention in bonobos (Kret et al., 2016). Other socio-emotional behaviors that are relevant to bonobo society are sex and grooming. Bonobos use sex to prevent or resolve conflicts and reduce stress levels (De Waal, 1988). Grooming is an important social behavior used to form and strengthen social bonds between individuals (Dunbar, 1991). As such, emotional scenes in our task consisted of one or more bonobos playing, having sex or displaying an erection (male) or a large swelling (female), grooming, displaying distress, and yawning. Neutral scenes consisted of one or more bonobos lying down, sitting or walking with a neutral facial expression (see Tables S1 and S2 in supplements).

To make direct comparisons between bonobos and humans possible, we selected emotional scenes of humans that were equivalent to or an approximation of the emotional bonobo scenes. The stimuli consisted of humans playing, having sex (specifically: engaged in a romantic embrace), embracing ("grooming"), displaying distress (crying), and yawning. As there is no clear human equivalent for grooming in humans, we opted to use embracing as it is a reflection of social closeness and involves physical contact, just like grooming (Forsell & Åström, 2012). Neutral scenes of humans depicted one or more individuals lying down on grass, sitting, walking, or cycling with a neutral facial expression (see Table S3 for more information on the composition of the scenes). In total there were 10 unique stimuli per emotional category (5) and per species (2), as well as 100 unique neutral stimuli, as each emotional stimulus was matched with a neutral stimulus. Stimuli consisted of a subset of the validated sets by Kret et al. (2016) and van Berlo & Kret (2021). They were colored pictures with a dimension of 500x430 pixels, matched on luminance level and number of individuals depicted as much as possible.

Calibration

Before commencing testing, we conducted a manual two-point calibration using the infant calibration procedure in Tobii Studio. We used a relatively small number of reference points because apes tended to look only very briefly at the points. However, two-point calibrations are often used in great ape research as they are reasonably sufficient for the research questions asked, and also attainable given the constraints of working with animals (Hopper et al., 2021b). A small video displaying penguins (270x155 px) was used for the reference points. Calibrations were repeated until a sufficient calibration was obtained (i.e., Tobii Studio indicated no large calibration errors). For each individual, we continued using their first successful calibration throughout the entire experiment. To make sure that the calibration remained sufficient over time, we showed bonobos a 9-point grid before the start of each test session and visually inspected the accuracy of the calibration (see supplements for more information regarding calibration).

Procedure

Before commencing the experiment, bonobos were familiarized with the setup by showing each individual at least two sets of 10 trials with stimuli of animals and objects. Due to time constraints, once four individuals were able to drink from the setup during most of the practice sessions, we moved on to the experiment. Bonobos then participated in an experiment in which they could freely view socio-emotional and neutral scenes (presented at the same time) of unfamiliar conspecifics and of unfamiliar humans (Figure 2). Because the bonobos were not physically separated from other group members, the progression from trial to trial was manually controlled by the experimenter. This was done to ensure that data would only be collected when bonobos were attending the screen, and not when there were disturbances such as individuals moving away from the setup or individuals being distracted by others.

Each test session consisted of 10 trials and started with a 9-point grid to check for calibration accuracy, shown until the experimenter manually continued the experiment. The presentation of the grid was followed with a black screen displayed for 4 seconds, and subsequently followed by a fixation video (a sped-up nature movie) positioned in the middle of the screen. Only when the participant's fixation was on the video for more than one second, the experimenter moved on to the next trial sequence. Bonobos were then presented with two stimuli on the left and right side of the screen; one emotional and one neutral image (location was counterbalanced). Stimuli were presented for 3 seconds, in accordance with previous eye tracking tasks with great apes (e.g., Kano et al., 2015, and see Hopper et al., 2020, for a review). After 3 seconds, the experiment continued with a black screen shown for 4 seconds, and this concluded a trial. After 10 trials, the task ended automatically. Bonobos first completed all the trials with bonobos before moving on to the human stimulus set.

On average, the bonobos were tested on 33.5 sessions (*SD* = 3.12), and 355 trials (trials with bonobo stimuli: *M* = 191.5 , *SD* = 23.84; trials with human stimuli: *M* = 163.5, *SD* = 46.57). Furthermore, trials were repeated in order to compensate for data loss (e.g., due to disruptions by other bonobos). On average, all unique stimulus combinations were repeated 3.59 times (*SD* = 1.49).

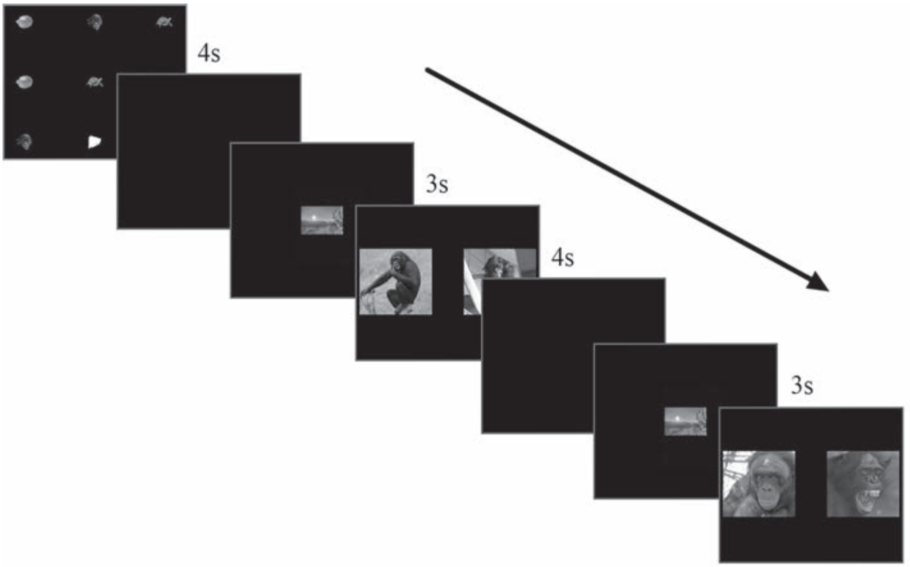

Figure 2. Trial sequence for participants. The test started with a 9-point grid, and each trial started with a black screen (4s), followed by a fixation video. Finally, two stimuli were shown on both sides of the screen (3s).

Data preparation

Because we used only one calibration per bonobo throughout the entire experiment rather than re-calibrating the bonobos for each experimental session, before analyzing the data, we checked whether the raw fixation data per bonobo and per session reasonably matched with the areas of the stimuli on the screen. We plotted all the gaze data for each individual onto a mapping of our screen and the location of the stimuli on the screen. We found that for two apes, in some sessions there were consistent shifts in gaze data to the left or to the right relative to the position of the stimuli on the screen.

Using K-means clustering in a custom script in Python, we established the difference between the gaze data collected by the eye tracker and the true centroids of the stimuli displayed on the left and right side of the screen. Based on these findings, we corrected 37/54 sessions for Monyama (average offset of +134 pixels), and 39/46 sessions for Zuani (average offset of -141 pixels) (see supplements for more information on how we corrected these sessions).

Next, two regions of interest (ROIs) were defined in Tobii Studio. We drew a 500x512 square around each of the stimuli (sized 500x430, thus the ROI was slightly larger in length than the stimuli to compensate y-axis inaccuracies in the gaze data; Figure S3). Through Tobii Studio's *Statistics* option, we extracted data on *Total Fixation Duration* per ROI using the Tobii Fixation Filter. Finally, after processing the *Total Fixation Duration* gaze data, we noticed that there were 19 trials where the total fixation duration was higher than 3 seconds (M = 4.47s, SD = 1.09), possibly due to Tobii registering a fixation that extended beyond the duration of the stimulus presentation. These isolated cases were removed from further analyses.

Statistical analyses

We used Bayesian mixed modeling in order to assess support for our hypotheses. We were interested in the total looking duration to emotional stimuli across trials. Our dependent variable was therefore the proportional looking duration to emotional stimuli (based on Tobii Studio's *Total Fixation Duration*. From here on: $\mathrm{PLD}_{\mathrm{emotion}}$), calculated by dividing the looking duration to the target by the sum of the looking duration to the target and distractor. The target was the emotional stimulus, and the distractor a neutral stimulus of the same species. A $\mathrm{PLD}_{\mathrm{emotion}}$ higher than 0.5 indicates a longer looking duration to the target.

Within a three-second trial, bonobos on average looked 1.93 (*SD* =0.78) seconds to the target and distractor combined (raw, unweighted values) when bonobo stimuli were displayed, and 2.04 seconds (*SD* = 0.72) when human stimuli were displayed. Thus, as some trials were more reliable than other trials, and to account for variation in overall attention to the target and distractor during the trial, we also calculated a weight for each trial. We calculated the weight by dividing the sum of the looking duration to the target and distractor by the average looking duration to the target and distractor per participant. The weight gives more importance to trials in which the participant paid more attention to the stimuli, and less importance to trials where participants were relatively inattentive. Weights were added to our models for all measures of interest (*M* = 1, *SD* = 0.38, range [0.01 – 1.69]).

We used zero-one-inflated beta (ZOIB) regression to account for 0's, 1's, and the data between the range [0 , 1]. For our measure of interest, proportional looking duration to emotional stimuli ($PLD_{emotion}$) across trials, we ran three separate models. In the first model, we examined whether the $PLD_{emotion}$ was higher than 0.5, i.e., whether participants look more than 50% of the time to emotional stimuli. In the second model, we assessed whether the $\mathrm{PLD}_{\mathrm{emotion}}$ differed between *Species* displayed on the stimuli (i.e., human or bonobo). In the third model, we zoomed in on the specific emotion categories and examined whether there was an interactive effect between *Species* and *Emotion Category* on the $\mathrm{PLD}_{\mathrm{emotion}}$.

In all of our models, we used weakly informative priors, specifically a student-t (default) prior (*df* = 3, *M* = 0, *SD* = 2.5) for the standard deviation coefficient, and a normal distribution (*M* = 0, *SD* = 1) for all other coefficients. *Species* and *Emotion Category* were treatment (dummy) coded. For each model, we report the median estimate coefficient, together with the 89% credible interval (either the Highest-Density Credible Interval [HDI; a "summary credible interval" for the posterior distribution] or the 89% Highest-Posterior Density [HPD; the shortest possible credible interval]). For comparisons between conditions, we report the odds ratio (*OR*). We also report the probability of direction (pd), which indicates the certainty that an effect goes in a specific direction (Makowski et al., 2019c, 2019a).

To establish model convergence, we followed the guidelines set out in the WAMBS checklist by de Paoli & van de Schoot (2017). We assessed trace and autocorrelation plots, the Gelman-Rubin diagnostic values (convergence indicated by a value close to 1), and density histograms for the posterior distributions. We conducted all of our analyses using RStudio (v. 1.4.1106, R Core Team, 2020) and the package brms (Bürkner, 2017, 2018).

Results

In model 1, we did not find evidence that the looking duration to emotional stimuli ($\mathrm{PLD}_{\mathrm{emotion}}$) of bonobos is higher than 50% (*Mdn* = 0.50, *89% CI* [0.46 – 0.53], pd_+= 58%, Table 1), meaning that bonobos did not reliably look longer at emotional stimuli of other bonobos and humans compared to neutral stimuli.

In our second model, we examined the effect of *Species* on the stimulus (bonobo or human) on $\mathrm{PLD}_{\mathrm{emotion}}$. We found that for both species, the $\mathrm{PLD}_{\mathrm{emotion}}$ did not reliably deviate from 50% (bonobo stimuli: *Mdn* = 0.52, *89% CI* [0.48 − 0.55], pd_+ = 82%; human stimuli: *Mdn* = 0.47, *89% CI* [0.44 − 0.51], pd_+ = 90%, Table 1). However, we found robust evidence for a difference between the $\mathrm{PLD}_{\mathrm{emotion}}$ for stimuli depicting humans and those depicting bonobos (*OR* = 1.17, *89% HPD* [1.09 − 1.26]); bonobos looked relatively longer to emotional stimuli of other bonobos than to emotional stimuli of humans.

In our third model where we zoomed in on the specific emotion categories, we found robust evidence for a longer $\mathrm{PLD}_{\mathrm{emotion}}$ of stimuli depicting distressed bonobos (*Mdn* = 0.54, *89% CI* [0.51 − 0.58], pd_+ = 96%, Table 1 and Figure 3a). For the sex category, the effect was in the expected direction (as indicated by the probability of direction; pd_+ = 93%), but not very strong (*Mdn* = 0.54, *89% CI* [0.50 – 0.57]). Finally, we found robust evidence for a lower $\mathrm{PLD}_{\mathrm{emotion}}$ of stimuli depicting humans having sex (*Mdn* = 0.40, *89% CI* [0.36 – 0.43], pd_+ = 100%, Table 1, Table S4, and Figure 3c).

Table 1. Overview of results per factor's level of interest for the three models. Robust effects are in bold.

Model	Species on stimulus	Emotion Category	*Median*	*89% CI*	*pd*
1 (Intercept)	Bonobo and human	All	0.50	0.46 – 0.53	0.58
2 (Species)	Bonobo	All	0.52	0.48 – 0.55	0.82
	Human	All	0.47	0.44 – 0.51	0.90
3 (Species*Emotion Category)	Bonobo	**Distress**	**0.54**	**0.51 – 0.58**	**0.96**
		Yawn	0.50	0.46 – 0.54	0.51
		Groom	0.49	0.45 – 0.53	0.61
		Sex	0.54	0.50 – 0.57	0.93
		Play	0.49	0.46 – 0.53	0.61
	Human	Distress	0.53	0.48 – 0.56	0.85
		Yawn	0.47	0.43 – 0.51	0.91
		Groom	0.49	0.45 – 0.53	0.73
		Sex	**0.40**	**0.36 – 0.43**	**1.00**
		Yawn	0.50	0.46 – 0.54	0.55

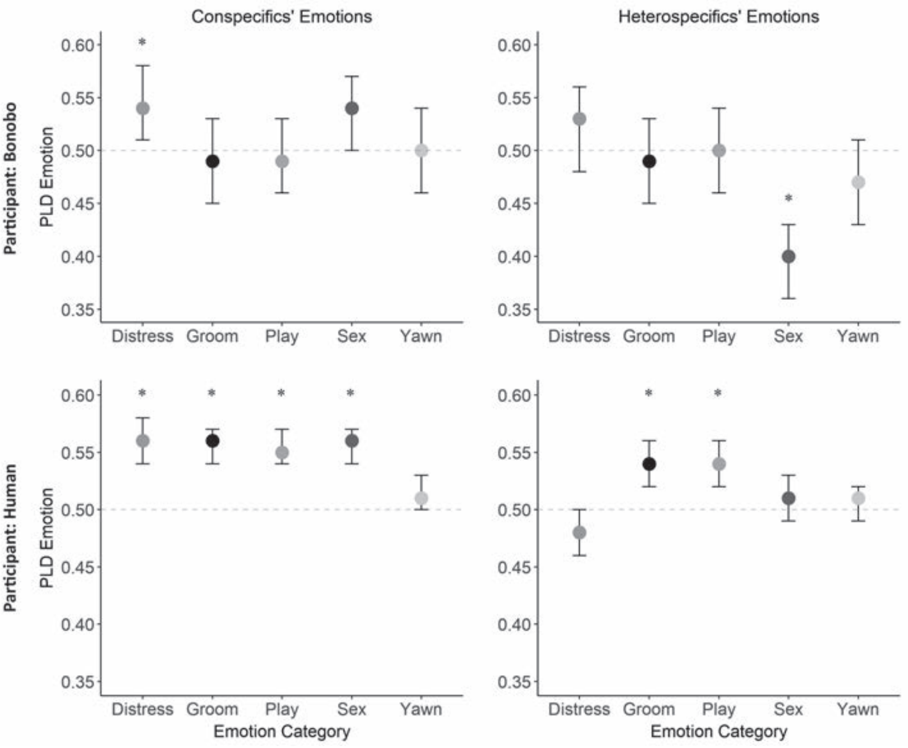

Figure 3. Graphs displaying the proportional looking duration to emotional stimuli (PLD Emotion) of conspecifics and heterospecifics by bonobos and humans. Error bars reflect the 89% credible interval, dots represent the median. Asterisks indicate robust effects.

Conclusion

Overall, we found that bonobos attended longer to emotional scenes of conspecifics (i.e., other bonobos) than to emotional scenes of heterospecifics (i.e., humans). When viewing emotional scenes of conspecifics, bonobos preferred to look at distressed others and sexual scenes compared to neutral scenes.

Experiment 2: Examining biased attention to emotions in humans

Method

Participants

Participants were visitors of primate park Apenheul. In total, 100 adults participated (Age category 18-30: $N = 57$, 31-50: $N = 33$, 51-80: $N = 9$; 58 women, 41 men). We tested participants in the visitor's area of the bonobo enclosure, where we set up a long table with cubicles in which we could test participants. We actively recruited participants by approaching them when they walked past the indoor bonobo enclosures and our setup. Participants were told that the bonobos participated in several experiments, and that we were now collecting human data using the same tasks. Data were collected between April and May 2017.

Stimuli

The same stimulus material was used as in Experiment 1. Like the bonobos, human participants saw both bonobo and human stimuli (see supplements Tables S2 and S3 for more information on the stimuli).

Equipment

Humans were tested near the indoor enclosures of the bonobos. We had a special corner dedicated to comparative research, consisting of two cubicles. One cubicle was specifically for this study. We tested participants using a 19" laptop (1920x1200 pixels) and a Tobii X2-60 eye tracker with Tobii Studio.

Calibration

Human participants were calibrated using the 5-point automated calibration procedure in Tobii Studio. Calibrations were accepted when the error displayed after finishing the calibration was minimal (less than a degree).

Procedure

Human participants were actively recruited by research assistants in the park. Visitors were approached and asked if they were interested in participating in a short, 10 minute task that was also completed by the bonobos. If visitors were interested, they were given a consent form to sign, thereby giving the experimenter permission to use their data for further analyses and publication. Participants then sat down behind the laptop and the experimenter started the 5-point calibration procedure. After finishing the calibration, participants filled in their age and sex in the task, and then the experimenter started the task. After finishing the task, participants were given the opportunity to ask more questions about the study and were given a debrief form containing the explanation and goal of the study.

To make direct comparisons between bonobos and humans possible, the difference in the task completed by both species was kept to a minimum. Whereas bonobos first completed all trials with bonobo stimuli and then trials with human stimuli, human participants first completed 10 trials of either bonobo or human stimuli, followed by 10 trials of the opposite species, and then followed by yet another 10 trial of the species that they started out with. Human participants thus completed 30 trials in one session. We created 10 versions of the task to control for order effects. In version 1, participants completed 10 trials with human scenes, followed by 10 trials containing bonobo scenes, and then another 10 trials with human scenes. In version 2 of the task, participants started with 10 trials with bonobo scenes, followed by 10 trials with human scenes, and again 10 trials with bonobo scenes. We continued alternating this sequence for the remaining 8 versions. We tested 10 participants per version of the task. This meant that every participant saw stimuli only once, but since we had 100 unique stimuli (50 combinations) and 10 versions of the task, every stimulus combination was repeated three times overall, resulting in 30 datapoints per stimulus combination.

In human participants, the trial sequence was fully automated. Because the bonobos could not be instructed, humans received minimal instructions as well, namely that they should pay attention to the screen and not move their head too much. Similar to the bonobo version of the task, humans started out with a 9-point grid that was shown for 3 seconds. The grid was followed by a black screen for 4 seconds, and then the fixation video for 3 seconds. Next, two stimuli of an emotional and a neutral bonobo or human were shown for 3 seconds, followed by a black screen shown for 4 seconds (Figure 2). Participants could take a short break between every set of 10 trials where they were allowed to move their head, but were requested to remain seated. When ready, participants could continue to the next 10 trials by pressing the space bar, followed by the 4 seconds black screen indicating the start of a new trial. At the end of the last set of 10 trials, participants saw a screen on which they were thanked for their participation.

Data preparation

After data collection finished, we realized that in version 3, 6, and 9 of the task, we accidentally showed one stimulus twice. These repetitions were removed from further analyses (31 datapoints). Furthermore, for five participants, there was a technical malfunction with the eye tracker resulting in 60% or more data loss. Thus, the data of these participants were excluded from further analyses. Similar to Experiment 1, we created ROIs in Tobii Studio, and extracted data on *Total Fixation Duration* per ROI using the Tobii Fixation Filter (see Experiment 1).

Statistical analyses

The analysis procedure for humans was similar to that of the bonobos. We were interested in the total looking duration to emotional stimuli across trials, thus calculated the proportional looking duration to emotional stimuli ($PLD_{emotion}$). Within the three-second trial window, human participants looked on average 2.66 s (*SD* = 0.38) to the target and distractor combined (raw, unweighted values) when human stimuli were displayed, and 2.64 s (*SD* = 0.43) when bonobo stimuli were displayed. Similar to what we did for the bonobos, we calculated the weight of a trial depending on how long a participant looked at the stimuli relative to their average looking duration to the stimuli (*M* = 1, *SD* = 0.15, range [0.005 – 1.51]).

For the $PLD_{emotion}$ across trials, we ran Bayesian zero-one-inflated beta regression models, similar to Experiment 1. Model 1 involved only the intercept, model 2 examined effects of *Species* displayed on the stimulus, and model 3 assessed an interaction effect of *Species* and *Emotional Category*. All models included a random intercept for *ID* (participant), and used weakly informative priors. Each model was checked using the WAMBS checklist (Depaoli & van de Schoot, 2017). We conducted all of our analyses using RStudio (v. 1.4.1106, R Core Team, 2020) and the package brms (Bürkner, 2017, 2018).

Results

In model 1, we found robust evidence for a longer $PLD_{emotion}$ in human participants (*Mdn* = 0.53, *89%CI* [0.52 – 0.54], pd_+ = 100%, Table 2), meaning that humans looked relatively longer to emotional stimuli than to neutral stimuli. In the second model with *Species* included as a factor, we found robust evidence for longer $PLD_{emotion}$ of stimuli depicting humans (*Mdn* = 0.55, *89% CI* [0.54 – 0.56], pd_+

= 100%, as well as those depicting bonobos (*Mdn* = 0.52, *89%CI* [0.51 – 0.52], pd_+ = 100%, Table 2). Additionally, we found robust evidence for a difference between $PLD_{emotion}$ to stimuli depicting bonobos or humans (*OR* = 0.88, *89% HDI* [0.84 – 0.91]). Thus, the looking duration to emotional stimuli was higher for human emotions than for bonobo emotions.

When examining the specific emotion categories per species in model 3, we found robust evidence that humans looked longer at other humans in distress (*Mdn* = 0.56, *89% CI* [0.54 – 0.58], pd_+ = 100%), having sex (*Mdn* = 0.56, *89% CI* [0.54 – 0.57], pd_+ = 100%, or humans playing (*Mdn* = 0.55, *89% CI* [0.54 – 0.57], pd_+ = 100%), grooming/embracing (*Mdn* = 0.56, *89% CI* [0.54 – 0.57], pd_+ = 100%, Table 2 and Figure 3b). For the yawning category, the effect was in the expected direction (pd_+ = 90%), but it was weak (*Mdn* = 0.51, *89% CI* [0.50 – 0.53]). For the bonobo category, we found robust evidence for $PLD_{emotion}$ of stimuli of grooming (*Mdn* = 0.54, *89%CI* [0.52 – 0.56], pd_+ = 100%) and playing bonobos (*Mdn* = 0.52, *89% CI* [0.52 – 0.56], pd_+ = 100%). We also find a weak effect of humans looking towards neutral scenes that were matched with distressed bonobos (*Mdn* = 0.48, *89% CI* [0.46 – 0.50], pd_- = 96%, Table 2, Table S5, and Figure 3d).

Table 2. Overview of results per factor level of interest for the three models. Robust effects are in bold.

Model	Species on stimulus	Emotion Category	*Median*	*CI 89%*	*pd*
1 (Intercept)	Bonobo and human	**All**	**0.53**	**[0.52 – 0.54]**	**1.00**
2 (Species)	Bonobo	**All**	**0.52**	**[0.51 – 0.52]**	**1.00**
	Human	**All**	**0.55**	**[0.54 – 0.56]**	**1.00**
3 (Species*Emotion Category)	Bonobo	Distress	0.48	[0.46 – 0.50]	0.96
		Yawn	0.51	[0.49 – 0.52]	0.78
		Groom	**0.54**	**[0.52 – 0.56]**	**1.00**
		Sex	0.51	[0.49 – 0.53]	0.87
		Play	**0.54**	**[0.52 – 0.56]**	**1.00**
	Human	**Distress**	**0.56**	**[0.54 – 0.58]**	**1.00**
		Yawn	0.51	[0.50 – 0.53]	0.90
		Groom	**0.56**	**[0.54 – 0.57]**	**1.00**
		Sex	**0.56**	**[0.54 – 0.57]**	**1.00**
		Play	**0.55**	**[0.54 – 0.57]**	**1.00**

Conclusion

In general, humans attended longer to emotional scenes compared to neutral scenes. This general emotion bias was also present for scenes of bonobos, although it was less pronounced. Humans tended to look longer at all types of emotional scenes involving humans, although evidence for a bias towards yawning was not robust.

Discussion

Emotions and their perception in non-human animals are intriguing, yet elusive (Anderson & Adolphs, 2014). To progress our understanding of when and how the brain evolved to efficiently process emotionally salient cues, we set out to study attention for emotions in our closest relatives, bonobos, and in humans. We found that both species preferentially attended to conspecific over heterospecific emotional scenes. Moreover, attention appeared to be strongly tuned to conspecifics in distress. Furthermore, bonobos showed an (albeit weak) attentional bias towards sex stimuli, while humans tended to look longer at emotional scenes across all categories. Below, we first discuss the findings in experiment 1, followed by a comparison between the results of humans (experiment 2) and bonobos.

In the first experiment, we partially confirmed our expectation that bonobos preferentially look at emotional scenes over neutral scenes of other bonobos and humans. Seeing distressed others can be a very salient cue, for instance, because detecting potential social or environmental threats can be crucial to survival (Öhman et al., 2001a). Similarly, in bonobos, socio-sexual interactions play a major role in preserving stability in the group (for instance to ameliorate tension) (Genty et al., 2015), and sexual stimuli may therefore receive enhanced attention. Bonobos showed no pronounced attention bias towards playful, grooming, or yawning scenes. These results are somewhat surprising, as a previous study found an implicit attentional bias towards scenes depicting yawning and grooming (in addition to sexual scenes) (Kret et al., 2016), and one study found that playful scenes interfered with bonobos' attention in an emotional Stroop task (Laméris et al., 2022). However, this could be explained by the results capturing different attentional processes, with reaction time paradigms possibly tapping into bottom-up attention, and eye-tracking paradigms having the potential to also measure top-down attention

(Belopolsky et al., 2011). Our data are not fine-grained enough to disentangle the two processes, as eye-tracking is not yet optimized for apes. However, given that bonobos do appear to have an immediate bias towards playful and yawning scenes, but not attend to them longer when given the opportunity to do so, these categories likely elicit a bottom-up attentional process. Future studies could focus on distinguishing between bottom-up and top-down attentional processes, especially now that new eye-tracker models allow for greater sampling rates and are more forgiving in terms head movements (which is important when working with animals).

We expected a similar (but less pronounced) attentional bias pattern when the bonobos viewed emotional scenes involving humans. Although we did not find robust evidence that bonobos looked longer at any of the human emotions compared to neutral scenes, the looking duration pattern was similar to viewing bonobo scenes. Specifically, bonobos seemed to look slightly longer at humans expressing distress. These findings may be explained by human expressions of distress sharing similar morphological action tendencies as bonobo expressions of distress. For instance, a general feature of fearful expressions is the tendency to make oneself small, indicating weakness or submissiveness, and this occurs in humans and many other primates (Kret et al., 2020). Moreover, the scream face of apes shares a lot of morphological similarities with its human equivalent (Parr et al., 2007). Furthermore, the finding that bonobos looked longer to the neutral scenes that accompanied a human sex scene is curious. In the neutral scenes, people wore more clothing, which may provide a salient cue (e.g., due to more variation in patterns) than seeing people without clothes in the sex scenes (Van Renswoude et al., 2019). Finally, as scenes showing bonobos engaged in play or a grooming bout did not hold attention longer than neutral scenes, the human variant of these scenes is likely also not very salient to bonobos.

In experiment 2 with human participants, we found an overall preference for viewing emotional scenes over neutral scenes, and with human emotional scenes receiving slightly more attention than bonobo scenes. Humans showed the most pronounced effect in the distress category, with longer looking durations towards distressed conspecifics compared to the neutral scenes. Moreover, humans also preferentially looked at individuals that were embracing each other, playing, or having sex. An implicit attentional bias for threatening signals has been studied a great deal in humans. Most studies indicate that in highly anxious individuals, attention to negative or threatening stimuli is strongly prioritized (Armstrong & Olatunji, 2012). Results on non-anxious individuals are mixed, showing that an implicit bias towards positive emotional expressions also occurs (Becker et al., 2011). In a previous study, we found an implicit attentional bias towards stressful scenes in a heterogeneous human population, as well as to scenes involving sex and yawning (Kret & Van Berlo, 2021). Here, we add to the existing literature by showing that emotional scenes also spontaneously hold attention for longer durations in a task without a clear goal to the participants, and even when a competing social, but emotionally-neutral stimulus is present.

Interestingly, the attentional pattern of humans for human emotional scenes differed from that for bonobo emotional scenes. Humans looked longer at bonobos engaged in grooming and play compared to neutral scenes, but not at sex or yawning scenes, even though we found an effect for these two categories within the human scenes. Furthermore, we found weak evidence that humans looked longer at neutral scenes rather than bonobo distress scenes; the opposite from what we found for distress scenes of humans. In a previous study, adults rated distress scenes of bonobos as negative and highly arousing (similar to ratings of distress scenes of chimpanzees (Kret et al., 2018)), possibly due to canine visibility (Kret & Van Berlo, 2021). In our study, participants may have looked away from the distress scenes because they are intense in terms of emotional arousal, but this remains speculative. To date, very little work has examined how humans view (other) emotional expressions of primates (see e.g., Kret & Van Berlo (2021); Maréchal et al. (2017)). As such, future work on attentional biases could benefit from including questionnaires that measure participants' interpretation of and feelings towards the stimuli.

Compared to our bonobo sample, humans appear to preferentially sustain attention to emotional scenes across all categories. A possible explanation for this difference is that humans have evolved exceptionally distinctive and exaggerated communicative faces in order to communicate more effectively (Kret et al., 2020), and therefore also have a sensitivity to a wider range of expressions. Nevertheless, alternative explanations, particularly relating to our methodology, must be considered.

We report several limitations to our study. First, we used static images of emotional expressions instead of dynamic scenes. Studies have suggested that the dynamic facial expressions of emotion provide richer information than static expressions, causing stronger activation in brain regions associated with emotion recognition (Arsalidou et al., 2011). Second, we made use of more complex social and emotional scenes rather than isolated facial expressions, potentially providing more contextual information. However, it is possible that by providing this context, we increase the complexity of the stimuli, therefore making the interpretation of the stimuli more ambiguous (Tottenham et al., 2013). A combination of these two interpretations may explain our bonobo results in that our stimuli may underrepresent the interest bonobos have in emotionally-salient information. Nevertheless, it is important to note that humans do show an emotional bias across all the categories of emotions even with a similarly prepared stimulus set. Moreover, in a follow-up experiment where we zoomed in on facial expressions rather than scenes (as well as investigating effects of expression channels such as face vs. body), an emotional bias was not observed (in prep, (Kim et al., 2021)). At the moment, It is difficult to know how bonobos interpret emotional images and whether emotional scenes are better at providing more salience than isolated faces. Future research could use dynamic emotional cues using videos or a combination of images with sound, as this has previously proved to be successful in uncovering an emotion bias in for instance chimpanzees (Kano & Tomonaga, 2010a).

Another limitation of our study is the small sample size. Moreover, we were only able to test female bonobos. The reason for this is that bonobos are rarely found in zoos and face a high risk of extinction (Fruth et al., 2016), and even fewer are accessible for research purposes. As such, we cannot extrapolate our findings to the entire species. Nevertheless, our results convergence with a small, but growing body of experimental studies indicating that bonobos and other apes are sensitive to the emotional cues of others (Kano et al., 2016; Kret et al., 2016; Laméris et al., 2022; Pritsch et al., 2017; Van Berlo et al., 2020a), and showing that bonobos have remarkably well-developed brain structures that are important for emotion processing (Issa et al., 2019; Stimpson et al., 2016).

Perceiving emotions in others is at the foundation of more complex socio-cognitive abilities such as cooperation and empathy (Levine et al., 2018). Our findings show that bonobos, like humans, voluntarily look longer at emotionally salient signals such as distress and sex. Our findings converge with previous studies, suggesting that the groundwork for higher social cognition is likely shared with our closest living relatives.

Part II:

Spontaneous Mimicry

Chapter 5

Low relationship quality predicts self-scratch contagion during tense situations in orangutans

Abstract

Primates show various forms of mimicry that are stronger between kin and friends. As a result, mimicry is thought to promote group coordination, social cohesion, and possibly state matching (or: emotion contagion). Aside from contagious yawning, little is known about the contagious effect of other behaviors. Self-scratching is commonly observed during arousal and as such may play a role within group dynamics. While the Bornean orangutan (Pongo pygmaeus) is commonly considered the least social great ape, orangutans do engage in social interactions. Therefore, their social organization makes them a suitable case for studying the social function of mimicry. Through behavioral observations of captive orangutans, we recorded all yawn and self-scratch events together with the corresponding behavior of all bystander group-members. As yawning was rarely observed, no conclusions could be drawn regarding this behavior. Self-scratching was contagious and occurred within 90 s after the triggering self-scratch. Specifically, orangutans showed increased self-scratch contagion when they had seen a weakly bonded individual self-scratch during tense contexts. When the orangutan had not seen the triggering self-scratch, the contagiousness of self-scratching was not affected by context or relationship quality. Our results indicate that mimicry is not simply higher between individuals with stronger social relationships, but that the contagiousness of behaviors may vary based on the context and on social factors. We discuss these findings in light of an adaptive function that may reduce aggression.

Based on:
Laméris, D. W., Van Berlo, E., Sterck, E. H. M., Bionda, T., & Kret, M. E. (2020). Low relationship quality predicts scratch contagion during tense situations in orangutans (*Pongo pygmaeus*). *American Journal of Primatology, 82,* e23138. https://doi.org/10.1002/ajp.23138

Introduction

Mimicry is a phenomenon in which a behavior is automatically triggered by the perception of a similar behavior in others (Zentall, 2003). From a proximate perspective, such contagion of behaviors can be explained by mechanisms rooted in primitive forms of state matching and empathetic processing (Joly-Mascheroni et al., 2008; Palagi et al., 2009). The perception-action mechanism explains that if such behaviors are manifestations of emotions, mimicry can result in emotional state-matching, a phenomenon known as emotional contagion (Preston & De Waal, 2002). However, mimicry can also be explained more parsimoniously as the nonconscious copying of a partner's behavior (Massen & Gallup, 2017; Yoon & Tennie, 2010). Interestingly, forms of mimicry are commonly found to be stronger between kin and friends (Campbell & de Waal, 2011; Demuru & Palagi, 2012; Massen et al., 2012; Palagi et al., 2009, 2014). Such enhanced mimicry between individuals that share social connections is thought to facilitate group coordination and social cohesion (Clay & De Waal, 2013; Lakin et al., 2003; Preston & De Waal, 2002; Prochazkova & Kret, 2017).

Probably the most well-studied behavior within the mimicry literature is yawning. While spontaneous yawning (i.e., nonsocial yawning) is widespread across vertebrates and may function in promoting cortical arousal (Baenninger, 1997; Guggisberg et al., 2010; Vick & Paukner, 2010), and/or changing emotional states through decreasing brain temperature (Gallup & Gallup, 2008; Massen et al., 2014; Massen & Gallup, 2017), contagious yawning is restricted to fewer species in which this trait may have evolved independently (Massen & Gallup, 2017).

Thus far, contagious yawning is observed in several primate species, including chimpanzees (*Pan troglodytes*; Anderson et al., 2004; Campbell et al., 2009; Campbell & de Waal, 2011; Massen et al., 2012), bonobos (*P. paniscus*; Demuru & Palagi, 2012; Palagi et al., 2014), and gelada baboons (*Theropithecus gelada*; E. Palagi et al., 2009). Other species in which contagious yawning is observed include domesticated dogs (*Canis lupus familiaris*; Joly-Mascheroni et al., 2008; Madsen & Persson, 2013), wolves (*C. lupus lupus*; Romero et al., 2013, 2014), budgerigars (*Melopsittacus undulates*; Gallup et al., 2015; Miller et al., 2012), and elephant seals (*Mirounga leonina*; Wojczulanis-Jakubas et al., 2019). However, some experimental studies have failed to provide convincing evidence for yawn contagion in orangutans (*Pongo abelli*) and gorillas (*Gorilla gorilla*; Amici et al., 2014), stump-tailed macaques (*Macaca arctoides*; Paukner & Anderson, 2006), ring-tailed lemurs (*Lemur catta*) and ruffed lemurs (*Varecia variegata*; Reddy et al., 2016), dogs (Harr et al., 2009), and red-footed tortoises (*Geochelone carbonaria*;

Wilkinson et al., 2011). This illustrates the ongoing debate on the possible mechanism underlying contagious yawning.

Although not receiving as much attention as contagious yawning, self-scratching may be another interesting behavior for contagion studies. Self-scratching is commonly associated with the presence of psychological and physiological stress (Maestripieri et al., 1992; Schino et al., 1991; Troisi, 1999). For example, increased self-scratch rates have been reported during aggressive interactions (Palagi & Norscia, 2011), post-conflict interactions without reconciliation (reviewed in Aureli et al., 2002), dominance-related interactions (Kaburu et al., 2012; Peignot et al., 2004), and predation attempts (Palagi & Norscia, 2011). Concurrently, self-scratching behavior is reduced after play bouts (Norscia & Palagi, 2011b), during affiliative interactions (Aureli & Yates, 2010), and after reconciliation following aggressive interactions (Aureli et al., 1989). However, a recent study also found that self-scratching increases with positive arousal (e.g., during play bouts), suggesting that it may be a marker of general emotional arousal, rather than an indicator of negative emotions specifically (Neal & Caine, 2016).

Apart from benefits for the expresser (Koolhaas et al., 1999), self-scratching potentially signals arousal to other group-members (Bradshaw, 1993). In rhesus macaques (*Macaca mulatta*), for example, self-scratching reduces the likelihood of subsequent aggression and increases the chance of affiliative interactions (Whitehouse et al., 2017). Furthermore, stressed individuals are a potential threat to group-members as they tend to behave unpredictably (Aureli et al., 1992). As such, the recognition and acquisition of the emotions of aroused individuals can result in fewer costly interactions (Whitehouse et al., 2016). While these studies suggest that self-scratching may play an important role within social groups, the contagious effect of self-scratching and its potential function is poorly understood.

Most studies on mimicry in great apes focused on bonobos and chimpanzees, probably because of their complex social structures, advanced cognitive capacities, and evolutionary proximity to humans (Maclean, 2016). However, the orangutan too, is one of our closest living relatives with highly developed cognitive skills (Damerius et al., 2019; Van Schaik et al., 2003), yet is considered semi-solitary as it does not live in stable social groups (Delgado & Van Schaik, 2000; Galdikas, 1985; Mitra Setia et al., 2009; Singleton et al., 2009; Van Schaik, 1999). Nonetheless, orangutans still form temporary parties for social reasons, e.g., for mating opportunities, protection from male coercion, and socialization opportunities for infants (Mitani et al., 1991; Mitra Setia et al., 2009; Singleton et al., 2009; Van Schaik, 1999). Furthermore, zoo-housed orangutans show increased frequencies of social behavior, including agonistic interactions (Edwards & Snowdon, 1980; Tajima & Kurotori, 2010; Zucker, 1987). This suggests that orangutans show a certain degree of behavioral flexibility under social contexts which makes them an interesting case for a study on mimicry and its possible social function.

Research on mimicry in orangutans, however, is scarce. One study found that orangutans show rapid facial mimicry during play events (Davila-Ross et al., 2008), while another study did not find evidence of yawn contagion in an experimental setup (Amici et al., 2014). In this study, we aimed to enhance our understanding of the function of mimicry in the orangutan. To do so, we focused on yawning as this behavior is commonly studied in mimicry research. In addition, we decided to focus on self-scratching behavior because of its possible link to arousal (Elder & Menzel, 2001). As such, we recorded all yawning and self-scratching events in a group of zoo-housed Bornean orangutans (*P. pygmaeus*) with the aim to investigate whether (a) yawning and self-scratching is contagious and (b) whether contagion has a social function in this species. Based on a previous study reporting the presence of rapid facial mimicry (Davila-Ross et al., 2008), we hypothesize that mimicry is present in orangutans in the forms of yawn and self-scratch contagion. Furthermore, if these behaviors have a social function, we expect that the contagion of yawning and self-scratching will be influenced by the relationship quality of the expresser and observer and that contagion is higher between kin and friends.

Method

Participants and data collection

Behavioral data were collected from February to May 2017 on nine adult Bornean orangutans (three males and six females, mean age= 23.2, range= 7–52 years old, all born and raised in captivity. See Table S1) living in primate park Apenheul, The Netherlands. The animals were housed in a building consisting of four indoor enclosures that were each connected to outdoor islands. The four enclosures could be disconnected from and connected to two adjacent enclosures, which allowed the zookeepers to alter group composition on a daily basis, based on the animals' preferences.

Usually, there were four separate groups (ranging from one to four individuals) that differed in composition and occasionally three groups (ranging from two to five individuals). This housing environment aims to mimic the natural social structure of orangutans in which they form temporary parties but no stable social groups. Some individuals were never housed together to avoid conflict (e.g., the two adult males). Focal-animal sampling of 10 min sessions was used to score behavioral patterns including social behaviors (e.g., grooming, agonistic interactions, and sexual behaviors), locomotion (e.g., walking and climbing), and food-associated behavior (e.g., foraging and feeding; ~18.5 hr per focal; see Table S2 for the ethogram). We used all-occurrence sampling to record all yawning and self-scratching events of group-members in the subgroup of the focal animal for 165 hr in total (Altmann, 1974).

Observations were performed by one trained researcher from the visitor's area in both indoor and outdoor enclosures. Due to the relatively low temperatures during the observation period, the orangutans were kept inside and as such most observations were performed in the indoor enclosures. The indoor enclosures were ~60 m2 in which observation conditions were excellent; the researcher had full view of the enclosure and its individuals as there were no big constructions blocking the line of sight. In addition, because subgroups had a maximum of five individuals, and because yawning and self-scratching could be considered 'attention-attracting' behaviors (Demuru & Palagi, 2012), it was possible for the researcher to record all yawning and self-scratching events.

The following variables were recorded whenever a yawn or scratch occurred: (a) time of occurrence; (b) identity of the expresser; (c) identity of all possible observers (i.e., individuals that were within the same enclosure); (d) presence/absence of a contagious response (i.e., a congruent behavior) within 3 min following the last triggering event (i.e., a spontaneous yawn or scratch. We selected a 3 min window in line with other work, and because a longer window would introduce the chance of 'self-contagion', meaning your own yawns or scratches cause you to perform these behaviors again (Gallo et al., 2021)) ; (e) time latency in contagious response measured in seconds (s); (f) duration of self-scratching behavior (short; 5 s); (g) if the observer could see the triggering event or not, based on the facial direction of the observer; (h) estimated distance between the expresser and observer (10 m); and (i) the context in which the triggering event occurred, categorized as 'tense' or 'relaxed'. The context categorization was based on the behavior of the expresser before and after the yawning or self-scratching behavior. Behaviors that indicated tension included display behavior (e.g., charging and shaking of climbing structures), high arousal vocalizations (long-calls or kiss squeaks), or agonistic behaviors (direct aggression and chasing). Because we rarely observed agonistic interactions, we consider yawning and self-scratching to be related to levels of increased arousal, but not aggression. Relaxed contexts were characterized by behaviors such as foraging, resting, or socio-positive interactions (e.g., grooming). To ensure the reliability of our data, we restricted our data set to the indoor observations and excluded cases for which the expresser and observer were at a greater distance than 10 m.

The care and housing of the orangutans was adherent to the guidelines of the EAZA Ex situ Program. Only observational data were collected, therefore there was no need for the approval of the Ethics Committee of Apenheul. The study complied with the requirements of the Dutch Animal Care and Use Committee and conformed to the American Society of Primatologists Principles for the Ethical Treatment of Non-Human Primates.

Relationship quality

Scan-sampling was performed every 30 min to score allogrooming, contact sitting, social play, and sexual behaviors (e.g., mounting and genital contact) to calculate relationship quality with a corrected composite sociality index (CSI; Silk et al., 2006). Relationship quality was based on two levels: kinship and CSI (Demuru & Palagi, 2012; Palagi et al., 2014). Regarding kinship, only maternal lineages were considered ($r = .5$), resulting in four dyads. However, the dyad involving a juvenile male was excluded from the analyses and only three kin dyads remained. One of these dyads was a mother that, in the past, already had an offspring and took on the role of surrogate mother for another juvenile of the same age as her own.

The CSI is a useful measure for scoring how much the positive relationship of a particular dyad deviates from the average of all dyads. We calculated it by summing up the following variables: frequency of grooming for a specific dyad divided by the mean frequency of grooming of all dyads, frequency of contact sit for a specific dyad divided by the mean frequency of contact sitting of all dyads, and frequency of days spent in the same compartment for the dyad divided by the total number of observation days. The summation of these variables was further divided by 3 (as per (Micheletta et al., 2013)). Since group composition for the orangutans was regularly changed and based on the preferences of the orangutans, we corrected for the total number of days spent together per dyad.

With the CSI score, we could identify high and low relationship qualities (Silk et al., 2006). Dyads with CSI scores in the top quartile were considered to have a high relationship quality, $N = 5$ (Demuru & Palagi, 2012), which included the kin dyads. Because of the low number of kin dyads, we did not separately test the influence of kinship on the degree of contagion. All other dyads were considered to have a low relationship quality, $N = 9$.

Statistical analyses

Yawn and scratch rates were extracted for two conditions: the baseline condition and the contagious condition. The baseline condition included spontaneous yawn and scratch events (i.e., when subgroup-members did not show yawning or self-scratching) which were extracted from the focal- animal observations. The contagious condition included those yawn/ scratch events that occurred within a 3-min period after a congruent triggering behavior, hence after spontaneous yawning/self-scratching behavior. By means of all-occurrence sampling, a total of 95 yawn and 597 scratch events were recorded. We had insufficient data to statistically analyze yawn contagion (baseline $N = 52$ and contagion $N = 4$) and therefore focused on the contagiousness of self-scratching. To test the data for normality, the Shapiro-Wilk test was used and Levene's test for equality of variances was used to test for homoscedasticity.

The use of long timeframes to study contagious responses have been discussed (Massen & Gallup, 2017). For this reason, we investigated the temporal boundaries of scratch contagion (i.e., during which time period following a triggering scratch of a group-member were scratch rates higher as compared to scratch rates observed during baseline). As such, we divided the scratch rates during the 3 min contagious condition into six intervals of each 30 s and calculated individual contagious scratch rates for each of the six 30 s intervals. In addition, for each individual, we calculated one baseline scratch rate per 30 s (i.e., number of spontaneous scratches per 30 s, derived from the focal sampling data). Due to the small sample size, we used bootstrapped paired samples t tests to compare each 30 s interval in the contagious condition to their matched 30 s baseline scratch rate. We employed Bonferroni corrections to adjust for multiple comparisons with the 30 s baseline scratch rate. From this, we found that contagious scratch rates were only higher than baseline scratch rates during the first three intervals (i.e., the first 90 s after a triggering scratch; Figure S1). Therefore, we only considered those scratches happening within 90 s after a triggering scratch as contagious and excluded the scratches that occurred after 90 s ($N = 37$). We then pooled the contagious scratches that occurred within 90 s together and calculated individual scratch rates during this period. We also calculated a baseline scratch rate per 90 s and compared this to the contagious scratch rates using a bootstrapped paired samples t test.

We created a generalized linear mixed model (GLMM) that included the identity of the expresser and observer as random effect and “context” (categorical; tense vs. relaxed), “relationship quality” (categorical; high versus low relationship quality) as fixed factors to test their effect on the occurrence of scratch contagion. Furthermore, we decided to include “seeing the triggering scratch” (categorical; seen vs. unseen) as additional fixed factor since auditory cues of self-scratching can already be sufficient to induce a contagious response in humans (Swithenbank et al., 2016). We included a three-way interaction for context, relationship quality and seeing the triggering scratch because we hypothesized that contagious responses triggered by unseen scratches would not be influenced by relationship quality, simply because the observer did not have information about the expresser. Sex of the expresser, observer, and sex combination were considered as additional fixed factors, but due to the low sample sizes (three males and six females), we decided to leave them out. The models used a binomial distribution (contagion or no contagion) and a logit link function. Likelihood ratio tests and a χ^2 distribution were used to compare the full model with the null model. Multicollinearity between independent variables was tested and variables with a variance inflation factor (VIF) of > 5 were rejected from the model (O’Brien, 2007). None of the factors showed high VIF values. Analyses were conducted using R version 3.6.1 (R Core Team, 2020), with the GLMM calculated using the lme4 package (Bates et al., 2015).

Results

Orangutans are susceptible to scratch contagion

We compared the scratch rates during the baseline condition with the scratch rates in each of the 30 s intervals during the contagious condition. Orangutans scratched more during the first 90 s after a triggering scratch (Figure S1; bootstrapped paired samples *t* test: Baseline vs. 0–30 s: $p < .001$; Baseline vs. 31-60 s: $p < .001$; Baseline vs. 61-90: $p = .002$). Furthermore, the scratch rates over the 90 s contagious condition were higher than the 90 s baseline condition (Figure 1; bootstrapped paired samples *t* test: $p < .001$). This suggests that only those scratches happening within 90 s after another scratch can be considered contagious.

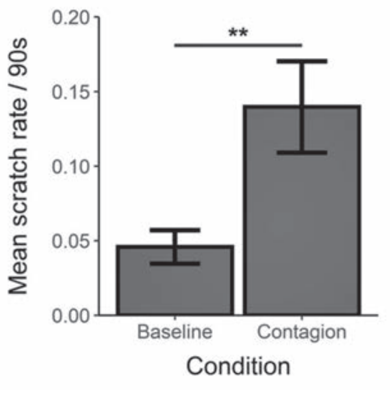

Figure 1. Mean scratch rates (±*SEM*) per 90 s in the baseline and contagion condition. *SEM*, standard error of mean. $^{**}p < .01$

Factors influencing scratch contagion

We further assessed potential factors explaining the occurrence of scratch contagion. Overall, the full model fitted the data better than the null model, as the likelihood ratio test (LRT) revealed a significant effect of the predictors on the occurrence of contagious self-scratching (LRT: $\chi^2_7 = 16.29$, $p = .023$). We found a significant interaction between whether the triggering scratch was seen or not, context and relationship quality (Table 1). Specifically, we found no difference in scratch contagion between context and relationship quality when the observer had not seen the triggering scratch. However, using simple contrasts, we found that during tense contexts, scratch contagion is more likely to occur between individuals that share a low relationship quality when the observer had seen the triggering scratch compared with when the observer had not seen the scratch (Figure 2; $z = 3.62$, $p < .001$). Furthermore, when only considering the cases where the observer had seen the triggering scratch, we found that self-scratching is more contagious between individuals that shared a low relationship quality during tense contexts compared with relaxed contexts ($z = 2.30$, $p = .021$) and during tense context between individuals that shared a low relationship quality compared with a high relationship quality ($z = 2.35$, $p = .019$). Follow-up analyses suggest that this effect is not a by-product of increased visual attention towards individuals with a low relationship quality as more scratches were observed when the expresser and observer shared a high relationship quality ($\chi^2_1 = 17.87$, $p < .001$).

It is possible that the increased scratch rates during tense context do not reflect contagion, but are simply a by-product of increased arousal levels during tense contexts in general (Castles & Whiten, 1998). Follow-up analyses revealed that contagious scratch rates did not differ between tense and relaxed context (bootstrapped paired samples *t* test: $p = .795$), suggesting that the observed effect of context is not just a by-product of increased self-scratching due to increased stress levels during tension.

Table 1. Type III tests for fixed effects on the occurrence of scratch contagion

	Estimate	SE	χ^2_1	p
Intercept	−1.897	0.38	24.864	*<.001*
Context (tense)	0.088	0.457	0.038	0.846
Relationship quality (low)	−0.228	0.428	0.283	0.595
Seen/unseen (seen)	0.24	0.418	0.33	0.566
Context * relationship quality (tense * low)	−0.576	0.725	0.631	0.427
Context * seen/unseen (tense * seen)	−0.653	0.956	0.466	0.495
Relationship quality * seen/unseen (low * seen)	0.384	0.675	0.324	0.569
*Context * relationship quality * seen/unseen (tense * low * seen)*	2.869	1.334	4.627	*0.032*

Note: GLMMs were used with a binomial distribution and logit link function. Effects with $p < .05$ are depicted in italics. Abbreviations: GLMMs, generalized linear mixed models; *SE*, standard error.

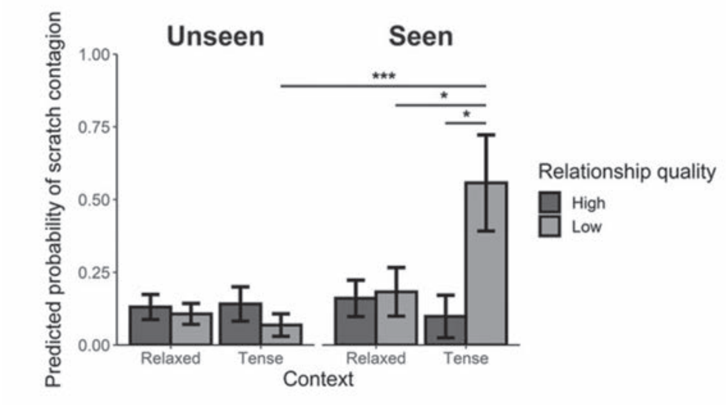

Figure 2. Predicted probability of scratch contagion (±*SEM*) based on the three-way interaction between seeing the triggering scratch, context and relationship quality. *SEM*, standard error of mean. $^*p < .05$; $^{***}p < .001$

Discussion

The contagion of behaviors such as yawning and self-scratching and their possible social function remain poorly understood. The current study aimed to investigate whether yawning and self-scratching are contagious in the orangutan and whether the contagion of these behaviors is linked to the context in which these behaviors occur and the quality of the bond between individuals. Orangutans showed increased self-scratch rates after a group-member scratched, indicating mimicry. This effect was visible within the first 90 s after the triggering scratch. Furthermore, when the relationship quality between the expresser and observer was low, and the observer had seen the triggering scratch, scratch contagion was more likely to occur during tense situations.

Our observation that scratch contagion is stronger in a tense context between weakly bonded individuals is novel, as most other studies report increased mimicry between individuals with a high relationship quality (Campbell & de Waal, 2011; Demuru & Palagi, 2012; Massen et al., 2012; Palagi et al., 2009, 2014). Yet, these studies predominantly looked at yawn contagion for which the social function and emotional load is debated and for which it is unknown how others perceive this behavior (Gallup, 2011; Massen & Gallup, 2017; Palagi, Celeghin, et al., 2020). Self-scratching, on the other hand, is often associated with physiological and psychological stress (Maestripieri et al., 1992; Schino et al., 1996; Troisi et al., 1991) although there is growing evidence that self-scratching also increases during positive arousing events, such as during play bouts (Neal & Caine, 2016). Without further measures (e.g., changes in emotional valence with cognitive bias testing as done by Adriaense et al. (2019) and Saito et al. (2016), we cannot conclude which emotions underlie self-scratching and if scratch contagion is truly linked to emotional contagion. Nonetheless, emotional contagion consists of simpler processes such as behavioral and physiological contagion (Edgar & Nicol, 2018) and the reported link between self-scratching and emotional arousal may suggest that the observed contagious effect of self-scratching in this study is a behavioral manifestation of emotional contagion.

If self-scratching is indeed an expression of emotional arousal, then this behavior could serve as a social cue for others (Laidre & Johnstone, 2013). Some other studies have reported on the potential signaling function of self-scratching. For instance, recent studies show that self-scratching can be used as a signal to coordinate joint travel, for example, between a mother and infant (Fröhlich et al., 2016, 2019a; Hobaiter & Byrne, 2014), and may be used to initiate grooming (Hobaiter & Byrne, 2014). Another possible communicative function of self-scratching is to signal social distress, which in turn reduces the likelihood of receiving aggression (Whitehouse et al., 2017). In our study, it is possible that orangutans use self-scratching in others as a marker of arousal and that the automatic contagion of such information from weakly bonded individuals during tension has an adaptive value. There was no difference in the probability of scratch contagion between contexts and relationship quality when the orangutan had not seen the triggering scratch, and hence only had auditory cues of this behavior. This can be explained by the fact that the observer had no information about the identity of the initial scratcher which may further highlight a possible link between contagious self-scratching and a social function.

If self-scratching indeed serves as a social signal (Fröhlich et al., 2016, 2019a; Hobaiter & Byrne, 2014), it is likely intended to change the behavior of the observer with the ultimate goal to benefit the expresser (Bradshaw, 1993; Laidre & Johnstone, 2013). A similar function of self-scratching is observed during agonistic interactions, where self-scratching rhesus macaques are less likely to receive aggression (Whitehouse et al., 2017). Because stressed individuals often behave unpredictably (McEwen & Sapolsky, 1995) they can become a potential social stressor (Aureli et al., 1992), especially when they are nonfriends or non-kin (Whitehouse et al., 2017) . Hence, increased awareness of such individuals through mimicry may be beneficial for observers and adaptive within group dynamics. Although such an adaptive function of mimicry warrants further investigation, we speculate that orangutans can benefit from increased self-scratch contagion, and potentially contagion of arousal, of weakly bonded individuals during tense contexts, as it may help individuals to prepare for potential unpredictable behaviors of the expresser. This way, scratch contagion becomes adaptive for both the expresser and observer by increasing social cohesion through reducing possible aggression (Rauchbauer et al., 2016). While we could not test such aggression-reducing hypothesis of scratch contagion, this would be interesting to explore in more detail.

It is important to recognize that increased scratch rates have often been observed during tense situations in general, independent of the identity of the individual providing the triggering scratch (Castles & Whiten, 1998; Kaburu et al., 2012; Palagi & Norscia, 2011; Peignot et al., 2004), although there are a number of studies that actually do not find increased scratch rates during anxiety-provoking circumstances (Aureli & De Waal, 1997; Duboscq et al., 2014; Judge et al., 2006; Pearson et al., 2015). Hence, it is essential to rule out that the heightened scratch contagion between weakly bonded individuals during tense contexts is not merely a by-product of increased arousal during these contexts. If this were the case, we would expect increased chances of scratch contagion during tense contexts regardless of the relationship quality and whether the triggering scratch was seen or not. This was not the case (see Figure 2). As such, it seems unlikely that the increased contagion observed in our study is a by-product of higher scratch rates induced by tension, but that it is truly an effect of the context and the relationship quality between the expresser and observer.

In conclusion, this study is the first to provide evidence for the presence of scratch contagion in the orangutan, possibly suggesting a case of emotional contagion. We show that scratch contagion is stronger between weakly bonded individuals when there is tension, demonstrating that it has a possible social function. Our results are relevant for future research on mimicry and its link with emotional contagion as they highlight that contagion is not simply stronger between individuals with a high relationship quality, as is commonly suggested. Furthermore, the variety of contexts in which self-scratching is observed throughout the literature highlight the complexity of this behavior and the mechanism underlying its contagious effect. Importantly, the degree of scratch contagion may depend on the interaction between contextual factors and social relationships.

Chapter 6

Experimental evidence for yawn contagion in orangutans

Abstract

Yawning is highly contagious, yet both its proximate mechanism(s) and its ultimate causation remain poorly understood. Scholars have suggested a link between contagious yawning (CY) and sociality due to its appearance in mostly social species. Nevertheless, as findings are inconsistent, CY's function and evolution remains heavily debated. One way to understand the evolution of CY is by studying it in hominids. Although CY has been found in chimpanzees and bonobos, but is absent in gorillas, data on orangutans are missing despite them being the least social hominid. Orangutans are thus interesting for understanding CY's phylogeny. Here, we experimentally tested whether orangutans yawn contagiously in response to videos of conspecifics yawning. Furthermore, we investigated whether CY was affected by familiarity with the yawning individual (i.e., a familiar or unfamiliar conspecific and a 3D orangutan avatar). In 700 trials across 8 individuals, we found that orangutans are more likely to yawn in response to yawn videos compared to control videos of conspecifics, but not to yawn videos of the avatar. Interestingly, CY occurred regardless of whether a conspecific was familiar or unfamiliar. We conclude that CY was likely already present in the last common ancestor of humans and great apes, though more converging evidence is needed.

Based on:
Van Berlo, E., Díaz-Loyo, A. P., Juárez-Mora, O. E., Kret, M. E., & Massen, J. J. M. (2020). Experimental evidence for yawn contagion in orangutans (*Pongo pygmaeus*). *Scientific reports*, *10*, 22251. https://doi.org/10.1038/s41598-020-79160-x

Introduction

Yawning is an evolutionarily old phenomenon as its associated motor features can be recognized in different groups of animals (Baenninger, 1987). It follows a stereotyped pattern that, once started, is unstoppable (Provine, 1986). Apart from its spontaneous form, it is also highly contagious, at least for some species; i.e., individuals yawn as an unconscious and automatic response to seeing or hearing other individuals yawn (Massen et al., 2015). While a yawning-like pattern is observed in a wide range of vertebrates (Baenninger, 1987), contagious yawning (CY) is less wide-spread. To date, CY appears to be present in only a few, generally social species, including tonkean macaques (Palagi & Norscia, 2019) (and possibly stumptail macaques: Paukner & Anderson, 2006), gelada baboons (Palagi et al., 2009), chimpanzees (Amici et al., 2014; Anderson et al., 2004; Campbell et al., 2009; Campbell & Cox, 2019; Campbell & de Waal, 2011, 2014; Madsen et al., 2013; Massen et al., 2012), bonobos (Demuru & Palagi, 2012; Palagi et al., 2014), dogs and wolves (Joly-Mascheroni et al., 2008; Romero et al., 2013, 2014), sheep (Yonezawa et al., 2017), elephant seals (Wojczulanis-Jakubas et al., 2019), budgerigars (Gallup et al., 2015), and rats (Moyaho et al., 2015). In contrast, studies failed to show CY in grey-cheeked mangabeys and long-tailed macaques (Deputte, 1994), mandrills (Baenninger, 1987), common marmosets (Massen et al., 2016), lemurs (Reddy et al., 2016), horses (Malavasi, 2014), lions (Baenninger, 1987), tortoises (Wilkinson et al., 2011), and fish (Baenninger, 1987), even though some of these species are also very social. Despite growing interest in CY, both its proximate mechanisms (how it functions and develops) and ultimate causes (why and how it evolved) currently remain unclear.

Several hypotheses have been put forward, following a Tinbergian approach (Tinbergen, 1963). One view on the proximate mechanism underlying CY is that it is an automatic form of physiological or emotional state-matching between individuals. This synchrony of states between individuals may work via a perception–action mechanism (PAM), an adaptive mechanism that serves to create and maintain relationships in highly social species and that can give rise to higher-order cognitive phenomena such as empathy (Preston & De Waal, 2002). Some scholars argue that CY taps into the same PAM as emotion contagion (e.g., Anderson et al., 2004; Lehmann, 1979; Palagi et al., 2009; Platek et al., 2003), which is the tendency to automatically synchronize emotional states with another individual (Hatfield et al., 1993). Following this line of thought, CY can thus potentially be a proxy for empathy (i.e., the CY-empathy hypothesis) (Amici et al., 2014; Campbell & de Waal, 2011; Norscia et al., 2020; Norscia & Palagi, 2011a; Palagi et al., 2009; Platek et al., 2003; Romero et al., 2013). Indeed, neuroimaging studies have shown increased brain activity during CY in areas involved in theory of mind and social cognition (Cooper et al., 2012; Haker et al., 2013; Platek et al., 2005), corroborating the idea that CY is linked with emotional state-matching and perhaps even empathy.

Furthermore, individuals who score low on empathy scales (e.g., individuals on the autism spectrum) are less likely to show CY (Senju et al., 2007), and females yawn more frequently in response to seeing others yawn than males do, reflecting the idea that females show higher levels of empathy than males because of their investment in offspring care (Norscia et al., 2016a). Nevertheless, there are some studies that do not find such a clear link between CY and empathy. For instance, when people with autism spectrum disorder (ASD) are instructed to pay attention to the eyes (avoidance of the eyes is one of the characteristics of ASD), they are just as likely to yawn contagiously as neurotypical individuals (Usui et al., 2013). Furthermore, the gender bias is not consistently found (e.g., Bartholomew & Cirulli, 2014; Massen et al., 2012) and heavily debated (Gallup & Massen, 2016; Norscia et al., 2016a). For instance, in chimpanzees, it appears that males yawn more frequently than females in response to seeing other males yawn (Massen et al., 2012). Finally, while dogs do engage in CY, its presence is not affected by whether the yawner is prosocial versus the yawner being antisocial (Neilands et al., 2020). The mixed findings in the studies investigating the relationship between CY and a complex construct such as empathy show that the topic deserves more attention, and that it is still debated (see Massen & Gallup, 2017, for a critical review).

The emotional bias hypothesis is a more detailed specification of how CY can be socially modulated through a shared PAM, namely via social closeness and familiarity. The hypothesis predicts that individuals who are socially, and thus emotionally close are also more likely to yawn contagiously in response to each other (Demuru & Palagi, 2012; Hatfield et al., 1993; Palagi et al., 2004, 2014; Romero et al., 2013, 2014; Silva et al., 2012). Additionally, individuals from a group (i.e., familiar others) are more likely to yawn in response to each other than to unfamiliar others (Campbell & de Waal, 2011; Romero et al., 2013). A potential issue that has been raised is that these studies often fail to rule out simple alternative explanations for CY that do not require higher-order cognition (Massen & Gallup, 2017). For instance, effects of familiarity on CY may be explained by a general tendency to bias attention to familiar and socially close others (Massen & Gallup, 2017). Nevertheless, in a recent study investigating auditory yawn contagion in humans, yawns were most contagious between family and friends while controlling for the potential effects of increased attention to socially close others using a non-visual stimuli (Norscia & Palagi, 2011a). Still, in quite some social species, the linkage between CY and social closeness or familiarity is not found (Madsen & Persson, 2013; Massen et al., 2012; Miller et al., 2012b; Neilands et al., 2020; O'Hara & Reeve, 2011). For example, a recent study analyzing a large dataset on CY in dogs shows CY is present in dogs, but is not affected by familiarity or other potential mediators such as sex or prosociality (Neilands et al., 2020). It therefore remains possible that mechanisms other than the same PAM that underlies emotion contagion or empathy are mediating CY. For instance, CY may result from stress induced by a common stressor in the environment (Buttner & Strasser, 2014; Paukner & Anderson, 2006). Thus, rather than being mediated by seeing others yawn, yawning occurs as a response to the stressor. Individuals that are stressed are known to show higher rates of self-directed behaviors, of which yawning and self-scratching are examples (Troisi, 2002), and indeed, in one study involving stumptail macaques, monkeys yawned more frequently in response to a video clip of yawns as compared to a control, but also scratched more (Paukner & Anderson, 2006). The authors concluded that tension was most likely mediating the occurrence of yawning in the yawn condition. In short, while it is likely that CY is a social phenomenon, its exact mechanisms remain an active field of investigation.

Notwithstanding the debate on proximate mechanisms, little attention has been given to more ultimate explanations for CY. One of the few hypotheses out there is that CY is an adaptive mechanism that helps with social coordination (Miller et al., 2012a). Accumulating evidence suggests that yawning itself serves to cool the brain as to maintain homeostasis (Eguibar et al., 2017; Gallup & Eldakar, 2012; Gallup & Gallup, 2007, 2008; Gallup & Gallup, 2010; Massen et al., 2014) and consequently may increase alertness and aid in vigilance. Within this social coordination hypothesis, CY, in turn, may help to spread vigilance within the group, for instance to remain alert for potential predators (Gallup & Gallup, 2007; Miller, Gallup, Vogel, & Clark, 2012). Specifically, it may be adaptive to match the state of a vigilant conspecific as it may have sensed a predator, which the individual itself did not yet sense. To date, however, the social coordination hypothesis remains untested, and the thermoregulatory function of yawning is still debated (e.g., Elo, 2011; Guggisberg et al., 2011, but see Gallup & Eldakar, 2012, for a response to the critique).

Another fruitful way to explore evolutionary hypotheses is through phylogenetic comparisons. Palagi et al. (2019) proposed the *common trait among hominids* hypothesis which states that, given the shared phylogeny between humans and great apes, CY may find its roots in a shared underlying socio-cognitive mechanism that was already present in at least the last common ancestor (LCA) of all hominids. Moreover, since CY is also present in some Old-World monkeys (Palagi et al., 2009; Paukner & Anderson, 2006) and non-primate species (e.g., Gallup et al., 2015; Madsen & Persson, 2013; Norscia et al., 2021; Silva et al., 2012), its roots could be much older, or CY is an example of convergent evolution. To date, few data exist to perform comparisons and most interestingly, the picture among the great apes is not yet clear. There is convincing evidence for CY in chimpanzees (Amici et al., 2014; Anderson et al., 2004; Campbell et al., 2009; Campbell & Cox, 2019; Massen et al., 2012). In bonobos, two observational studies (Demuru & Palagi, 2012; Palagi et al., 2014) and an experiment (Tan et al., 2017) show clear evidence for CY, while one experimental study did not (Amici et al., 2014). However, the latter study only tested four individuals, thus making it very likely that CY is, indeed, present in bonobos. Finally, the first comprehensive study on gorillas combining an experimental and naturalistic approach found no evidence for CY (Palagi et al., 2019a). Unfortunately, data on CY in orangutans are scarce, which, considering their semi-solitary lifestyle (Van Schaik, 1999) may be of comparative interest for a social phenomenon like CY. To date, the only existing study involving orangutans failed to find evidence for CY (Amici et al., 2014), yet the sample size was too small to be conclusive. In general, orangutans in the wild roam mostly solitarily: males travel alone, and mothers travel with their offspring (Singleton & Van Schaik, 2002; Te Broekhorst et al., 1990). Due to overlapping home ranges, occasional encounters and affiliation are possible, but generally do not occur frequently (Singleton & Van Schaik, 2002; Te Broekhorst et al., 1990). Consequently, finding out whether CY is present in orangutans will further help elucidate the hypotheses previously discussed.

The current study attempts to clear up the picture of CY in hominids in two ways. First, we aim to find a convincing answer to whether CY is present in orangutans or not via an experimental design involving the presentation of yawning and neutral stimuli of orangutans to 8 orangutans. Second, we also investigate whether this potential yawn contagion is affected by a familiarity bias, i.e., whether CY is stronger between individuals that know each other versus unfamiliar individuals. To this end, we exposed orangutans to videos showing either yawn or control clips of familiar (i.e., conspecifics living in close proximity) and unfamiliar orangutans, as well as a 3D avatar (Kolbrink, 2017) and measured their response (yawns). Additionally, we also measured the occurrence of self-scratching to rule out potential effects of stress on the occurrence of yawning (Troisi, 2002). So far, CY appears to be exclusively present in highly social species, and because orangutans do not show high affiliative tendencies, we therefore expected that orangutans do not show CY.

Method

Participants

Eight orangutans (ages: 15 months-36 years, 4 males) housed at Apenheul (The Netherlands), were tested (see Table S1 for more information about the participants). Individuals were divided over four neighboring enclosures and group composition varied weekly (Figure 1). The two adult females that had dependent offspring were always housed with their offspring and sometimes with one adult male. Experiments took place in the visitor area but while the park was closed to visitors. We tested individuals using a movable 47″ TV (LG 47LH5000, 1920 * 1080 pixels) placed in front of the enclosures to which the orangutans were habituated before commencing

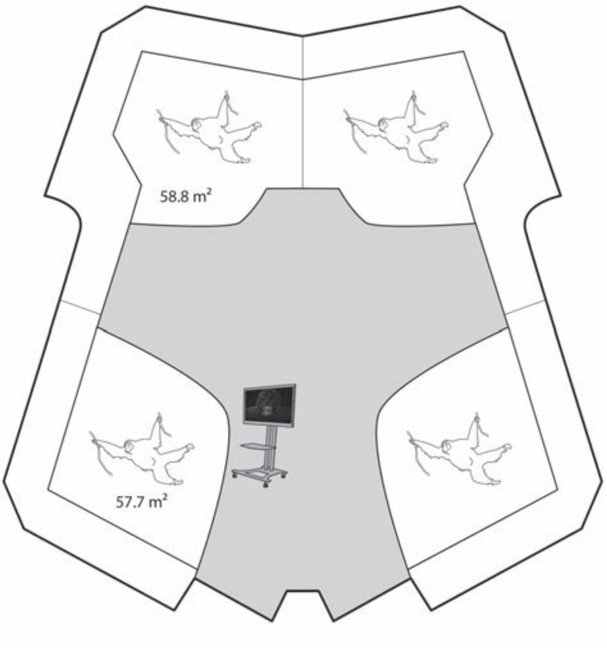

Figure 1. Abstract representation of the testing area and group composition. Experiments took place in the visitor area but while the park was closed to visitors. Only the experimenters and, occasionally, a caretaker, were present during testing. The TV screen was always directed at one of the four enclosures, which prevented orangutans in the other enclosures from seeing the videos.

testing. The screen was always directed at one of the four enclosures, which prevented orangutans in the other enclosures from seeing the videos. Food was provided four to six times a day and consisted of a variety of vegetables, and sometimes nuts, hay, and fruit, hidden in the enclosure for foraging purposes. Water was available ad libitum.

The care and housing of the orangutans was adherent to the guidelines of the EAZA Ex situ Program (EEP). As the study was non-invasive in nature, there was no need for the approval of the Ethics Committee of Apenheul primate park and the study complied with the requirements of the Dutch Animal Care and Use Committee.

Stimuli

The experiment involved three categories of mute, full-screen videos, each consisting of both a yawn and control condition (see Figure 2 for examples). We used mute videos as the enclosures were sealed with thick glass that dampened most of the sound both ways. Yawn videos showed clear yawns either filmed from the front or side, whereas control videos consisted of individuals with a neutral face and in a relaxed body position. Both types of videos involved movement, with yawn videos showing a wide gaping of the mouth followed by a relaxation of the mouth and jaw (Barbizet, 1958), including display of the teeth, and control videos showing an individual with a closed mouth with random movements of the lips. Both control and yawn videos were always of the same individual, and therefore the body position and face were identical.

The *familiar* video category consisted of two adult males housed in the zoo. For the *unfamiliar* video category, we used two adult males taken from clips on YouTube. Finally, in the *avatar* video category we used two mirrored videos of a computer-generated adult male. The 3D orangutan was created by Paul Kolbrink from XYZ-Animation and designed in Autodesk 3ds Max (2017) using the Octane render engine. Using these videos, we created video sequences starting with a primer video that depicted caretakers beckoning the orangutans towards the TV screen, which were created to grab the orangutans' attention right before the start of a trial. As we repeated the presentation of our video database four times during the course of the experiment, there were four different primers; one for every repetition.

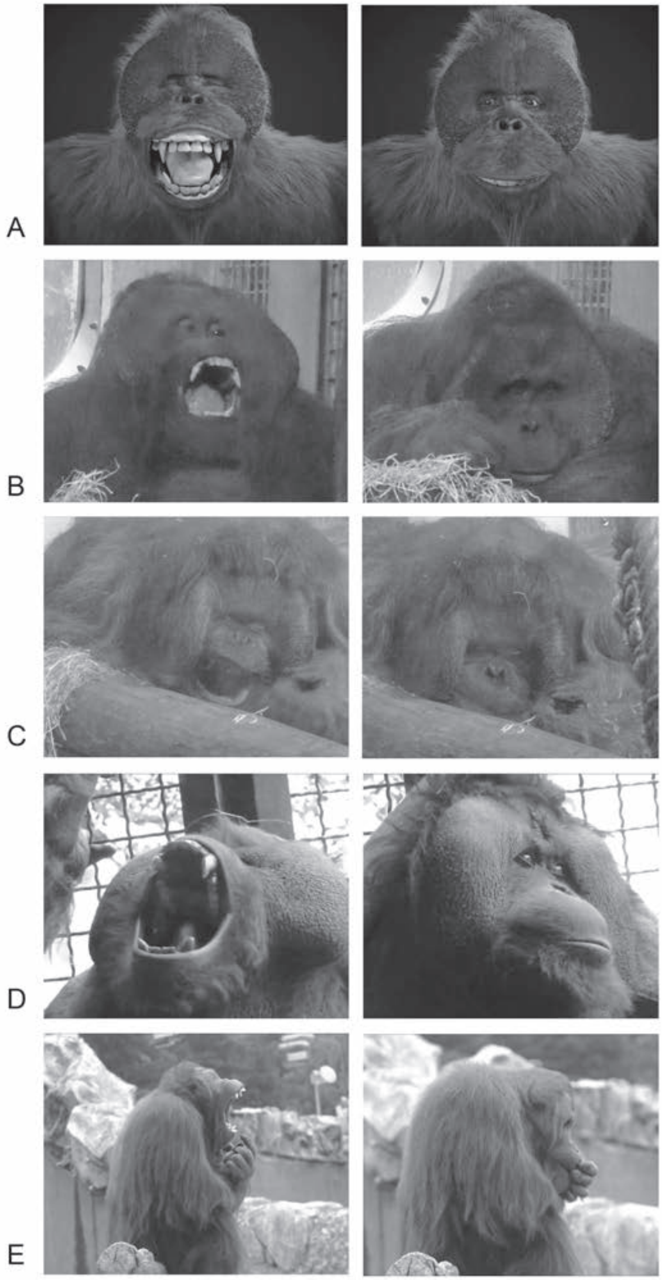

Figure 2. Stills of the videos used in the experiment with yawns on the left and controls on the right. A: Avatar*, B: Familiar adult 1, C: Familiar adult 2, D: Unfamiliar adult 1, E: Unfamiliar adult 2.
**To decrease the chances of pseudo-replication within this category we created horizontally mirrored copies of the yawn and control videos of the avatar such that – similar to the other triggers – we had two yawn and two control videos in total.*

Procedure

The experiment was carried out between 21-01-2019 and 13-03-2019. In this period, the park was closed for visitors. A test session involved the presentation of two different trials, each consisting of a specific video sequence, and each trial followed by an observation period. The video sequence consisted of a primer, followed by either a yawn or control video (lasting 14 s), which was repeated 4 times and with a colored screen (again to grab attention) for 1 s in between each video. The length of one video sequence was thus 90 s (cf. Massen et al., 2013): primer (30 s) – colored screen (1 s) – yawn/control video (14 s) – colored screen (1 s) – yawn/control video (14 s) – colored screen (1 s) – yawn/control video (14 s) – colored screen (1 s) – yawn/control video (14 s). The presentation of one video sequence (representing one trial) was then followed by a 3.5-min observation period, after which the second trial started. If the first trial involved yawn videos, the second trial involved control videos and vice versa. The second trial was also followed by a 3.5-min observation period, completing one test session. Within one test session we always showed the same stimulus individual. See supplements for more information on the procedure.

We cycled through the entire video database four times (i.e., 4 blocks) over the course of the experiment to ensure sufficient data points. The order of control and yawn trials were counterbalanced per subject, and was further counterbalanced over the subjects per block. Within each block, *trigger* (i.e., familiar/unfamiliar/avatar) was also randomized per subject. We designed a testing schedule based on eight test subjects, but two of those subjects involved a mother-infant pair and a mother-juvenile pair in which the infant/juvenile never left the mother. As such, we created a test schedule for six individuals rather than eight. With these six test subjects, three types of triggers, two conditions (yawn and control), two orders of condition presentation (yawn-control, or control-yawn), and finally four repetitions, we had a total of 288 test sessions and 576 trials planned (see Tables S2.1 to S2.3 and S3 for an overview). However, one video sequence was accidentally presented an extra time, resulting in 289 rather than the planned 288 sessions after data collection finished. On any given testing day, individuals participated in one or two sessions with 30 min breaks between video presentations to the tested subject. Furthermore, subjects never saw a video sequence more than once on any given day.

APDL and OEJM recorded all occurrences of yawning and self-scratching, and self-scratching was recorded as a measure for arousal and tension (Troisi, 2002). It was not possible to reliably quantify the amount of time spent looking at the screen due to the lack of continuous visibility of the gaze of the orangutans. To nonetheless ensure maximum attention to the screen, we presented primers before video sequences and colored videos in-between yawn and control clips, and we only started testing when orangutans had a direct line of sight towards the screen. Additionally, before each trial, we observed the orangutans for five minutes, and only started a trial if there were no yawns before the presentation so as to rule out that yawns within a trial were potentially caused by a previous yawn outside of the trial. Furthermore, yawns were scored in response to either the yawn or control video only if a subject looked at least once to the screen during presentation. If bystanders in the same enclosure attended to the screen, their behaviors were also scored. Data collection ended after 10 min, concluding one test session. Finally, EvB coded 15% of the videotapes for inter-rater reliability purposes. Results showed a good agreement on occurrences of yawning (ICC = 0.764, $p < 0.001$) and self-scratching (ICC = 0.894, $p < 0.001$). In subsequent analyses, only yawns on which the raters agreed were used.

Statistical analyses

The dependent variable was whether a subject yawned in response to a video or not. Because it is difficult to disentangle between whether multiple yawns occurring in succession are caused by another individual, or whether they are simply the result of an urge to yawn multiple times perhaps because of self-contagion (i.e., where your own yawns cause you to yawn again), we did not compare rates of yawning to establish CY (Kapitány & Nielsen, 2017). Rather, we looked at the likelihood of yawning within the yawn and control condition to establish the presence or absence of CY in orangutans. Nevertheless, when contagion indeed occurred, yawning rate could inform about the *strength* of contagion (Kapitány & Nielsen, 2017). As such, we analyzed our data using hurdle models in R (lme4 package, Bates et al., 2015). Hurdle models follow a two-step method that first deals with zero-inflated count data and subsequently with positive counts once the initial hurdle is crossed (Cameron & Trivedi, 2013), which make them applicable to our dataset.

In the first hurdle model we focused on whether CY is present or absent in orangutans by comparing the likelihood of yawning in the yawn and control condition using a binomial GLMM, in which we added *condition* as a fixed effect and *subject* nested in *trial* as a random effect. In the second step of the model, we analyzed the rates of yawning using a negative binomial GLMM only in those cases where at least one yawn occurred. Again, we entered *condition* as a fixed effect and *subject* nested in *trial* as a random effect. In the second hurdle model, we tested for potential effects of both *condition* and *trigger* (i.e., familiar/unfamiliar/avatar) and their interaction on the likelihood of yawning using a binomial GLMM, entering *condition* and *trigger* and their interaction as fixed effects, and again *subject* nested in *trial* as random effect. In the second step of the model, we were interested in how the conditions and triggers affected yawning rates in those cases that at least one yawn occurred. To investigate this, we entered *condition* and *trigger* and their interaction as fixed effects and *subject* nested in *trial* as random effect using a negative binomial GLMM.

It is possible that the likelihood of yawning in the conditions is due to the stimuli somehow being arousing to the observers, complicating the interpretation of the underpinnings of CY (see e.g., (Paukner & Anderson, 2006). For instance, yawning often involves display of the canines, which may be arousing for the orangutans (Plavcan, 2001). Therefore, as a control analysis, we looked at self-self-scratching behavior as this is indicative of arousal in primates (Troisi, 2002). In a third hurdle model, we checked whether the likelihood of self-scratching is affected by *condition* (fixed factor), with *subject* nested in *trial* as random factor and using a binomial GLMM. In the second step of the model using a negative binomial GLMM with *subject* nested in *trial* as random factor, we investigated whether self-scratching rate was affected by *condition, trigger*, and their interaction as fixed factors only in those cases when self-scratching occurred.

In all analyses, we compared the models to their respective null-models (i.e., including only the random effects) and only report on significant values if the models and null-models differ significantly from each other (Forstmeier & Schielzeth, 2011). For post-hoc contrasts of interaction effects we report corrected p-values using Tukey-adjustments. Alpha was set to 0.05.

Results

In total, we witnessed 83 yawns across 8 individuals and 289 sessions (see Tables S2.1 to S2.3). First, we investigated the likelihood of yawning in the two conditions. We found a significant effect of *condition*; yawning was more likely to occur in the yawn versus the control condition ($\beta = 3.45$, $SE = 1.06$, $p = 0.001$). Next, we compared the yawning rate between the two conditions in those cases that at least one yawn occurred, but this alternative model did not deviate significantly ($\chi^2(1) = 3.09$, $p = 0.079$) from its respective null-model.

Assessing whether familiarity affects the occurrence of CY, we found a significant interaction effect of *trigger* (familiar, unfamiliar, avatar) with *condition*.

Specifically, we found a significant contrast of yawns between the yawn and control condition in the familiar ($\beta = 6.62$, $SE = 1.59$, $p < 0.001$) and unfamiliar trigger ($\beta = 3.45$, $SE = 1.52$, $p = 0.023$), but not in the avatar trigger ($\beta = 0.09$, $SE = 1.58$, $p = 0.950$) (Figure 3). Hence, orangutans are more likely to yawn in response to yawning videos rather than to control videos, but only when the yawning individual is a "real" orangutan (i.e., a familiar or unfamiliar conspecific), and are less likely to yawn in response to the avatar. To investigate whether the likelihood of CY differed with regard to familiarity with the "real" orangutan stimuli, we also ran an additional binomial model on a reduced dataset that excluded all trials with the avatar (see supplemental materials). Whereas this model confirmed the previously found effect of *condition*, here we did not find a significant interaction between *condition* and *familiarity*, suggesting that the likelihood of CY was not being modulated by the familiarity with the "real" orangutan. We also investigated the effect of *familiarity* on yawning rate using the same reduced dataset, but the model including the interaction between *condition* and *trigger* did not significantly improve the null model ($\chi^2(3) = 3.50$, $p = 0.321$). As such, while we can establish that orangutans do show CY in response to yawn videos of familiar and unfamiliar conspecifics, this likelihood of CY is not modulated by familiarity and we cannot draw any conclusions regarding the *strength* of CY in relation to familiarity.

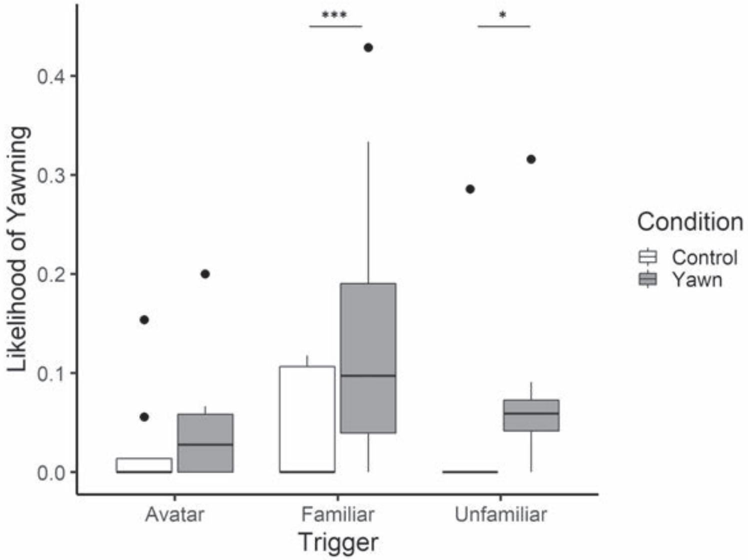

Figure 3. Likelihood of yawning across conditions and triggers. Boxplots show the median (solid line), 25th-75th percentile (box) and the largest and smallest value within 1.5 times the interquartile ranges respectively (whiskers). Dots reflect outliers.

Looking at self-scratching, we first investigated the likelihood of self-scratching when viewing yawning and control videos, and found that the *occurrence* of self-scratching did not differ between conditions ($\beta=-0.17$, $SE=0.17$, $p=0.319$). Similarly, self-scratching *rates* were not significantly higher in the yawn versus control condition ($\beta=0.10$, $SE=0.09$, $p=0.301$). Moreover, both models did not deviate significantly from their null-model ($\chi^2(1)=1.06$, $p=0.303$). Hence, it is unlikely that orangutans were more aroused viewing yawn videos compared to viewing control videos, at least when measured via self-scratching. Additionally, we also included self-scratching in our original models on yawning as a covariate, and found it to not significantly explain the likelihood of yawning, nor to influence the found effects of condition and the lack thereof in the avatar treatment (see supplements for these analyses).

Discussion

Here we find that orangutans yawn contagiously in response to conspecifics yawning, independent of whether the conspecific is a familiar or unfamiliar individual. Furthermore, orangutans were not susceptible to yawns of an avatar. Additionally, the videos used in our experiment appeared to be similarly arousing. That is, there was no difference in self-scratching (an indicator of stress) between the conditions. We here discuss the consequences of our findings for the different proximate and ultimate hypotheses that currently exist.

CY has thus far been observed in highly social species (Anderson et al., 2004; Demuru & Palagi, 2012; Joly-Mascheroni et al., 2008; Palagi et al., 2009; Romero et al., 2014; Wojczulanis-Jakubas et al., 2019a; Yonezawa et al., 2017, but see: Baenninger, 1987; Deputte, 1978; Malavasi, 2014; Massen et al., 2016; Reddy et al., 2016). Orangutans have meaningful social interactions that occur more often than is expected by chance alone (Roth et al., 2020), but these interactions occur at a much lower frequency compared to bonobos and chimpanzees (Te Broekhorst et al., 1990; Van Schaik, 1999). Interestingly, our results show that orangutans exhibit CY, suggesting that a high degree of affiliation within a species is not necessary for CY to occur. This also indicates that more studies are needed that investigate the presence or, importantly, absence of CY in a variety of species that differ on their social organization and affiliative tendencies. At the same time, it has to be noted that our sample consists of zoo-housed orangutans that were also born in captivity. In captivity, frequencies of affiliation can exceed those observed in the wild (Edwards & Snowdon, 1980), thus potentially increasing the likelihood of CY to occur. Nevertheless, our results do show the presence of CY in orangutans and the few generations of zoo-living individuals cannot inform us about any selection pressures that have resulted in this tendency in orangutans. Our results must therefore be discussed in light of the orangutans" natural behavior and social environment.

In our study, we did not find an effect of familiarity on CY, suggesting that at least in orangutans, social modulation of CY may not be present. While presence of social modulation of CY is often used as confirmation of CY and emotion contagion sharing the same underlying perception–action mechanism (Campbell & de Waal, 2011; Demuru & Palagi, 2012; Norscia & Palagi, 2011a; Palagi et al., 2014; Romero et al., 2013), its absence in our data makes it more difficult to interpret the emotional bias hypothesis. Orangutans do have some preferences when it comes to their interaction partners, thus one could expect social modulation of CY under the emotional bias hypothesis. For instance, related female orangutans are known to associate more often than unrelated females (Van Noordwijk et al., 2012), and prefer the long-calls of dominant males (Setia & van Schaik, 2007). Additionally, in a recent study, orangutans were shown to scratch contagiously in response to conspecifics self-scratching, suggesting a potential case of emotion contagion (Laméris et al., 2020). Interestingly, scratch contagion was stronger between weakly bonded individuals during tense situations, which shows a social closeness bias in the opposite direction. This suggests that a familiarity bias may be more flexible depending on the situation individuals are in (e.g., relaxed versus stressful contexts) and the nature of the behavior that is copied (e.g., self-scratching as an expression of tension).

At the same time, there are other studies on highly social species that do not show a familiarity bias (e.g., chimpanzees: Massen et al., 2012, dogs: Neilands et al., 2020, macaques: Deputte, 1978, and marmosets: Massen et al., 2016). As such, there may be (currently unknown) species-specific traits that determine whether a familiarity bias occurs or not. The exact (social) function of CY remains unclear and thus alternative explanations that do not involve the PAM that is underlying empathy may still be possible (e.g., spreading of vigilance). As has been pointed out by others, solving this issue requires a more systematic study of CY that includes a bigger variety of animals, including solitary animals such as reptiles and amphibians (Massen & Gallup, 2017).

From an evolutionary perspective, our results pose an interesting conundrum: while we found CY in orangutans, it is not present in gorillas, even though the split between orangutans and other hominids is evolutionarily older than the split between gorillas and other hominids (Schwartz, 1987). It is possible that the number of trials in the study by Palagi et al. (2019) were not sufficient to detect CY, as in our study, even with a large number of trials, we only detected yawns in 11.9% of all cases. Nevertheless, studies with chimpanzees that have few trials were able to establish CY in the past, albeit with a relatively large number of subjects (Amici et al., 2014; Campbell et al., 2009; Massen et al., 2012), and there was also no evidence for CY in naturalistic observations in gorillas (Palagi et al., 2019a).

Interestingly, it has been argued that in the past, orangutans may have been more social, but that due to long periods of low food availability, orangutan gregariousness was no longer viable (Harrison & Chivers, 2007). This may suggest that the ancestor of all apes already possessed the mechanism underlying CY. However, based on observational and relatedness data, it has been suggested that this hominid lived in a group with gorilla-like structure in which one male could monopolize multiple females (Harrison & Chivers, 2007). In this sense, it is difficult to explain why, given a similar social structure, gorillas do not show CY and orangutans do. It is possible that CY was somehow lost in the gorilla lineage, or that CY evolved multiple times over the course of evolution. The loss of CY is theoretically possible, given that CY has been found in some, but not all primates (Baenninger, 1987; Palagi et al., 2019a). Here, there is a role for the type of social system that characterizes a species in the loss (or occurrence) of CY (Palagi et al., 2019a). There is, however, not yet enough variation in data on CY in different species of primates to draw clear conclusions. Furthermore, it is possible that the measures to detect CY in certain species are simply not sensitive enough. All these explanations can be true, given that the occurrence of CY is highly variable in primates in general. It is clear that more studies are needed in order to draw robust conclusions about the evolution of CY.

In our study orangutans did not significantly respond to the avatar, which contrasts with findings in chimpanzees (Campbell et al., 2009). Potentially, orangutans experienced the uncanny valley phenomenon in which the avatar looks very realistic, yet fails to behave like a real orangutan, therefore violating natural expectations of orangutan behavior. Indeed, previous research on monkeys showed that they preferentially looked at real or completely unrealistic 3D model monkeys compared to very realistic 3D models (Steckenfinger & Ghazanfar, 2009). Nevertheless, this would likely have increased self-scratching when viewing the avatar, which was not evident in our study. Furthermore, a recent study investigating the uncanny effect in macaques showed that looking times did not differ between the Primatar (3D monkey head) and real or unrealistic images, indicating that the use of virtual stimuli can still be a promising way to study social cognition (Wilson et al., 2020). Future studies will have to verify whether the lack of evidence for CY using an avatar in our study is because the effect is truly absent, for instance by looking specifically at how similarity with another individual (on a physical level) affects CY. In humans, there is ample evidence that the more similar that individuals are in terms of physical characteristics, but also personal convictions and views, the more likely they are to automatically mimic behavior (Chartrand & Bargh, 1999).

Future studies can improve on the current study design in several ways. First, we only used orangutan males as stimuli. In previous studies with chimpanzees (Massen et al., 2012) and bonobos (Demuru & Palagi, 2012), the sex of the triggering yawner affected the occurrence of CY; i.e., in chimpanzees, male yawns were more contagious whereas in bonobos, female yawns were more contagious. In gelada baboons, CY is more prevalent among females, especially when they are closely bonded (Palagi et al., 2009). It is possible that these results can be explained by emotional closeness between individuals, as in chimpanzees males typically form strong social relationships (Mitani et al., 2000), and in bonobos and gelada baboons it is mostly females that bond (Dunbar & Dunbar, 1975; Furuichi, 1989).

Alternatively, results could be explained by the differences in hierarchy with chimpanzees being male dominant (Goldberg & Wrangham, 1997) and bonobos female dominant (Furuichi, 1989), and by the strong matrilineal bonds between gelada baboons (Silk et al., 2004). Investigating whether there is an interaction between sex of the stimulus and of the responder in orangutans could help elucidate the roots of the observed sex effects in CY in some species. The restricted selection of stimuli and the low sample size did unfortunately not allow us to perform such analyses. It is noteworthy, however, that the males in our study yawned more frequently than the females (i.e., the total yawning rate of males was 74, whereas females yawned only 9 times. See Table S2A). Yawns occur more frequently in males of species with canine polymorphism, and also during aggressive contexts (Leone et al., 2014). Given that all our stimuli were male, perhaps there is a role for dominance or rivalry in the occurrence of CY in orangutans (Moyaho et al., 2015). Nevertheless, one could argue that this leads to tense situations, thus leading to more self-scratching when observing yawns of others, which is not what we found.

Additionally, all of our videos contained flanged males. Flanged adult males are often preferred over unflanged males by receptive female orangutans (Knott et al., 2010), and can be viewed as threatening by unflanged males (Delgado & Van Schaik, 2000). As such, in addition to interactions between the different sexes and CY, it may also be interesting to study potential effects of the two different morphs of orangutan males on CY.

Furthermore, due to power issues, we could not reliably test effects of age on CY. In humans, while spontaneous yawns can occur already before birth (De Vries et al., 1982), CY does not seem to appear until the age of four to five (Anderson & Meno, 2003; Millen & Anderson, 2011), although when children of 3 years old are specifically told to look at the eyes of the stimulus they show CY as well (Hoogenhout, 2013). Similar developmental trajectories of CY have been reported in other animals (Anderson et al., 2004; Madsen et al., 2013; Madsen & Persson, 2013; Palagi et al., 2009). In our study, there were only two individuals younger than 5; one 15 months (Indah) and one three-year old (Baju). We observed one yawn occurrence in Indah (in the yawn condition), in Baju we observed six events (four in the yawn and two in the control condition). We decided to include these individuals in our study because while it is true that CY shows a relatively slow developmental pattern in humans, orangutans are born more precocial, and developmental rates in nonhuman primates are much faster compared to humans (Clancy et al., 2007). Therefore, CY may possibly also occur earlier in development in orangutans, but with only anecdotal evidence we cannot verify this in our study.

Third, while we tested effects of familiarity in our study by including both familiar and unfamiliar yawners, the fact that we only had yawns from the two adult males to use as stimuli restricted any potential investigation of the potential link between social closeness of the responders and the familiar individuals on the stimuli. The positive effect of social closeness on the occurrence of CY is well established in humans (Provine & Hamernik, 1986), chimpanzees (Campbell & de Waal, 2014, but see Massen et al., 2012), and bonobos (Demuru & Palagi, 2012), but is strongly debated in other species such as dogs (Neilands et al., 2020) and budgerigars (Miller et al., 2012b). For dogs, it should be noted that CY is interspecific, and that domestication might have had influential effects on how CY is modulated. Inverse effects have also been reported. For instance, a large study in rats has shown a familiarity bias in the opposite direction with rats being more likely to yawn in response to unfamiliar yawns (Moyaho et al., 2015). Similarly, a recent study investigating scratch contagion in orangutans found that during tense situations, orangutans are more likely to take over self-scratching from individuals with whom they have a weak bond (Laméris et al., 2020), indicating a (negative) correlation between social closeness and the contagiousness of a behavior or motor pattern. Thus, it remains possible that social modulation of CY is present in orangutans, at least in those living with conspecifics in captivity, although its presence was not shown in our sample. Yet, given our small sample size, replications that test for the presence and subsequent direction of social modulation of CY in orangutans are needed.

Finally, we could not quantify attention to the screen, which is one of the common methodological issues raised by Massen et al. (2017). We tried to maximize attention to the screen by using attention-grabbing videos of caretakers at the start of every video sequence, and by adding colored screens in-between stimulus presentations. Furthermore, we made sure that orangutans had a direct line of sight towards the screen at the start of the experiment, and only recorded yawns when they directed their attention to the screen at least once during stimulus presentation. Nevertheless, quantification of attention to the stimuli (either measured as a continuous variable or a frequency of gazes) remains the most robust way to control for potential effects of attentional bias.

To summarize, our findings contribute to understanding the evolutionary basis of CY in hominids by showing that orangutans, like humans, chimpanzees and bonobos, yawn contagiously.

Part III:

Implicit Associations

Chapter 7

Validation of a pictorial version of the Implicit Association Test (IAT) for comparative research

Abstract

The Implicit Association Test (IAT) is frequently used to measure implicit associations, but one issue with current forms of the IAT is that they require at least some comprehension of written and/or spoken language, making them difficult to use in illiterate or pre-verbal populations. We therefore designed a self-explanatory, touchscreen-based, fully pictorial IAT that we validated across two experiments. In Experiment 1 we tested the PIAT's ability to measure implicit inter-ethnic attitudes in 129 Dutch adults and in 143 Dutch children visiting a zoo. In Experiment 2, we validated the PIAT by directly comparing its results to a word IAT in an online, within-subjects experiment involving 141 adults. D-score analyses showed that the PIAT can reliably tap into the same implicit biases as its verbal counterpart. We believe that the PIAT provides a good adaptation to the original IAT, offering a standardized test that could potentially be suitable for quantitative, cross-cultural, and cross-species comparisons.

Based on:
Van Berlo, E., Otten, M., Roth, T. S., Binnekamp, J., Van der Ven, E. J., & Kret, M. E. (2021). Validation of a pictorial version of the IAT. *Manuscript submitted for publication*.

Introduction

Since the publication of the seminal paper by Greenwald, McGhee, and Schwartz (1998), the Implicit Association Test (IAT) has been one of the most well-established tasks to measure implicit attitudes (Cunningham et al., 2001; Greenwald et al., 2003, 2009; Lane et al., 2007). Implicit tasks such as the IAT are crucial for providing a window into unconscious processes that drive behavior. Unfortunately, most versions of the IAT require comprehension of written or spoken language, thereby limiting its usability in non-verbal or illiterate populations such as young children, clinical populations such as individuals on the autism spectrum or cognitively impaired individuals, and possibly in comparative research with non-human animals as well. To offer researchers who work with these populations an adaptation to verbal IATs, we developed a fully pictorial, intuitive implicit association test, and validated it against its classical counterpart in a population of children and adults.

The original IAT has received extensive psychometric evaluation and is a widely used tool to assess implicit attitudes and stereotypes (Cunningham et al., 2001; Greenwald et al., 2003, 2009; Lane et al., 2007). It measures the strength of implicit associations between two concepts (e.g., names of African-Americans vs. White Americans) and two attribute dimensions (e.g., pleasant vs. unpleasant words) by comparing reaction times in a categorization task consisting of a series of testing blocks. In the practice blocks, participants learn to categorize exemplars of the concepts and attributes into their superordinate categories. For example, they categorize a name such as “Tyrone” as African-American and “Hannah” as White-American, and a word such as “happiness” as pleasant and “suffering” as unpleasant. In the critical blocks, these superordinate categories are combined. For instance, in the first critical block participants categorize names and faces into the combined superordinate categories “White” + “unpleasant” and “Black” + “pleasant”. In the subsequent critical block(s), this combination is reversed (e.g., “White” + “unpleasant”). In general, participants respond faster in critical blocks congruent with their implicit associations, and slower in incongruent critical blocks (Greenwald et al., 1998). Importantly, in order to complete a typical IAT, a good understanding of written and/or spoken language is necessary, as the task activates implicit attitudes through the use of words and names representing (a subset of) the superordinate categories.

Several studies have made elegant adaptations to the IAT to partly overcome the necessity for understanding written or spoken language, especially within the developmental sciences. For instance, one IAT used pictures of flowers and insects to represent the concept categories, but retained words for describing the attribute dimensions (e.g., “good” and “bad”) in order to test implicit attitudes in 6- and 10-year-olds (Baron & Banaji, 2006, but also see a pictorial and touchscreen adaptation by Thomas et al., 2007). In the pre-school IAT, Cvencek and colleagues (2011) used schematic representations of smiling and sad faces to indicate the attribute dimensions “good” and “bad”, and pictures of flowers and insects for the concepts. Furthermore, during the task pictures and (spoken) words were alternated (Cvencek et al., 2011). In another child-friendly IAT, only pictures of White and Black faces were used, together with line drawings of happy and sad faces (Rutland et al., 2005). Instead of pushing buttons that were mapped with the left and right superordinate categories, children had to move their mouse towards the target locations. Using a touchscreen, another adaptation of the IAT used pictures of faces that were shown in the middle of the screen with a happy and sad emoticon on the left and right side on the bottom of the screen. In critical blocks, children were told whether to press the “happy” or “sad” emoticon (Qian et al., 2016; Setoh et al., 2019). Notwithstanding the high usefulness of these tasks, a commonality in the adaptations is that child participants still received (extensive) instructions (for instance, extra verbal instruction during critical blocks (Qian et al., 2016; Setoh et al., 2019)), or the tasks combined words with pictures (Baron & Banaji, 2006; Cvencek et al., 2011). To this end, we focused on creating an entirely non-verbal pictorial version of the IAT (from hereon: PIAT) with stimuli that are proven to be interpreted similarly across cultures, and that requires a minimal amount of instructions to complete.

The aim of our study was to assess our PIAT’s validity in comparison to its classic counterpart. In Experiment 1, we test the PIAT in a large, diverse population consisting of Dutch adults and children that are visiting a zoo with two aims in mind, i) to assess whether our version of the PIAT can indeed measure implicit attitudes, and ii) to assess whether it is suitable for use in a heterogeneous sample in a more naturalistic environment (i.e., not in the lab). In Experiment 2, we directly compare the performance of an online version of our PIAT to an online word IAT (WIAT) with the aim to further validate it as a tool for measuring implicit attitudes. We have chosen to measure implicit attitudes towards different ethnicities, as these have also been extensively studied in traditional IATs (Baron & Banaji, 2006; Greenwald et al., 1998). In both experiments, we assess participants’ implicit attitudes towards individuals of Moroccan and Dutch descent using images from validated face and emotional images databases (Lang et al., 2007; Langner et al., 2010). We chose for these ethnicities because negative opinions about individuals of Moroccan descent are pervasive in Dutch society (Andriessen et al., 2020). For instance, compared to other minorities, individuals of Moroccan descent were rated as most negative in a Dutch national survey (Vrooman et al., 2014). As such, we expect to find that Dutch participants (i.e., participants with 2 parents born in the Netherlands) have negative implicit associations with individuals of Moroccan descent, compared with individuals of Dutch descent. Furthermore, we expect to find this bias in both adults and children. Finally, we expect that the results of the PIAT are comparable to those found in word-based IATs.

Experiment 1: PIAT in the zoo

Method

Participants

129 Adults (73 females) and 143 children (72 females) took part in this study. Participants were visitors of a zoo in the Netherlands, and were all Dutch-speaking and of Dutch descent. The majority of individuals were right-handed (adults: 113 participants (87.6%); children: 121 (84.6%)). Children were between ages 5 to 17 (*M* = 10, *SD* = 2.29) and adults between ages 18-76 (*M* = 33.83, *SD* = 13.95). Consent for participating in the study was received from all adult participants and parents of participating children. Participants took part in the study on a voluntary basis and thus were not compensated for their participation. Data were collected between April and May 2017 (see Figure S1 in supplements for photos of the setup).

Task

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) measures the strength of an association between a pair of concepts based on the reaction times of the participant. In a series of blocks, participants categorize concepts into two opposing categories. In the original version, seven blocks of trials are used, but versions with five blocks are common as well, especially when the target group consists mainly of children (Nosek et al., 2005). We thus opted for the five-block version of the task. Block one, two, and four were practice blocks in which participants learned to categorize images; block three and five were critical blocks in which the speed of categorization was measured (Figure 1).

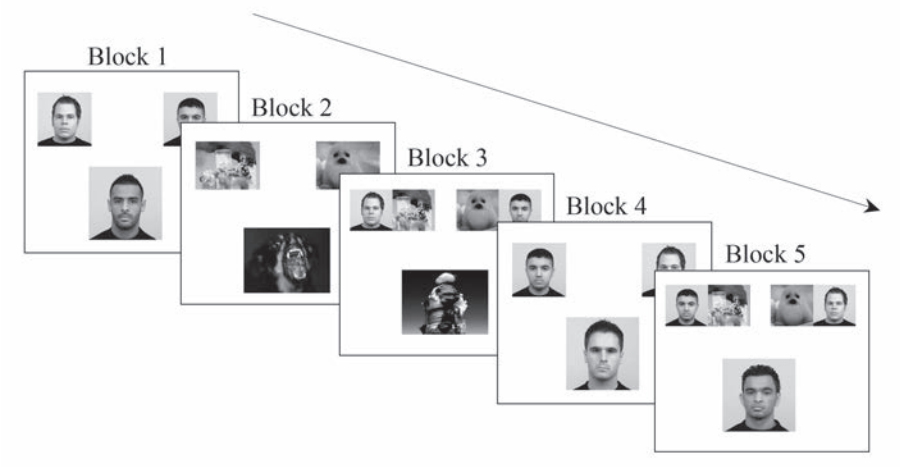

Figure 1. Block design of the PIAT. Here, the adult version of the PIAT is shown. Block 1 and 2 always consisted of trials in which participants had to categorize faces and positive and negative images (in random order). Block 3 is the first critical block in which concepts and attributes are combined. In Block 4, the position of the faces is reversed. Block 5 is the last critical block.

Practice blocks consisted of 20 trials, and critical blocks included 40 trials. Before the start of each trial, a "fixation" dot appeared in the lower-middle part of the screen, which had to be touched in order to start or continue the task (Figure 2A, I). The dot functioned as a fixation cross that directed the participants' gaze and hand towards it, thereby preventing an attentional bias to the left or right side of the screen.

In the first practice block, participants learned to categorize images into superordinate categories (i.e., faces were categorized into the superordinate "Moroccan" and "Dutch" categories, each represented by one exemplar image of a man of Moroccan or Dutch descent). These concepts were presented on the top left and right side of the screen, while the images that had to be categorized were presented in the lower-middle part of the screen. Correct categories could be indicated by pressing on the exemplar image on the screen (Figure 2A, II), and feedback was given in the form of a thumbs-up or -down image[2] (2s), indicating a correct or incorrect answer (Figure 2A, III). Next, a dot appeared that had to be touched to start the next trial (Figure 2A, IV-V). In the second practice block of the PIAT, participants again categorized images into two categories, but this time the images represented a positive or negative attribute dimension. In block three (the first critical block), the

---

[2] Populations unfamiliar to this sign will in a separate learning phase need to learn that thumbs up means 'good' and down, 'bad'. Alternatively, different culture-specific signs can be implemented instead.

concepts from block one and attribute dimensions of block two were combined (Figure 2B, and see also Figure S2A in the supplements). In practice block four, the spatial location of the concepts (Moroccan or Dutch descent) was switched, so that the concept that was on the right side in block 1 was now on the left side, and vice versa. Block five was similar to block three and formed the second critical block, but

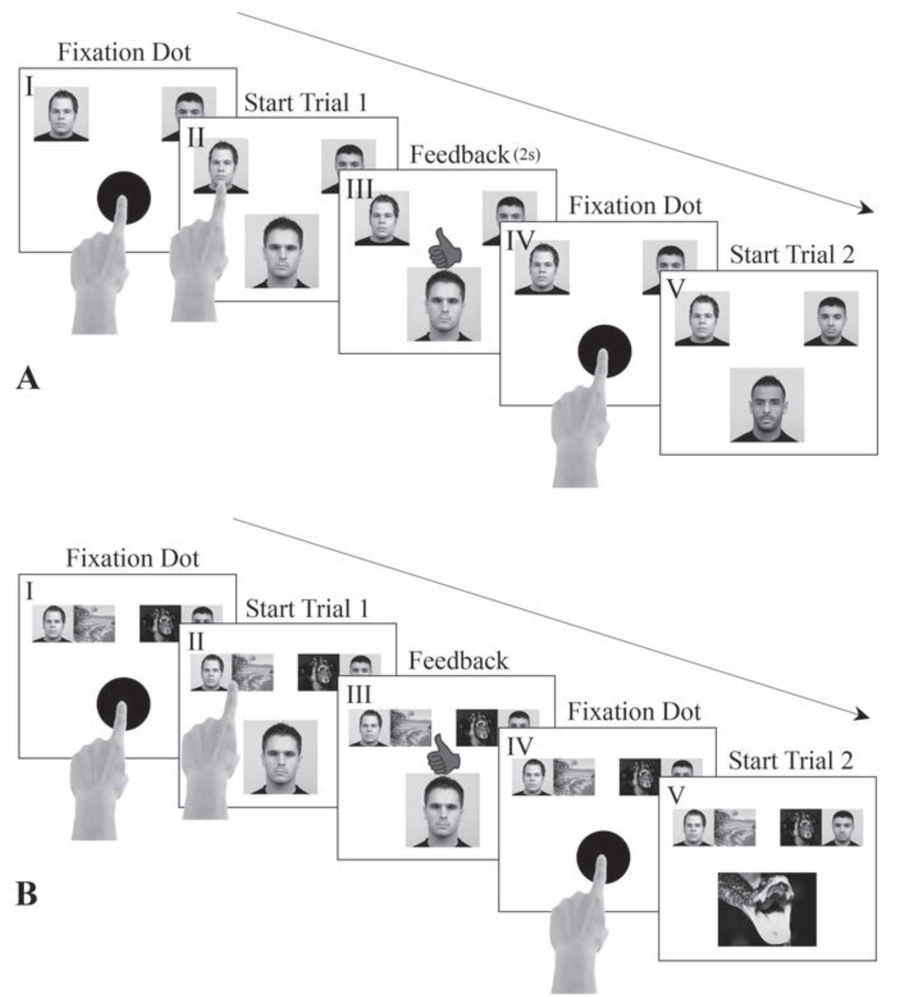

Figure 2. (A) Trial outline for a practice block in the PIAT. After pressing the dot (I), participants categorize the image they are presented with by tapping on the correct concept (faces), attribute dimension (positive/negative scenes), or a combination of the two appearing in the top left and right corner of the screen (II). Participants then receive feedback in the form of a thumbs-up or -down (III, presented for 2 s). Next, a new dot appears in order to start the next trial (IV, V). (B) Trial outline for the critical blocks of the adult PIAT. The trials in the critical blocks follow the same procedure as in the other blocks. In the child PIAT, attribute images are replaced by cartoon figures.

this time the concept-attribute combination was switched (i.e., if participants saw the face of a man of Moroccan descent combined with a negative image in block three, they now saw the face of a man of Moroccan descent combined with a positive image, Figure 2A and 2B).

The IAT is known for its order effects such that when congruent trials are presented first in the critical block, larger IAT effects are found than when incongruent trials are presented first (Nosek et al., 2005). Order effects can be counteracted by counterbalancing the presentation order of the critical blocks, and therefore half of the participants started out with congruent trials (i.e., "Dutch" + positive, "Moroccan" + negative), while the other half started out with incongruent trials (i.e., "Dutch" + negative, "Moroccan" + positive). Furthermore, in the practice blocks, participants either started out with categorizing faces (concepts) or with categorizing positive and negative scenes (attributes). Reaction times on trials in the critical blocks were expected to depend on the congruency of the trials, i.e., following our hypothesis, participants were expected to respond slower to incongruent trials in which faces of men of Dutch descent were linked to negative attributes, and the faces of men of Moroccan descent to positive attributes.

The PIAT was performed on a touchscreen using only one hand. This is different from the typical procedure where participants use their left and right hand to press a left and right key on a keyboard. Previous studies have shown that handedness and the assignments of the left or right response key to a particular IAT category have little to no influence on the IAT effect (Greenwald, 2001; Greenwald et al., 1998), thus we expect no difficulties with using one hand in the PIAT.

Stimuli[3]

Adult

In non-verbal versions of the IAT, stimuli usually consist of pleasant and unpleasant words (attributes) and words referring to the two categories that are being investigated (concepts). In the PIAT, concepts and attributes were replaced with images. Concepts

---

[3] Given the sensitivity of topics such as ethnic prejudice and discrimination, I wish to provide some more context to our choice for measuring implicit racial attitudes (and not, for instance, implicit attitudes towards insects, established in prior studies (Greenwald et al., 1998)). Originally, when designing the study, we aimed to validate it in humans and subsequently test bonobos and chimpanzees on their implicit associations with familiar group mates (the so called "ingroup") and unfamiliar others ("outgroups"). By doing so, we wanted to gain more insights into the evolutionary roots of discrimination. We therefore chose to test the implicit association that humans may have with their ingroup (e.g., individuals of the same descent) versus one potential outgroup (e.g., individuals of Moroccan descent).

consisted of six images of faces of men of Moroccan or Dutch descent with a neutral expression and formed a subset of the Radboud Faces Database (Langner et al., 2010) (see Table S1.1 and Figure S2A in the supplements). For the attributes, we selected the six most negatively and most positively rated images from the International Affective Picture System (Lang et al., 2007), excluding images showing humans as those may interfere with measuring inter-ethnic attitudes (Table S1.2 and Figure S2B in supplements). All images were presented in color, and all images were presented twice during the two critical blocks.

To control for an effect of attribute image type on the results, three different stimulus sets were created (Figure S3A in supplements). In all sets, we used the same faces to depict the "Moroccan" and "Dutch" superordinate category, but varied the images depicting the positive and negative category. For example, in the first version of the stimulus set, the positive attribute was represented by a seal pup and the negative attribute by a building on fire. For each stimulus set, we created four different versions in order to control for order effects and stimulus location effects within the task.

Images that served as category indicators were 300x300 pixels and were presented in the top-left and top-right corner of the screen. In the experimental trials (block 3 and 5, but see *Task*) a combined image of a face and a positive or negative attribute was shown on the top-left and top-right corner of the screen with a dimension of 500x225 pixels (see Figure 1B). Finally, images that needed to be categorized appeared in the lower-center part of the screen and were 400x300 pixels (attributes, i.e., positive/negative scenes) or 450x450 pixels (concepts, i.e., faces).

Children

The faces used as stimuli in the adult PIAT were also used in the child PIAT. Some IAPS images can be upsetting or frightening for young individuals, thus we opted to use cartoons instead. Positive and negative attribute images were changed to cartoon heroes and villains from animated tv-shows (see Table S1.3 and Figure S2C in the supplements for an overview of the selected images). All these attribute images were rated based on valence and arousal by children, and thus validated before we commenced the study (see Table S1.3 in supplements for an overview of the results). Like in the adult PIAT, we used six different positive and negative attributes, each consisting of a hero or villain from the same cartoon (Figure S2C in supplements). The child PIAT followed the same procedure and task presentation as the adult PIAT (i.e., the presentation order between practice and critical blocks were counterbalanced).

Furthermore, like in the adult PIAT, three different stimulus sets were created in which the positive and negative attribute images were varied. Again, for each stimulus set, we created four versions to control for order effects (see Table S2 in supplements and Figure S3B).

Equipment

The adult PIAT was performed on a Dell S2240Tb touchscreen (21.5 inch, 1920x1080 pixels, 12 ms response time). Children performed the PIAT on an Iiyama T1931SR-B1 touchscreen (19 inch, 1280x1024 pixels, 5 ms response time). Validation of the child PIAT attribute images was conducted on a Panasonic FZ-G1 ToughPad tablet (10.1 inch, 1920x1200 pixels).

Procedure

Participants were recruited by student assistants who approached zoo visitors that passed the test location during their visit. The assistants approached the visitors with information about the studies taking place in the zoo, and visitors were then asked if they were willing to participate in the current study. The goals of the study were deliberately kept vague and only minimal instructions were provided. Participants were told that the current test required them to categorize the big picture (of either a face or scene) into one of two categories on the upper left or right side of the screen, and that the test itself would provide them with feedback on their performance. They were also instructed to only use only one hand. After receiving consent from adult participants and the caregiver of child participants, individuals were seated behind the touchscreen. Participants started out with five practice trials in which they sorted images of flowers and bunnies to get a better idea of how the task looks and works. Next, they completed the five blocks of the PIAT. The task took about 10 minutes to complete. After completion, participants were thanked for their participation and fully debriefed.

Analyses

Analyses were performed in R, using the IATscores package (Richetin et al., 2015). We calculated a D-score using RobustScores (a function within IATscores) based on the following minimum performance criteria: reaction lower than 10,000 ms, and an error rate below 40% for the critical blocks (Greenwald et al., 2003; Nosek et al., 2014). Furthermore, in one of the stimulus sets used in 15 children (Stimulus Set 2, Version 2), one of the trials in critical block 3 wrongly presented participants with superordinate category images that belonged to the last critical block (block 5). For each of the affected participants, the reaction time of the trial was manually set to 10,001 ms to ensure that it would be discarded in subsequent analyses.

For adults, two trials were removed due to the 10,000 ms cutoff, and one participant was excluded due to a high error rate (Mean error in our sample M_{error} = 4.91, *SD* = 6.31). For children, 44 Trials were excluded based on the 10,000 ms criterion (15 trials within one child; the rest divided over 29 children), and one child was excluded based on the 40% error cutoff (M_{error} = 6.92, *SD* = 7.23). Remaining reaction time data were then 10% winsorized (Richetin et al., 2015). Note that we did not exclude erroneous trials; inclusion of erroneous trials increases validity and reliability of the scoring method (Richetin et al., 2015). The D-score represents the difference in reaction times (after processing) between critical blocks divided by the standard deviation of the datapoints in both critical blocks. Positive D-scores indicated an association between faces of men of Moroccan descent and negative images, and faces of men of Dutch descent and positive images.

For adults and children we performed two separate one sample t-tests to establish whether D-scores significantly differed from 0. Furthermore, as presentation order of the critical blocks and task version may affect D-scores, we also fit two separate linear models using sum-to-zero coding. We used *Congruency* (i.e., congruent or incongruent block first) and *Task Version* (Version 1, 2 or 3, reflecting the three different stimulus sets) as fixed effects, with the intercept reflecting the average D-score in our sample. Next, to assess the consistency of results across all items, we correlated results of the first half of the trials within the critical blocks with the second half of the trials using the SplitHalf function in IATScores.

Furthermore, the results of sensitivity power analyses for our main hypotheses can be found in the supplements (supplemental Figure S5).

Results

Adult PIAT

For adults we found a significantly positive D-score average of .24 (*95% CI* [.16, .32], $t(127) = 5.94$, $p < .001$, *Cohen's d* = .52), indicating that adults respond significantly faster on congruent trials ("Moroccan" + negative and "Dutch" + positive) versus incongruent trials ("Moroccan" + positive and "Dutch" + negative; Figure 3 and Table 1). After controlling for *Congruency* and *Task Version*, this effect remained present

(D-score: .24, *95% CI* [.16, .32], $t(124) = 5.97$, $p < .001$). We also found a main effect of order ($F(1, 124) = 14.22$, $p < .001$), meaning that adults who receive incongruent trials first show a higher D-score on average ($M = .39$, $SD = .41$) than adults who receive congruent trials first ($M = .09$, $SD = .46$). We found no main effect of task version on D-score ($F(2, 124) = .07$, $p = .936$, Table 1 as well as supplemental Figure S4). Lastly, when assessing the internal consistency of the adult PIAT, we found split-half reliability of $r = .84$.

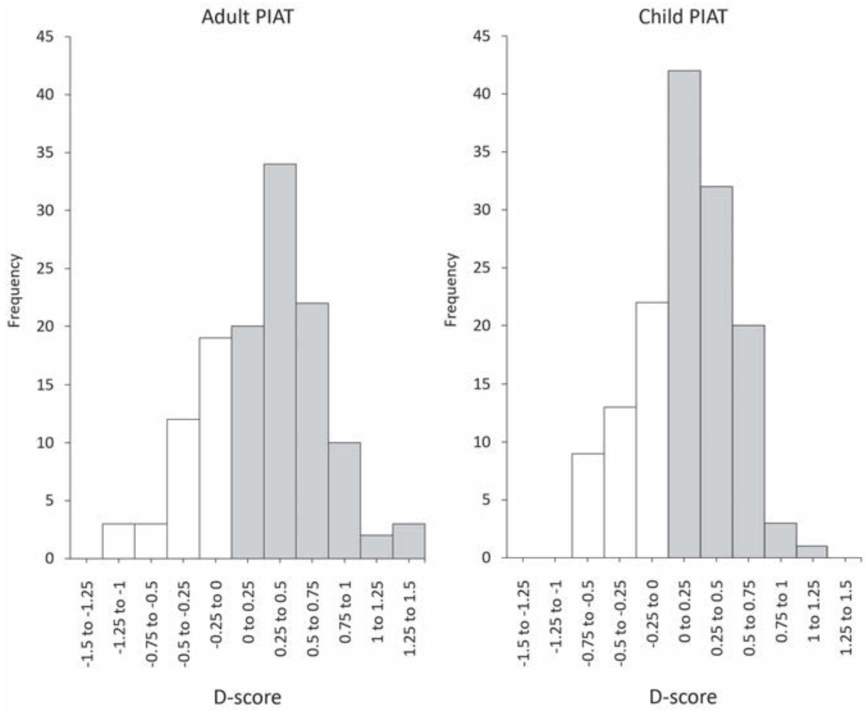

Figure 3. D-score distribution for the adult (left) and child PIAT (right). Positive values represent stronger associations between pictures of men of Dutch descent and positive scenes, and pictures of men of Moroccan descent and negative scenes.

Child PIAT

Children showed statistically significant positive D-score average of .14 (*95% CI* [.08, .21], $t(141) = 4.56$, $p < .001$, *Cohen's d* = .38), meaning that children associate faces of men of Dutch descent with positivity, and faces of men of Moroccan descent with negativity (Figure 3, Table 1). Using a linear model with *Congruency* and *Task Version* added, this effect remained present (D-score: .15, *95% CI* [.09, .21], $t(138) = 4.83$, $p < .001$). We also found a significant effect of *Congruency* on D-score ($F(1, 138) = 9.18$, $p = .003$); children who received incongruent trials first responded significantly faster than children who received congruent trials first (incongruent: $M = .25$, $SD = .34$; congruent: $M = .06$, $SD = .39$, supplemental Figure S4). Additionally, there was a significant effect of *Task Version* on D-scores ($F(2, 138) = 3.08$, $p = .05$, Figure S4). When plotting the data, all task versions cause the IAT effect in the expected direction (i.e., an average D-score above 0), but task version 1 causes the lowest ($M = .05$, $SD = .38$) and task version 3 the highest effect ($M = .18$, SD = .34, task version 2: $M = .17$, $SD = .40$, S3). Indeed, post-hoc comparisons using the Tukey procedure revealed a significant difference between task version 1 and 3 (V1-3: $p = .040$, V1-2: $p = .233$, V2-3: $p = .627$).

Finally, to assess the consistency of results across all items, we correlated D-scores based on only the first half of the trials within the critical blocks with D-scores based on only the second half and found a split-half reliability of $r = .69$, indicating acceptable internal consistency in the child PIAT.

Table 1. Model results for the adult and child PIAT

				D-score		
Adult PIAT	*Predictors*	*Estimates*	*SE*	*95% CI*	*t*	*p*
	(Intercept)	.24	.04	.16 – .32	5.97	**<.001**
	Congruency (Congruent first)	.15	.04	.07 – .23	3.78	**<.001**
	Task Version (V2)	-.01	.06	-.12 – .11	-.13	.900
	Task Version (V3)	-.01	.06	-.12 – .10	-.21	.832
Child PIAT	(Intercept)	.15	.06	.09 – 0.21	4.83	**< .001**
	Congruency (Congruent first)	-.21	.06	.04 – 0.17	3.37	**.001**
	Task Version (V2)	.12	.08	-.19 – .02	-2.33	**.02**
	Task Version (V3)	.19	.08	-.06 – .10	.45	.65

Conclusion

The results show that both children and adults appeared to have a significantly negative implicit bias towards faces of men of Moroccan descent and a positive implicit bias towards faces of men of Dutch descent, but children appear to have a weaker implicit bias than do adults. The PIAT thus appears suitable for testing both adults and children in an environment with a lot of potential distractors (i.e., the zoo). Furthermore, in adults the different versions of the tasks did not significantly impact D-scores, showing that the specific stimuli used in the PIAT do not significantly impact the PIAT's ability to measure implicit attitudes, at least when using the IAPS and Radboud Faces database images. In the child PIAT we do find a significant effect of task version 1 on D-scores. Finally, the PIAT shows decent internal consistency with results that are in line with previous findings in image-based IATs (e.g., split-half reliability $r = .69$ in Palfai et al., 2016).

The IAT can suffer from order effects (Nosek et al., 2005) and despite counterbalancing critical block order, this is also apparent in our PIAT; Individuals who receive incongruent trials first show a much higher D-score average than individuals who receive congruent trials first. This contrasts with common IAT findings that show the reverse effect, i.e., higher IAT effects are found for tasks that present the congruent trials first (Nosek et al., 2005).

To further validate the PIAT, in the next experiment, we investigate whether the effects we find in the PIAT correlate with results on an IAT combining words and pictures (word-IAT or WIAT) within the same subjects, and also study whether both IATs are correlated with more explicit measures of inter-ethnic biases.

Experiment 2: Online PIAT and WIAT

Method

Participants

Initially, 158 adult participants took part in the online PIAT/WIAT study, but 17 did not complete the study, thus resulting in a final *N* of 141 (Age range: 19-68, *M* = 23.72, *SD* = 10.16, 114 females). All participants were native Dutch speakers with a Dutch nationality, and had parents with the Dutch nationality as well. Most of the participants were right-handed (i.e., 128 (81%)). All participants were recruited via an online recruitment system of Leiden University (SONA), through flyers and posters, and through social media. As part of the Psychology curriculum of Leiden University, participants received 1 course credit after completing the experiment. Data collection took place between June 2018 and March 2019.

Tasks

Participants took part in two 7-block IATs; a word-IAT in which individuals categorized pictures into categories represented with words (WIAT) and our picture-only IAT (PIAT). In both IATs, there were initially two training blocks (block 1 and 2), each consisting of 20 trials. Next, participants continued through two critical blocks (block 3: 20 trials, block 4: 40 trials), followed by another training block (block 5: 40 trials). Finally, participants completed two critical blocks again (block 6: 20 trials, block 7: 40 trials). In contrast with Experiment 1 that was conducted on touchscreens, participants now performed the tasks online while using a keyboard. Participants used the “E” and “I” keys to indicate the left and right superordinate categories of the concepts (faces) and attributes (positive and negative images), respectively, and we used their reaction time on the key presses as a measure of bias. Furthermore, whereas on the touchscreen-based task participants had to press a dot to continue, participants were now shown a fixation cross in the middle of the screen for 300 ms before the next trial started.

The task design for the WIAT and PIAT was similar to Experiment 1, but the two critical blocks each contained an extra 20 trials. In the first training block, participants categorized perceived ethnic concepts (“Moroccan” vs. “Dutch”), and in the second training block attributes (positive vs. negative). The third and fourth blocks, i.e., critical blocks, presented participants with the combined concepts and attributes. In the fifth (training) block, participants had to categorize attributes again, but this time the attributes switched positions on the screen (e.g., when the positive attribute was presented on the left side of the screen in block 2, it was now positioned on the right side of the screen). Critical blocks 6 and 7 once again presented participants with the combined concepts and attributes, but this time the position of the attributes was switched relative to the position in critical blocks 3 and 4 (i.e., if “Dutch” + negative was presented on the left side in block 3 and 4, it was now presented on the right side).

For both the WIAT and PIAT, participants were issued one of four versions that varied on the following randomized factors: the starting position of the concept (i.e., left side or right side of the screen), and whether the concept is expected to be congruent or incongruent with the outgroup negativity bias. See Table S3 for more details.

In the word-IAT, words were used to indicate the superordinate categories (concepts and attributes), and pictures for the to-be categorized stimuli. For the superordinate categories, the concepts were written in a black font as “Nederlands”

(*Dutch*) or “Marokkaans” (*Moroccan*). The attributes were written in a green font as “positief” *(positive)* and “negatief” (*negative*). In the critical blocks where combined categories were presented, the concept and attribute words were written on top of each other, separated with an “of” (*or*). Their order (top or bottom) was randomized. The to-be categorized stimuli were always presented in the lower part of the screen in the center, just like in the PIAT in Experiment 1, and consisted of the same images as in Experiment 1. The PIAT in Experiment 2 was similar to the one in Experiment 1, with the difference being that there were now two more critical blocks and answers were given via pressing “E” and “I”, rather than touching the screen.

Symbolic Racism 2000 Scale

The Symbolic Racism 2000 Scale (henceforth: SRS) was created to assess explicit inter-ethnic biases via a series of eight questions (Henry & Sears, 2002). The SRS specifically looks at a modern variant of discrimination in the form symbolic racism, or the belief that discrimination based on ethnicity is no longer impacting people of non-Dutch descent’s chances to thrive and that continuing disadvantages are attributable to their lack of responsibility for their own lives (Henry & Sears, 2002). The SRS issues participants with questions regarding work ethic and responsibility of outcomes, excessive demands, denial of continuing discrimination, and undeserved advantages. Participants answer most questions on a 4-point Likert scale ranging from “strongly disagree” to “strongly agree”, but questions 3, 4 and 5 involve different types of answers. The questions in the SRS are specifically attuned to the cultural history of people of color residing in the United States. To make the SRS applicable to our study group, we translated the English questions to Dutch, and replaced the words “African-American”, “United States”, “Irish, Italian, Jewish, and many other minorities” to “Marokkaans” (“Moroccan”), “Nederland” (The Netherlands), “Mensen met een Surinaamse of Poolse afkomst, of andere minderheden” (“People of Surinamese or Polish descent, and other minorities”), respectively.

Stimuli

The stimuli were the same as the ones used in the PIAT in Experiment 1. However, whereas in Experiment 1 there were different versions of the PIAT using different exemplars for the attribute categories, in Experiment 2 we used only one exemplar for the positive and negative attributes (i.e., the rabid dog and the seal pup from the IAPS) to ensure that we had enough participants per versions of the tasks. The PIAT and WIAT consisted of four versions, which each differed in a) whether incongruent trials were presented first or second, and b) the location of the stimuli on the screen (e.g., faces of men of Moroccan or Dutch descent, or positive/negative images and text on the left/right side of the screen, see Table S3 in supplements for an overview of all task versions). All images were presented in color, and all images were presented a maximum of two times during the two critical blocks.

Equipment

Experiment 2 was conducted online through Qualtrics and by using Iatgen, a pre-programmed survey-software implicit association test (Carpenter et al., 2019). As the original survey software uses words only, we adapted it to work with pictures. Participants using mobile devices such as tablets and smartphones were not allowed to participate in the study.

Procedure

Participants who signed up to take part in the study were sent a link to the tasks. Via Qualtrics, participants first received brief information about the goal of the study, namely to compare two types of categorization tasks containing faces and scenes. They were also told that they would receive more information about the study after completing the experiment. If participants were still interested in participating, they signed a digital consent form to allow us to use their data. Next, participants were issued questions about their age, gender, handedness, native language, and their own and their parents' nationality. The study would be terminated with a custom message if the participants were below of the age of 18, or if they or their parents did not have the Dutch nationality. If participants passed the screening, they were notified the first categorization task would start when they continued on to the next screen. To keep the online experiment as similar as possible to the PIAT in experiment 1, this was all the information that was given.

Every participant took part in both the WIAT and the PIAT. The order of the IATs was counterbalanced, and in-between each IAT the participant completed the SRS. Between each of the tasks, participants could decide to take a break for as long as they wanted. At the end of the experiment, participants were asked questions about their prior experience with people of Moroccan descent, i.e., whether they knew anyone of Moroccan descent and if they did, how well they knew this person or these persons. This information was not used in subsequent analyses, as it was part of a different research project. Participants were then given a full debriefing on the goals of the study, and were thanked for their participation.

Analyses

D-score calculations and analyses were performed in R, using the IATScores package (Richetin et al., 2015). As per the suggestion of Richetin et al. (2015), we did not distinguish between the first 20 trials ("practice" trials) in the critical block and the 40 following trials. The minimum performance criteria were an error rate below 40% for the critical blocks, and reaction times (RTs) higher than 400 ms and lower than 10,000 ms (Nosek et al., 2014). RTs below or above these criteria were discarded. Based on these criteria, we discarded 4 trials with an RT > 10,000, and 23 trials with an RT < 400 (divided relatively equally across 13 subjects) within the PIAT sample. For the WIAT sample, we removed 6 trials (divided over 4 individuals) based on the 10,000 ms cutoff, and 124 trials based on the 400 ms cutoff (divided over 16 individuals, of which one individual had 20 trials meeting this criterion, and another individual 80 trials). No participants were removed based on the 40% error rate cutoff in the PIAT, and one participant in the WIAT. This method is slightly different from Experiment 1, where only RTs above 10,000 ms were discarded and remaining RTs were 10% winsorized. While this is the most robust approach for treating IAT data (Richetin et al., 2015), this was not an option due to an error in data collection. During data collection, trials with where a wrong categorization was made were flagged by the software, but due to an error on our part, RTs for these trials were not saved and thus erroneous trials could not be included in the analysis. Instead, we discarded RTs lower than 400 ms (as per the original scoring method by Greenwald et al., 2003).

We used one sample t-tests to test whether D-scores significantly differed from zero. Furthermore, to assess internal validity, we also fit linear models using *Congruency* (incongruent first vs. congruent first) and *Location* (whether faces of men of Moroccan or Dutch descent were presented left or right) as fixed effects (sum-to-zero coded) in two separate analyses (PIAT and WIAT), with the intercepts reflecting the average D-score. Note that instead of *Location*, we used *Task Version* as fixed effect in Experiment 1. This was done because the versions in Experiment 1 differed on multiple fronts (i.e., stimuli and location on the screen), whereas in Experiment 2, versions of the tasks only differed in the location of the stimuli on the screen. Next, we correlated D-scores on the PIAT with the WIAT in order to assess test-retest reliability. In addition to comparing the IATs to each other, they were also compared with the explicit bias measure (SRS) in order to assess discriminant validity. The data of the SRS were first converted to a continuous 0-1 scale as described by Henry & Sears (2002), and subsequently used for a correlation calculation. Finally, we also investigated whether providing the SRS in-between the two tasks (and thus before one of the two tasks) affects subsequent D-scores on the final task (e.g., due to a priming effect). For this, we first calculated a difference score between performance on the tests by subtracting the D-score from the second task from the first. We then fitted a linear model using *First Task* (PIAT first or WIAT first) a sum-to-zero coded fixed effect and *Difference Score* as the dependent variable. Furthermore, see supplemental Figure S5 for our sensitivity power analysis results.

Results

Implicit associations

For the PIAT, we found a significantly positive D-score average of .18 (*95% CI* [.07, .28], *t*(139)= 3.34, *p* = .001, *Cohen's d* = .28), meaning that participants associated faces of men of Moroccan descent with negativity, and faces of men of Dutch descent with positivity (Figure 4). In the linear model controlling for *Congruency* and *Location*, this effect remained present (D-score: .17, *95% CI* [.07, .27], *t*(137)= 3.36, *p* = .001). We also found a significant effect of *Congruency* on D-scores (*F*(1, 137) =7.03, *p* = .009),

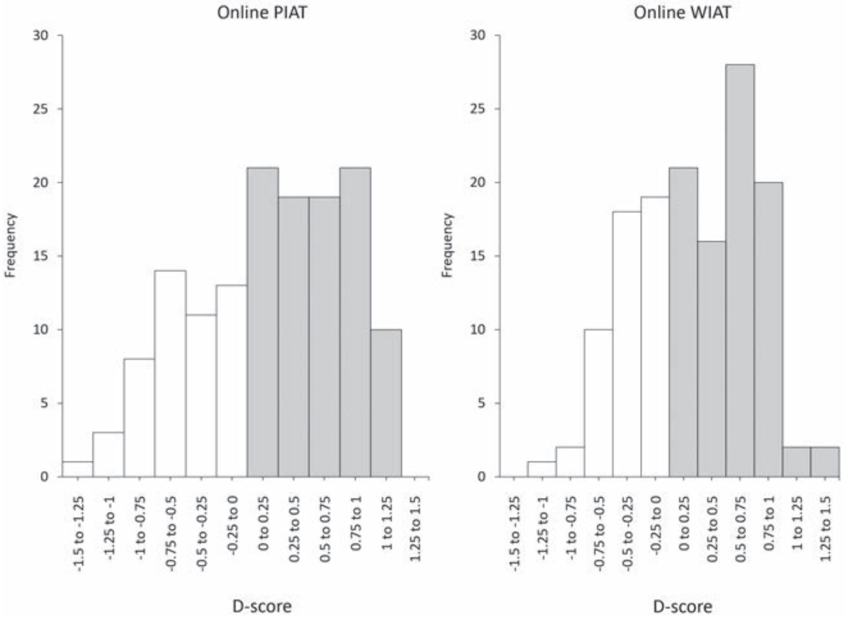

Figure 4. D-score distribution for the online PIAT (left) and online WIAT (right). Positive values represent stronger associations between faces of men of Dutch descent and positive scenes, and faces of men of Moroccan descent and negative scenes.

with individuals receiving congruent trials first having a higher D-score on average (*M* = .31, *SD* = .64) than individuals receiving incongruent trials first (*M* = .04, *SD* = .58, supplemental Figure S6 and Table 3). Finally, we did not find an effect of *Location* on D-score averages (*p* = .194, Table 3).

Results of the WIAT indicate a significantly positive D-score average of .22 (*95% CI* [.13, .31], *t*(138) = 4.91, *p* < .001, *Cohen's d* = .42), similar to what the PIAT showed (Figure 4). In the linear model this finding held up (D-score: .22, *95% CI* [.13, .31], *t*(136) = 4.88, *p* < .001). However, there was no *Congruency* effect (*p* = .269), nor an effect of *Location* (*p* = .409, supplemental Figure S6, Table 3).

Finally, there was a strong positive correlation between individuals' scores on the PIAT and WIAT (*N* = 139, *r* = .69, p < .001, Figure 5), indicating that participants who had an implicit bias in one of the tasks showed a similar bias in the other task.

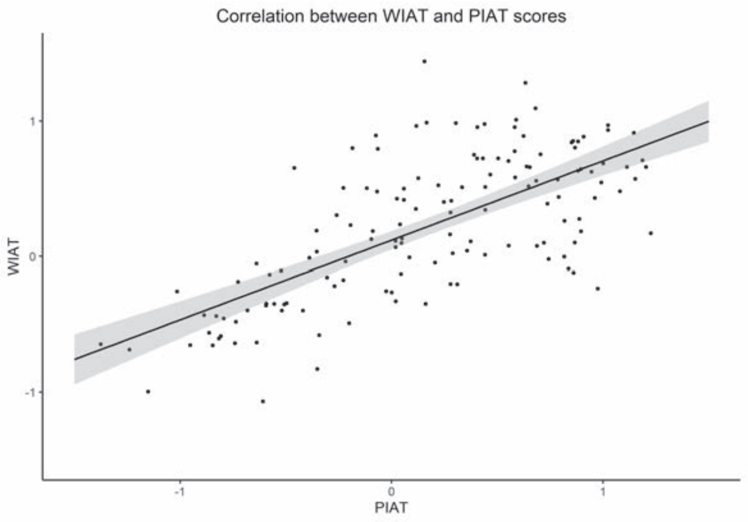

Figure 5. Scatter plot showing the correlation between WIAT and PIAT D-scores (*r* = .69, *p* < .001).

Correlation between the WIAT, PIAT and explicit measures

All 144 participants completed the Symbolic 2000 Racism Scale. On average, individuals had an SRS score of .33 (*SD* = .13, with scores ranging between .04-.64), meaning that explicit symbolic racism among our participants was low (Henry & Sears, 2002). For both the WIAT and the PIAT, we did not find a significant correlation between implicit biases (D-scores) and explicit symbolic racism score (WIAT: *r* = -.02, *p* = .801; PIAT: *r* = -.01, *p* = .934). To test whether our results indicated evidence for the null-hypothesis, we performed an additional correlation analysis using Bayesian statistics (using the R-packages BayesFactor (Rouder et al., 2009) and bayestestR (Makowski et al., 2019b). We found a BF_{01} = 5.1 for the correlation between PIAT and SRS scores, and a BF_{01} = 4,95 for the correlation between WIAT and SRS scores, meaning that in both cases the data are around 5 times more likely under the null hypothesis that there is no correlation between the variables. The results therefore indicate moderate evidence for the null hypothesis (i.e., there is no correlation between the measures (Lee & Wagenmakers, 2013). As such, the implicit biases reflected in the D-scores in both versions of the IAT do not seem to correlate with explicit inter-ethnic biases in our participant pool.

Finally, we also assessed whether performing the SRS in-between tasks affected D-scores on the final IAT. We found that the difference score did not significantly differ from zero ($t(137) = -1.14$, $p = .255$), meaning that D-scores in the second task were not significantly higher or lower than the D-scores in the first task. Furthermore, we found no significant effect of *First Task* ($t(137) = -1.27$, $p = .206$) on the difference score, showing that the difference score was not affected by whether a participant first started with the PIAT or the WIAT. In short, performing the SRS in the middle of two IATs did not significantly impact D-scores on the last IAT, and this was regardless of whether participants started with a PIAT or WIAT.

Table 3. Model results for the online PIAT

Task	D-score					
PIAT	*Predictors*	*Estimates*	*SE*	*95% CI*	*t*	*p*
	(Intercept)	.17	.05	.07 – .27	3.36	**.001**
	Congruency (Congruent first)	-.14	.05	-.24 – -.03	-2.63	**.009**
	Location ("Moroccan descent" Left/ "Dutch descent" right)	.07	.05	-.03 – .17	1.31	.194
WIAT	(Intercept)	.22	.04	.13 – .31	4.88	**<.001**
	Congruency (Congruent first)	-.05	.04	-.14 – .04	-1.11	.269
	Location ("Moroccan descent" Left/ "Dutch descent" right)	-.04	.04	-.13 – .05	-0.83	.409

Conclusion

Participants performed similarly on the PIAT and WIAT: in both IATs, participants appeared to have a an implicit bias in the predicted direction. Although the order in which participants completed the critical blocks significantly impacted the D-scores in the PIAT (i.e., participants with congruent trials first had a higher D-score average than participants with incongruent trials first), this effect was not present in the WIAT. The PIAT and WIAT results were significantly correlated and showed good (within-participant) test-retest reliability. Finally, implicit biases measured with the IATS did not correlate with explicit measures of inter-ethnic bias measured through the SRS, nor did the SRS impact performance on the IAT that followed it. The results extend our findings from Experiment 1, showing that the PIAT can tap into the same implicit biases as the more commonly used WIAT.

Discussion

The aim of this study was to design and validate a non-verbal, intuitive pictorial IAT (PIAT). Experiment 1 shows that the PIAT can tap into implicit inter-ethnic attitudes in a large group of participants including adults and children, and can do so reliably using different stimuli. Furthermore, the PIAT can do this outside of a lab setting and on a representative subject population involving participants of different ages (as opposed to only university students). In Experiment 2 using a within-subjects design, the performance of the PIAT was comparable to a more typical word IAT that has been rigorously tested in the last two decades (Dunham et al., 2006; Greenwald et al., 1998, 2003; Kurdi et al., 2019; Nosek et al., 2002a, 2013; Oswald et al., 2015). As such, the PIAT could be standardized tool that enables future studies to make direct comparisons across different cultures, age groups, and potentially also between species.

Internal validity

Although we counterbalanced the order of the presentation of critical blocks, participants who received incongruent trials first and congruent trials second in the PIAT in Experiment 1 showed higher D-scores on average than participants who received a reversed order. Interestingly, this effect was reversed in the PIAT of Experiment 2, and absent in the WIAT. IAT order effects are well documented, and may impact the magnitude of the found IAT effects (Greenwald, 2001). An explanation for their existence is a cost of switching tasks in the two critical blocks, namely in the form of increased reaction time latencies and error rates (Mierke, 2001; Mierke & Klauer, 2003). Furthermore, these task switching costs may remain for quite some time after switching, as switching requires the activation of the appropriate action and suppressing the previous, competing one. Nevertheless, we find diametrically opposed results in the order effects, which suggests that order effects likely occurred due to noise or random differences between the groups of participants. Our data therefore do not clearly support an effect of block sequence on IAT scores. Order effects remain a topic of debate in the IAT literature, and thus deserve more attention in future studies.

Psychometrics (internal consistency)

The PIAT shows an acceptable internal consistency in children, and good internal consistency in adults. The result of the child PIAT is somewhat lower than internal consistencies revealed in a comparable PIAT used in children (e.g., $\alpha = 86.5$ on average across two PIATs (Cvencek et al., 2011)), which could be explained by the more noisy setting in which the PIAT was distributed. Generally, our results are in line with the limited amount of studies using image-based IATS that report internal consistency values (Brand et al., 2014; Palfai et al., 2016; Slabbinck et al., 2011). Importantly, despite the less-controlled circumstances in which we conducted the experiment, the PIAT reveals ethnicity-based implicit associations consistent with previous findings (Dunham et al., 2008; Greenwald et al., 1998).

Choice of stimulus material

For all IATs, we created several different versions of the task that differed on the stimuli that represented the attribute dimensions (in Experiment 1: PIAT Version 1, 2 and 3), or differed in the order of critical block presentation and the location of the concepts and attributes on the screen (All IATS in Experiment 1 and 2). We found that children showed a lower D-score average in version 1 of the PIAT in Experiment 1, and that the spread in D-scores was higher in this version than in the other versions. This effect was, however, not present in the adult PIAT nor in the online PIAT and WIAT. It is important that the differences in the attribute dimensions are clear (i.e., clearly negative and positive) as the IAT effect relies heavily on responses that do not require a lot of deliberation (Lane et al., 2007), which is why we used different types of stimuli that were rated by children in Experiment 1 on valence and or whether they were low or high in arousal. We found that the ratings for the different stimuli were very close to each other (see Table 1c in supplements), thus it is unclear what it is about the stimuli used in task version 1 that resulted in lower D-score averages compared to the other versions. The IAT represents a family of instruments where differences in e.g., choice of stimuli can result in entirely different results, even when the versions were built with the aim to measure the same underlying construct (Foroni & Bel-Bahar, 2009). As such, it remains crucial to choose stimuli that reflect the same superordinate category that one is interested in as much as possible (Lane et al., 2007), but also to include several different stimuli (Nosek et al., 2005) as we aimed to do in our study.

Discriminant validity

Both the PIAT and WIAT do not correlate with the explicit measure of inter-ethnic bias (the symbolic racism scale, or SRS), which suggests that IATs tap into a different kind of cognitive constructs than explicit biases. Indeed, this is the reason why IATs exist in the first place, namely to show that views and attitudes are partially driven by unconscious mechanisms (Greenwald et al., 2002). At the same time, we did not find high antipathy scores against individuals of Moroccan descent, which could for instance reflect that a) our participants indeed do not consciously view individuals of Moroccan descent as more negatively than individuals of Dutch descent, or b) participants answer in a socially desirable way, which can be true especially in case of a highly sensitive topic such as ethnic prejudice (Fazio & Olson, 2003). In the current study we cannot directly dissociate between these explanations, but a meta-analysis on the correlation between implicit and explicit measures of attitudes showed that correlations between implicit and explicit measures may be low when participants make their judgments deliberately or spontaneously. For instance, it takes more cognitive effort when asked to reflect on your evaluations of individuals with differing backgrounds than whether you are asked about more mundane things such as attitudes towards fruits and candies; in the latter case, correlations between explicit an implicit measures are higher (Hofmann et al., 2005) .

In general, IATs that find ethnicity biases do indeed report negative correlations between IAT effects and explicit biases, and our results are in line with these findings. Furthermore, the notion that the PIAT indeed uncovers implicit biases is supported by the fact that the Symbolic 2000 Racism Scale questionnaire, which primes participants to think more deeply about their inter-ethnic biases, did not seem to affect subsequent IAT effects. For further discussions on the correlations between explicitly and implicitly assessed attitudes we refer to the meta-analysis by Hofmann et al. (2005).

PIAT performance compared to the WIAT

In Experiment 2 we show that the PIAT performs similarly to a word-IAT, and that the test-retest reliability of the IAT measures was good. This is interesting considering the persistent debate on the relatively large range of test-retest reliability scores of the IAT (Lane et al., 2007). The benefit of a PIAT over IATs that use words or spoken language are that it is applicable to a wider variety of populations, and that the same test (i.e., without having to translate words) can be used even in populations that are very different (e.g., because of culture, language, cognitive ability), thus making direct comparisons possible. The stimuli we used for the (adult) PIAT were selected from the cross-culturally validated International Affective Picture System (Lang et al., 2007), and the IAPS is one example of what researchers can use to study e.g., cultural differences in implicit attitudes. At the same time, while the PIAT has not been validated for non-human animals, it could potentially be a useful tool to study implicit attitudes in for instance great apes, as they are highly capable of extracting emotionally relevant information from scenes and can be trained on the use of a touchscreen (Altschul et al., 2017; Kret et al., 2016; Perdue et al., 2012). For animals, more appropriate positive and negative images should then be selected (e.g., a favorite food item or an item that holds a negative association).

Conclusion

With the aim of validating a non-verbal PIAT, we found that it can be used to measure implicit biases reliably, and similarly to a standard verbal IAT. As such, we believe it can provide a practical way to study implicit associations in a wide variety of individuals, and conceivably in non-verbal populations. Pictorial adaptations to the IAT have the potential to answer important questions related to the ontogeny and evolutionary development of implicit attitudes, and to directly compare different groups of individuals on their implicit associations. Intergroup conflict in humans is still ubiquitous, and the discussions about the foundations of implicit associations are still ongoing. We therefore deem it crucial to find novel ways to probe these implicit attitudes and make within- and between-species comparisons possible, which we think is the most important role that pictorial adaptations to the IAT can fulfill. By validating the PIAT as a tool, our study sets a first step into that direction, but future studies should look into optimization of the task by testing different kinds of attitudes and using multiple different category exemplars. Ultimately, we hope the PIAT can be added to the steadily growing list of cognitive tasks that can be used in comparative research.

Chapter 8

General discussion

Detecting and correctly recognizing emotional expressions is pivotal to social life, as these expressions guide our thoughts and behaviors, and enable us to understand others' feelings and intentions. They are at the basis for complex phenomena such as empathy (De Waal & Preston, 2017; Koski & Sterck, 2010) and cooperation (Boone & Buck, 2003), which have long been considered to be uniquely human. However, when and why did our high sensitivity to emotions evolve, and are the processes that govern emotion perception truly unique to our species? Do emotions mean anything to other animals? One way to go about answering these questions is to reconstruct the lives of our ancestors, but this is difficult. The fossil record of early humans is sparse (Andrews, 2020) and fossilization of brain tissue is extremely rare. Nevertheless, our closest living relatives, the great apes, offer us an invaluable window into the past, allowing us to indirectly infer the social and cognitive characteristics of extinct humans and other apes (e.g., Wilson, 2021).

Studying great apes can not only inform us about our evolutionary past but can also provide a stepping stone towards understanding the evolutionary pressures that shaped the expression and processing of emotions throughout the animal kingdom. Here, it is important to remember that the social and cognitive abilities of hominids are just *examples* that are part of a diverse collection of exceptional skills that organisms can develop to deal with the social and physical demands of their specific environments (i.e., "There is not 'one cognition'", Bräuer et al. (2020)). That said, directly comparing different species is challenging. Preferably, we would measure emotion perception using the same method for all species involved, without neglecting species-specific characteristics that may impact the expression of emotion perception (for instance, if one species expresses emotions mainly through facial expressions and another species expresses them mainly through vocal signals, using the same testing method may not be optimal). Based on the literature on cognition and behavior in great apes, it is reasonable to assume that there is at least some continuity between how humans and great apes express and perceive emotions (Kret et al., 2020; Nieuwburg et al., 2021). Following this assumption, the studies reported in this dissertation focused on several unconscious and automatic mechanisms underlying emotion perception that can be studied similarly across humans and great apes: *attention*, *spontaneous mimicry*, and *implicit associations*.

The goal of this dissertation was to take a closer look at the differences and similarities in emotion perception in humans, bonobos, and orangutans. Specifically, we examined emotional modulation of *attention* (Figure 1i) and *spontaneous mimicry* (Figure 1ii), and developed a method that could potentially probe *implicit* *associations* in primates and other animals (Figure 1iii). In this final chapter, I will focus on integrating the key findings from each chapter and discuss the implications of the results. At the end of the dissertation, I will point towards new and unanswered questions that can help us move forward in understanding the continuity between human and animal emotions.

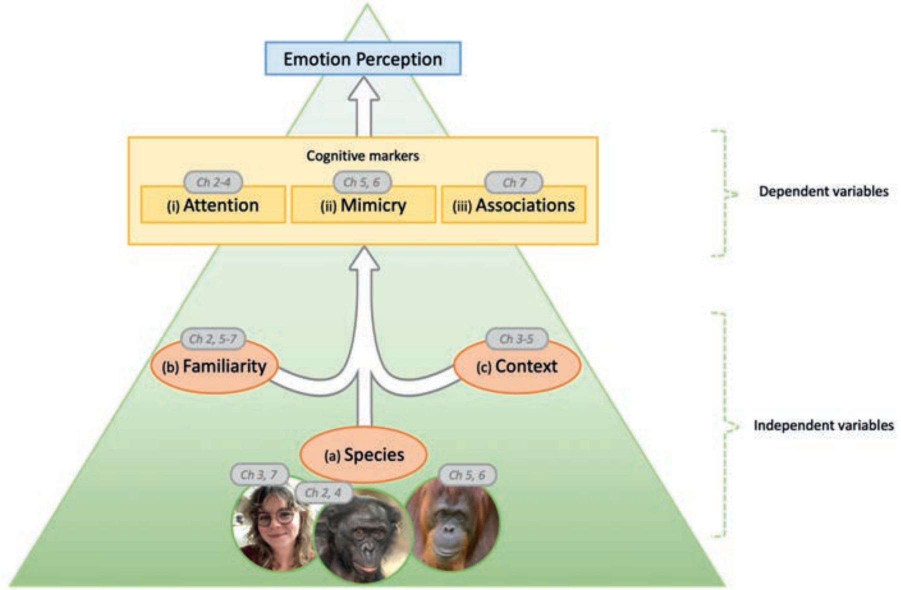

Figure 1. Schematic overview of the research topic. Emotion perception is a multifaceted phenomenon that is governed by many different cognitive mechanisms. Often, these mechanisms operate on an implicit level; automatically and unconsciously. To study emotion perception across species, I investigated its underlying implicit mechanisms or *cognitive markers*. The focus of this dissertation lies on *(i) attention, (ii) mimicry,* and *(iii) implicit associations*. Moreover, I investigated the effects of *species (a), familiarity (b),* and *context (c)* on these markers across six chapters (grey circles) in this dissertation.

Summary of key findings

In **Chapter 2**, I compared an attentional bias for emotions (Figure 1i) between bonobos and humans and examined whether this bias is affected *by familiarity with* or *the species of* the expressor (Figure 1a, b) across a series of three experiments using the dot-probe task. In Experiment 1, bonobos were presented with emotional and neutral scenes of familiar and unfamiliar conspecifics, and we found that the attention of bonobos was automatically tuned to emotions of unfamiliar conspecifics. In Experiment 2, we examined whether this emotion-biased attention in bonobos also occurred when bonobos were exposed to human facial expressions of emotions, again of familiar individuals (caretakers) or unfamiliar individuals. Bonobos did not show an attentional bias towards human expressions of emotions, even though in a control experiment using the same stimuli, humans did show this bias. Nevertheless, low power may explain the null findings in bonobos. In Experiment 3 using a heterogeneous human sample, we found that in contrast to bonobos, humans showed an attentional bias towards emotions of familiar, socially-close conspecifics. The findings of this study underline that an attentional bias for emotions is driven by factors that hold motivational relevance to the observer (i.e., socially-relevant characteristics such as the species of the expressor and social closeness) (Brosch et al., 2013). Importantly, it indicates continuity between the attentional mechanism (i.e., selective attention) underlying emotion processing in humans and bonobos, but also that attention to emotions may be most sensitive when the expressors of these emotions are conspecifics.

As we did not study whether humans have an attentional bias for emotions of bonobos, we attempted to close this knowledge gap in the study described in **Chapter 3** (Figure 1i). Moreover, most research into emotion-biased attention in humans has used isolated facial expressions of emotion, thereby not fully appreciating the role that whole-body emotional expressions play in recognizing emotional expressions, as these whole-body expressions embedded in a scene can provide more context (Figure 1a, c) (Kret et al., 2013a). Additionally, many studies have looked only at specific age classes or homogeneous populations such as university students, which are not good examples for generalizing about humans (Henrich et al., 2010). We partially tackled these issues by examining how a more heterogeneous (i.e., non-university) group of human adults and children rate emotional *scenes* of humans as well as bonobos on valence and arousal, and by measuring their attentional bias towards these scenes. Overall, humans perceived emotional scenes of other humans to be similar to scenes of bonobos in terms of valence (positivity or negativity) and arousal (intensity). However, children misinterpreted the bared-teeth display as a positive expression.Humans also showed an attentional bias towards emotional scenes of both species, but the bias was strongest for human emotional scenes. These findings could suggest a shared evolutionary origin for emotional expressions and their perception (Kret et al., 2020), but also show that in humans, a learning component may be important for understanding emotional expressions by other species such as bonobos.

As a next step, we examined whether an attentional bias towards emotions in bonobos and humans does not only occur initially when early attentional mechanisms (e.g., covert attention) are recruited but also when emotional scenes are presented for a longer duration (**Chapter 4**). Bonobos and humans were presented with emotional and neutral scenes of conspecifics and heterospecifics (Figure 1i, a, c). Each emotional and neutral stimulus combination was shown for three seconds, thereby recruiting not only initial, more reflexive attentional mechanisms, but also more voluntarily controlled attention. We found that humans overall showed a more pronounced emotion bias than bonobos, i.e., humans looked longer to emotional scenes of different categories (play, grooming and embracing, sex, distress, and yawning), whereas bonobos only looked longer to scenes of distressed bonobos or bonobos having sex. For both species, distress was the most salient emotional category, holding attention the longest. Moreover, we extended our previous findings (Chapter 2) indicating a lack of attentional bias to human emotional expressions in bonobos. Conversely, humans looked longer at grooming and playing bonobos: two categories also rated most positively in the study of Chapter 3. In general, humans and bonobos appeared to be most sensitive to the emotions of individuals of their own species. Additionally, humans appeared sensitive to the emotions of bonobos, but bonobos not to the emotions of humans.

In **Chapter 5**, I moved away from attention and focused on another unconscious process at stake during social interactions: spontaneous mimicry (Figure 1ii). Specifically, we investigated contagious yawning and self-scratching in relation to social closeness and context in orangutans: our most distantly related relatives within the great ape family (Figure 1a-c). Nine orangutans were observed for a period of four months, and all occurrences of yawning and self-scratching and the context within which they occurred were recorded. There was only sufficient data on self-scratching, as yawn occurrences are generally rarer. The results showed that in orangutans, self-scratching was indeed contagious, and most strongly so during situations where there was tension, and between individuals with a relatively weak social bond. The novelty of the findings is that this study presents a potential case of *negative* emotion contagion, rather than the more typically presented positive emotion contagion.

In **Chapter 6**, I reported on an experiment on contagious yawning in orangutans (Figure 1ii). Orangutans were presented with videos of real, unfamiliar orangutans, familiar orangutans (conspecifics that were also housed in Apenheul), and a 3D avatar orangutan (Figure 1a, b). The orangutans in the videos either yawned or showed a neutral expression, allowing us to investigate whether yawns occur more frequently in response to other yawns compared to neutral videos. For the first time, orangutans were shown to yawn contagiously, but only in response to yawn videos of *real* orangutans. The avatar did not elicit more yawns, and there was no difference between yawns in response to familiar or unfamiliar orangutans. These results are interesting for several reasons: i) orangutans show contagious yawning despite their more semi-solitary nature (previously, contagious yawning was only found in highly social species (e.g., Norscia et al., 2020)), ii) the yawning avatar did not elicit more yawns, even though a previous study found a contagious yawning effect in response to a virtual representation of a chimpanzee (Campbell et al., 2009), and iii) familiarity did not modulate the occurrence of yawn contagion, even though this is expected if contagious yawning serves some social function or is linked to emotion contagion (Palagi et al., 2020). Overall, the findings of Chapters 5 and 6 indicated that orangutans, despite being classified as the least-social great ape, are susceptible to yawn and self-scratch contagion.

Finally, in **Chapter 7**, I elaborate on a fully pictorial version of the *Implicit Association Test* (PIAT) that we created to eventually test implicit attitudes in great apes (Figure 1iii). The PIAT was designed to be intuitive and require minimal instructions (as the apes cannot be instructed). In this study, we validated the PIAT in children and adults and compared how it performed in relation to the classic IAT using words in an online study with mainly university students. Participants were tested on their implicit attitudes towards individuals of Dutch or Moroccan descent (Figure 1a, b). We chose these specific categories because a race bias is one of the most established IAT effects, thus testing for it was useful for examining the validity of the PIAT. We established that a pictorial version of the IAT can tap into the same implicit attitudes as the original word-IAT in human adults and children. Though we have not yet had the opportunity to test its utility in great apes, I hope that others may find the PIAT useful for inter-species or inter-cultural research.

Theoretical implications

The results in this dissertation converge to two main conclusions. Firstly, the mechanisms underlying attention for emotions as well as automatic, spontaneous mimicry are likely to be conserved ancestral traits among the hominids. Secondly, species-specific characteristics such as their evolutionary and social environment dramatically shape how these mechanisms are expressed.

Survival and reproduction of a species are driven by adaptive mechanisms that underlie individuals' behavior and cognition (Cosmides & Tooby, 1992). Attention for emotions as well as spontaneous mimicry of behaviors may be two adaptations that are preserved in the hominid lineage. What can the results on emotion-biased attention and mimicry tell us about the most recent common ancestor of humans, bonobos, and chimpanzees? Likely, this ape-like creature was sensitive to the emotions of others, having brain mechanisms in place that guided their attention towards emotional expressions, and that elicited automatic mimicry of facial and bodily expressions. Nevertheless, findings on an attentional bias towards emotions are mixed for chimpanzees, complicating this conclusion. For instance, some studies find no evidence for an immediate bias towards emotions in chimpanzees (Kret et al., 2018; Wilson & Tomonaga, 2018), whereas another study reports that chimpanzees preferentially look at agonistic scenes (Kano & Tomonaga, 2010a). Evidence from behavioral observations, however, corroborate the idea that chimpanzees and bonobos are both sensitive to the needs of others. For instance, both species are known to console each other after stressful situations (e.g., de Waal & van Roosmalen, 1979; Goldsborough et al., 2019; Palagi & Norscia, 2013). In addition, there is evidence for contagious yawning in bonobos and chimpanzees (e.g., Campbell & de Waal, 2014; Demuru & Palagi, 2012), as well as for facial mimicry (Palagi et al., 2019b, 2020a), again suggesting that the basic mechanisms for emotion perception (De Waal & Preston, 2017) were already present in the last common ancestor of bonobos, chimpanzees, and humans.

If we look further down the phylogenetic tree to the common ancestor of humans and orangutans (Figure 2 in the introduction), we still see some evidence for sensitivity to the emotions of others. Recently, we tested orangutans' attentional bias towards emotional scenes versus neutral scenes, and our preliminary findings indicated a lack of an attentional bias towards emotions in these great apes. The only other study to date that has looked at emotion-biased attention in orangutans found that they preferentially look at the silent bared-teeth face compared to the bulging-lip face (shown during aggressive encounters), but a direct comparison between the silent bared-teeth face and a neutral face was not significant (Pritsch et al., 2017). Clearly, more research is needed to understand how and to what extent orangutans perceive the emotional expressions of others. Importantly, we find evidence for yawn and self-scratch contagion in orangutans, and one other study has found mimicry of the play face in this species as well (Davila-Ross et al., 2008). Albeit limited, there is also evidence for facial mimicry in gorillas (Palagi et al., 2019b). Whether yawn or self-scratch contagion reflect true emotion contagion is still highly debated (see e.g., Gallup, 2021; Massen & Gallup, 2017; Palagi, Celeghin, et al., 2020), but given their presence in orangutans as well as evidence for facial mimicry in gorillas, it is likely that the last common ancestor of all great apes expressed similar behaviors in response to conspecifics.

Other important findings in this dissertation relate to the influence of social and contextual factors on emotion perception. Specifically, there are crucial differences in how these factors affect human and bonobo attention for emotions, how they modulate mimicry in orangutans. Attention for emotions in humans and bonobos is affected by *familiarity*, or more specifically, by whether the emotional expressions come from group mates or unfamiliar others (Chapter 2). As for orangutans, *familiarity* with individuals does not appear to impact the occurrence of contagious yawning (Chapter 5), but social closeness (or distance) does impact the occurrence of self-scratch contagion in stressful situations (Chapter 6). These findings highlight once again that context and social factors have an interactive effect on attention and mimicry. For instance, different evolutionary environments may have contributed to shaping the sensitivity to emotions of others, such that humans are more tuned to the ingroup, and bonobos more to the outgroup (Hrdy & Burkart, 2020).

Bonobos are remarkable in their xenophilic tendencies, which likely arose due to bonobos' relatively stable feeding environment (Hare et al., 2012). With less competition over valuable resources such as food, there was also less need to fight over these resources, which over time led to selection against aggressive tendencies. Crucially, this allows bonobos to have relatively peaceful interactions with strangers and even share food with them (Tan et al., 2017; Tan & Hare, 2013). In contrast, humans (and chimpanzees) had to adapt to more arduous environments and faced more severe competition over food and resources. Aggressive tendencies would therefore be beneficial for both species to protect the group (Bowles, 2009; Hare et al., 2012). Following this line of thought, as orangutans tend to affiliate with others less often compared to for instance bonobos, their perception of emotional expressions may also be less affected by socially close others, or only be affected by seeing unfamiliar or socially distant others. While this remains an open question for future research, I can at least say that our findings support the idea that self-scratch contagion in orangutans is affected by social distance, or specifically, having a weak social bond with the observed self-scratching individual(s).

In addition to an evolutionary account, contextual factors likely exert a great influence on emotion perception. For instance, recent findings show that bonobos and chimpanzees preferentially attend to high-ranking familiar individuals compared to unfamiliar individuals (for bonobos, these high-ranking familiar individuals were females, and for chimpanzees, males) (Lewis et al., 2021). Importantly, the researchers report differences not only between the species but also between the different populations that were tested. One population of chimpanzees lived with only one resident male whereas the other population had multiple males and females. The females from the first population showed a bias towards outgroup males rather than the familiar male, likely due to the difference in social environments. In the studies on emotion-biased attention in bonobos and humans, we aimed to provide more context to the emotional expressions on our stimuli by including whole-body expressions and by embedding expressions into *scenes*. Nevertheless, we did not yet look at how contextual factors that the participants find themselves in may affect their attentional bias. Future work should aim to consider these contextual factors more closely, and steps into this direction have already been made. See for instance work by Bethell et al. (2012) on how emotional states affect social attention in macaques. Moreover, we did consider contextual factors in our studies on mimicry with orangutans, and indeed see that for instance stress can affect the occurrence of self-scratch contagion in this species. As a comparison, a next step could be to look into how stress or tension affects the occurrence of self-scratch contagion (or other forms of mimicry) in bonobos and humans.

Finally, in addition to *familiarity* and *context* as moderators of emotion perception, I wish to highlight one more example, namely *similarity*. Specifically, similarity in facial or bodily expressions of emotions. The form and potentially also meaning of emotional expressions are, to some degree, shared between primates (e.g., Darwin, 1872; Preuschoft & van Hooff, 1995). Thus, due to a common heritage, it is conceivable that humans and great apes show an attentional bias towards some of the other species' emotional expressions, or that mimicry can occur between different species. For humans, we indeed found an attentional bias towards bonobo expressions in Chapters 3 and 4, but for bonobos, we could not confirm an attentional bias towards human expressions (Chapter 2 and 4). Given the evidence at hand, only humans may be sensitive to the feelings and needs of other species, with one explanation being that humans often ascribe human-like characteristics to other animals (Williams et al., 2020). However, it is more likely that our studies did not have enough power to detect an effect, given our small bonobo sample size. The evolutionary continuity between expressions of emotions in hominids remains an active topic of research (e.g., Waller et al., 2020), and future work could tackle the current knowledge gap by for instance measuring whether different species have similar implicit associations with emotional expressions of conspecifics and heterospecifics (e.g., humans or other closely related species), for which the PIAT (Chapter 7) could be an option. I discuss this in more detail in the final paragraphs of the next part.

Methodological considerations and future directions

Studying the emotional modulation of attention and spontaneous mimicry in bonobos, orangutans, and humans has revealed intriguing commonalities and differences between the species, and applying similar methods that make interspecies comparisons possible has proved to be a worthwhile approach. Future studies should consider improving the employed methods in several ways.

Firstly, although increasingly more effort is put into studying the perception of emotions in great apes and other animals, scientists are still only scratching the surface (see Kret et al. (2020) for a review). We based our stimulus selection in Chapters 2-4 on previous work that investigated how great apes perceive emotional expressions (e.g., De Waal, 1988; Kret et al., 2016; Parr et al., 1998, 2008), but several important candidates have yet to be studied. For instance, disgust is associated with a distinct, universal facial expression in humans that may have evolved as a response to harmful foods or other substances (Curtis et al., 2011). Moreover, disgust can also be used as an intentional signal to express strong disapproval of for instance immoral behavior (Chapman et al., 2009). There is some evidence that great apes show some features of the prototypical disgust expression, i.e., nose wrinkling and tongue protrusion (Case et al., 2020), but we still know surprisingly little about disgust in great apes and other primates. Similarly, anger is a core emotion in humans (Ekman, 1999), and the bulging lip face that for instance chimpanzees and bonobos produce may be a homologue of human anger (De Waal, 1988; Parr et al., 2007). In our studies, we lacked stimuli depicting anger and aggression or disgust, as they were very hard to come by, and therefore we could not measure to what extent these emotional cues modulate attention. Thus, future studies could include a wider range of emotional categories that also include more negatively-valenced emotional states such as anger and disgust. Moreover, we did not have the sample size to zoom in on specific emotion categories. Nevertheless, in humans, there are some mixed findings indicating either an attentional bias away or towards certain emotional categories (e.g., Pool et al., 2016; van Rooijen et al., 2017; Zvielli et al., 2014). Thus, future studies could focus on investigating how valence may impact attentional biases.

Secondly, we made use of emotional scenes that contain more contextual information compared to isolated facial expressions. Previous findings have shown that providing this context facilitates the recognition of emotions (De Gelder et al., 2010; Kret et al., 2013b). Although I did not explicitly test this, it is plausible that this is also true for apes' recognition of emotion. For instance, bonobos show a bared-teeth display when scared and nervous, but also during sex (De Waal, 1988). Providing contextual information by showing a scene rather than an isolated facial expression may thus facilitate the processing of its emotional content. Nevertheless, our stimuli were static (Chapters 2-4), and only contained social information coming from the visual modality (Chapters 2-4, 6). Although human facial expressions of emotions are highly ritualized and therefore salient (Kret et al., 2020), expressions of emotions are often multi-modal, consisting of vocalizations, gestures, and facial and bodily expressions. Therefore, the perception of emotional expressions may be enhanced when emotional information is coming from multiple channels (Paulmann & Pell, 2011). Indeed, human studies have shown that emotional information across different modalities is integrated, creating holistic, enhanced emotion recognition (De Gelder et al., 1999; De Gelder & Vroomen, 2000; Schirmer & Adolphs, 2017).

Similarly, there is some work suggesting that great apes use social information coming from different modalities to categorize expressions. For instance, Parr (2004) found that when chimpanzees categorized faces, they preferentially categorized pant-hoots and play faces based on their auditory components, and scream faces based on their visual components. From an evolutionary perspective, the results suggest that the auditory modality is more informative for pant-hoots and play faces because pant-hoots are used for long-distance communication, and play faces are often concealed during rough, close-contact play. Thus, vocalizations during play may be more salient than facial expressions for indicating playful intentions (Parr, 2004). Nevertheless, despite numerous studies in primates investigating expressions in a single domain (e.g., vocalizing, gesturing, and facial expressions), multi-modal signaling is virtually unexplored (Fröhlich et al., 2019b; Liebal et al., 2014). The call for a multi-modal approach is in line with recent work stressing the effects of different natural ecologies of animals (including humans) on the evolution and development of behavior and cognition (Bräuer et al., 2020). Moving forward, comparative studies could take a multi-componential approach, and for instance investigate how dynamic emotional scenes that also include auditory cues are viewed, what behaviors these dynamic scenes elicit, and how different modalities contribute to emotion perception.

In Chapters 5 and 6 I described our studies investigating yawn and self-scratch contagion in orangutans. Yawn contagion was virtually unexplored in this species, and only a limited amount of work has previously looked into self-scratch contagion. Nevertheless, a third consideration for future work has to do with the link between yawning and emotion contagion is currently weak (Massen & Gallup, 2017). Emotion contagion requires having an emotional experience, but it is unclear what the emotional state underlying yawning might be. Yawning has been linked to boredom (see a discussion in Burn, 2017) as well as stress (Maestripieri et al., 1992; Paukner & Anderson, 2006). Moreover, some researchers argue that yawning is entirely unrelated to emotions and may simply be a form of motor mimicry (i.e., devoid of any emotional content; Yoon & Tennie, 2010). To establish whether yawn contagion is a proxy for emotion contagion, future research should aim to measure which emotional state, then, is transferred (Adriaense et al., 2020). This is of course not an easy feat as animals cannot report on their feelings, but one way to move forward is to study more closely the (social) contexts in which yawn contagion (or other examples of mimicry) occurs to determine whether it has a communicative purpose, and what that communicative purpose is.

A final methodological consideration concerns the use of the Pictorial Implicit Association Test (PIAT) for comparative research (Chapter 7). What research questions could comparative scientists answer using the PIAT, and which type of stimuli could be used? IATs are widely used in social cognitive research, for instance to study implicit associations with ethnicity or gender (e.g., Baron & Banaji, 2006; Nosek et al., 2002). Nevertheless, IATs could potentially help uncover the implicit associations underlying emotion perception in great apes.

Often, we have to make assumptions about valence and arousal of emotional signals in animals based on the contexts in which emotions are expressed and how individuals respond to them (Kret et al., 2020). However, we currently do not yet fully understand whether great apes view certain emotional expressions or scenes as positive or negative, or how they are perceived in terms of arousal. Moreover, human studies on emotion perception suffer from the same interpretive issues as animal studies, as human participants are often directly asked (e.g., through standardized questionnaires) about how they interpret emotional expressions or how they experience them. Notwithstanding the strong psychometric properties of some questionnaires, how individuals answer questions can be confounded by for instance the tendency to give desirable answers or their ability to self-reflect and articulate their emotions (Stone et al., 2000). For instance, individuals from clinical populations (e.g., autism-spectrum disorders) may find it challenging to report felt emotions or interpret them (Cook et al., 2013). Probing of implicit associations with emotions may therefore offer a less-biased way to measure how emotions are perceived, and also allows for direct comparisons between different populations and species. As such, the PIAT could potentially measure the strength of associations between emotional scenes and their valence or arousal level.

Indeed, the IAT has been successfully used to establish for instance cultural differences in implicit attitudes towards emotion regulation (Deng et al., 2019), approach and avoidance of fearful expressions (Hammer & Marsh, 2015), attitudes towards emotional expressivity (Cameron & Zeman, 2019), and stereotypical associations between certain emotional expressions (e.g., anger) and ethnicity (Bijlstra et al., 2014). Recently, one study directly investigated implicit attitudes towards emotions in adolescence with and without high-functioning autism spectrum disorder (ASD). Emotional stimuli consisted of the six basic emotion expressions (i.e., happiness, sadness, disgust, fear, anger, and surprise) and neutral expressions. The emotional expressions also varied in intensity. In this IAT, emotional stimuli were paired with words that were positive (e.g., peace, success) or negative (e.g., war, failure). Results indicated that individuals with high-functioning ASD showed larger IAT-effects (i.e., stronger positive or negative associations with the different emotional images) than individuals without ASD, indicating that it is possible to probe implicit attitudes towards emotions in humans. For a comparative study involving e.g., great apes, the words would have to be replaced by images of objects or individuals that great apes would associate with positivity or negativity, e.g., a veterinarian or a tranquillizer gun (see e.g., Allritz et al. (2015)).

Conclusion

In this dissertation, I set out to probe the emotional landscape of our closest relatives, the great apes. This type of research can help us progress our understanding of the evolution of emotional capacities in our species as well as other animals. Unfortunately, the population of all wild-living great apes is dwindling, and all great ape species now face extinction. As a consequence, the unique window into our evolutionary past is slowly but surely disappearing. Motivated by this pressing issue, I investigated the similarities and differences in emotion perception in bonobos, orangutans, and humans. Specifically, I investigated attentional and behavioral mechanisms that underlie emotion perception, which has been suggested to lay at the foundation of higher cognitive processes such as empathy (e.g., De Waal & Preston, 2017).

The results in this dissertation indicate a shared evolutionary origin for emotion-biased attention and spontaneous mimicry in *at least* the hominids. Importantly, familiarity and similarity may affect how attentional and behavioral mechanisms drive emotion perception, and these “modulators” can themselves be affected by species-specific characteristics such as sociality. The work I describe in my dissertation has raised new, important questions that need addressing in future research. At this moment, we do not fully understand which emotional expressions are the most salient or relevant to specific species (Kret et al., 2020). Furthermore, to elucidate a link between yawn and self-scratch contagion and emotion contagion, we require more work involving different animal species and more direct ways of testing the link (Massen & Gallup, 2017). Finally, establishing how great apes view emotions in terms of valence and arousal is important not only for our fundamental understanding of emotion perception but also for questions relating to animal welfare (Adriaense et al., 2020). Thus, I hope that this dissertation provides a stepping stone towards more research on emotions in our closest living relatives, but also other animals.

Appendices

Appendix A

Supplementary Materials for Chapter 2

Experiment 1: Bonobos' attentional bias towards emotions of familiar and unfamiliar conspecifics.

Table S1. Definition of emotion categories used for bonobos in Experiment 1, and number of pictures per emotion category and per familiarity category.

Picture category	Description	No. of individuals per picture	No. of unfamiliar pictures	No. of familiar pictures
Distress	Aggressive displays (e.g., charges and direct displays); submissive behaviors (e.g., grin faces, fleeing and crouching)	$M = 1.55$, $SE = 0.16$	15	25
Sex	Mating, genito-genital rubbing, prominent full swelling, penile erection	$M = 2.25$, $SE = 0.25$	27	16
Playing	Together or alone, with a relaxed open mouth, without an object	$M = 2.25$, $SE = 0.25$	22	36
Grooming	Two or more individuals grooming	$M = 2.38$, $SE = 0.26$	53	50
Yawning	Wide-open mouth with or without teeth exposure	$M = 1.17$, $SE = 0.17$	20	16
Self-scratching	Roughly rubbing the body, face, or one of the limbs	$M = 1.00$, $SE = 0.00$	18	30
Neutral	Walking, lying down, or sitting	$M = 1.67$, $SE = 0.14$	155	173

Note. All bonobo participants saw the same number of unique pictures, but for the familiar models the composition of the stimulus set differed per participant because we replaced pictures of the participants themselves with pictures of other familiar individuals. Furthermore, the number of pictures differs per *Familiarity* and also per *Picture category* because some behaviours were easier to photograph or occurred more frequently than other behaviours.

Appendix to stimuli and validation

Six primate experts scored the pictures of bonobos based on valence and intensity. Three experts worked with the bonobos on a daily basis, and the others worked with bonobos or chimpanzees in the past. Experts were presented with one picture at a time and were asked to 1) rate how negative or positive they thought bonobos would experience each picture (1= very negative, 7 = very positive) and 2) how intense the picture was (1= not intense, 7 = very intense). Pictures were shown until a response was given. We calculated intraclass correlations (ICCs) for valence and intensity ratings using a two-way mixed model and a consistency definition and found a high reliability for both ratings ($ICC_{valence} = .82$, 95% CI: .79 - .84, $F(653, 3265) = 5.45$, $p < .001$; $ICC_{intensity} = .87$, 95% CI: .86-.89, $F(653, 3265) = 7.96$, $p < .001$). A generalized linear mixed model with *Emotionality* (emotion versus neutral) as fixed factor, *Rating* as target variable and *Rater Number* as random effect confirmed that emotional pictures are indeed rated higher in intensity than neutral pictures ($M_{\text{emotional}} = 3.28$, $SE = 0.35$, $M_{\text{neutral}} = 1.76$, $SE = 0.18$, $F(1, 3922) = 1337.81$, $p < .001$).

Table S2. Overview of average intensity and valence ratings per emotion category and per familiarity category in Experiment 1.

	Familiar		Unfamiliar	
Emotion category	Average Valence (SD)	Average Intensity (SD)	Average Valence (SD)	Average Intensity (SD)
Distress	3.07 (1.38)	4.73 (1.57)	3.49 (1.55)	4.74 (1.39)
Sex	5.93 (0.85)	4.90 (1.28)	5.88 (0.93)	4.94 (1.34)
Play	5.31 (0.95)	3.12 (1.58)	5.44 (0.92)	2.89 (1.44)
Groom	5.31 (0.82)	3.09 (1.64)	5.15 (0.83)	2.73 (1.57)
Yawn	3.96 (0.92)	3.31 (1.56)	3.97 (1.00)	3.53 (1.62)
Self-scratch	4.11 (1.07)	1.76 (0.71)	4.14 (1.02)	1.74 (0.78)
Neutral	4.63 (0.91)	1.64 (0.89)	4.62 (0.88)	1.74 (0.98)

Table S3. Number of incorrect trials per bonobo in Experiment 1.

Participant name	Tested trials (sessions)	Repetitions of trials*	Incorrect trials (% of grand total) †	Of which due to nose wipes	Of which due to self-scratching	Of which due to outliers‡
Besede	525 (21)	202	130 (6.0 %)	2	1	127
Kumbuka	582 (24)	264	181 (8.4%)	38	2	141
Monyama	537 (22)	212	104 (4.8%)	2	0	102
Yahimba	518 (21)	210	99 (4.6%)	4	0	95
Grand total	2162	888	514 (23.8 %)	46	3	465

* These reflect the number of trials that were repeated due to disruptions in the first/original trials.

† The number of incorrect trials consists of both erroneous trials within the first/original trials and erroneous trials within the repetitions.

‡ Outliers were disruptions during a trial other than due to the behaviours described above, e.g.: not attending to the screen during stimulus presentation, someone other than the participant pressing the probe, using the opposing hand, not sitting directly in front of the screen, or the screen not immediately responding to a touch. Outliers also contain extreme RTs (250 < RT < 5000) and RTs higher than the median RT per participant minus 2.5*MAD per participant.

Appendix to results

Table S4.1. Individual and overall reaction time means and standard deviations per condition in Experiment 1

Individuals	Familiar		Unfamiliar	
	Congruent *M (SD)*	Incongruent *M (SD)*	Congruent *M (SD)*	Incongruent *M (SD)*
Besede	699.73 (95.78)	702.63 (101.11)	697.43 (88.19)	714.96 (87.84)
Kumbuka	534.77 (64.87)	535.84 (69.44)	518.17 (75.58)	532.12 (68.29)
Monyama	460.44 (75.00)	449.59 (71.63)	456.89 (76.93)	473.03 (58.25)
Yahimba	431.61 (66.60)	421.19 (62.18)	422.49 (66.89)	423.12 (71.12)

Table S4.2. Model output from Experiment 1.

Predictors	β	*SE*	*t*	*p*
(Intercept)	532.16	55.31	9.62	.001***
Congruency (*congruent*)	-1.47	1.85	-.80	.426
Familiarity (*familiar*)	-.11	1.85	.06	.952
Congruency* Familiarity	3.77	1.85	2.04	.042*
Random effects	*Variance*	*SD*		
Session*ID (intercept)	264.2	16.25		
ID (intercept)	12090.4	109.95		
Residual	5597.4	74.82		

Number of observations: 1648, Session*ID: 8, ID:4. Note that by sum coding, the levels of all factors have been mean centred. As such, the intercept reflects the grand mean of all predictors.

Experiment 2: Bonobos' attentional bias towards emotions of familiar and unfamiliar humans.

Table S5.1. Overview of average agreement on emotionality (emotional/neutral), average rating scores for emotional intensity and authenticity rated by N=5 research assistants in Experiment 2.

	Caretaker (familiar)			NimStim (unfamiliar)		
Stimulus category	Agreement on emotionality (SD)	Intensity rating (SD)	Authenticity rating (SD)	Agreement on emotionality (SD)	Intensity rating (SD)	Authenticity rating (SD)
Emotional	92.8% (.26)	4.21 (1.36)	4.37 (1.60)	88.9% (.32)	5.09 (1.34)	3.66 (1.56)
Neutral	77.3% (.42)	2.84 (1.67)	5.04 (1.32)	100% (.00)	3.18 (1.58)	4.80 (1.39)

Note. Agreement on emotion refers to the average agreement between 5 raters on whether emotional stimuli were recognized as an emotion by them, and neutral stimuli as neutral. Agreement on whether a stimulus was emotional or neutral was not 100% on average, meaning raters sometimes rated a neutral expression as emotional and vice versa. Furthermore, low intensity scores for neutral stimuli and high intensity scores for emotional stimuli are preferred, because this means there is a clear discrepancy between two simultaneously presented stimuli and as such, highly intense (emotional) stimuli can capture attention faster than low-intensity (neutral) stimuli. A generalized linear mixed model with *Emotionality* (emotion versus neutral) as fixed factor, *Rating* as target variable and *Rater Number* as random effect confirmed that emotional pictures are indeed rated higher in intensity than neutral pictures ($M_{\text{emotional}} = 4.70$, $SE = 0.39$, $M_{\text{neutral}} = 2.79$, $SE = 0.24$, $F(1, 603) = 233.19$, $p < .001$).

Table S5.2. ICCs of scores on the different scales (intensity, emotionality (emotion/neutral), and authenticity using two-way mixed effects using a consistency definition in Experiment 2 (N=6).

Scale	Intraclass Correlation	95% confidence interval		*F* value (df1, df2)	*P* value
Intensity	.78	.71	.84	4.53 (120, 480)	.000
Emotion	.66	.55	.75	2.94 (120, 480)	.000
Authenticity	.69	.59	.77	3.20 (120, 480)	.000

Table S6. Number of incorrect trials per bonobo on the dot-probe with human stimuli Experiment 2.

Participant name	Tested trials (sessions)	Repetitions of trials*	Incorrect trials (% of grand total) †	Of which due to nose wipes	Of which due to self-scratching	Of which due to outliers‡
Besede	325 (13)	198	66 (20.3)	0	0	58
Kumbuka	326 (13)	198	106 (32.5)	15	0	73
Monyama	377 (15)	245	78 (20.7)	0	0	48
Yahimba	350 (14)	223	123 (35.1)	1	0	58
Grand total	1378	864	373	16	0	237

* These reflect the number of trials that were repeated due to disruptions in the first/original trials.

† The number of incorrect trials consists of both erroneous trials within the first/original trials and erroneous trials within the repetitions.

‡ Outliers were disruptions during a trial other than due to the behaviours described above, e.g.: not attending the screen during stimulus presentation, someone other than the participant pressing the probe, using the opposing hand, not sitting directly in front of the screen, or the screen not immediately responding to a touch. Outliers also contain extreme RTs (250 < RT < 5000) and RTs higher than the median RT per participant minus 2.5*MAD per participant.

Appendix to results

Table S7.1. Individual and overall reaction time means and standard deviations per condition

Individuals	Familiar (Caretaker)		Unfamiliar (NimStim)	
	Congruent *M (SD)*	Incongruent *M (SD)*	Congruent *M (SD)*	Incongruent *M (SD)*
Besede	628.21 (125.55)	631.84 (133.57)	630.15 (113.73)	635.15 (113.28)
Kumbuka	526.69 (52.74)	530.69 (49.84)	525.93 (61.46)	526.43 (52.41)
Monyama	429.28 (62.39)	430.03 (69.86)	437.04 (73.85)	438.94 (66.04)
Yahimba	399.79 (56.26)	400.79 (76.19)	379.63 (84.67)	394.47 (76.73)

Table S7.2. Model output from Experiment 2.

Predictors	β	*SE*	*t*	*P*
(Intercept)	491.33	41.69	11.79	.001***
Congruency (*congruent)*	-1.44	2.51	-.58	.565
Familiarity (*unfamiliar/familiar*)	-.50	2.50	-.20	.842
Congruency* Familiarity	.99	2.50	.40	.691
Random effects	*Variance*	*SD*		
Session*ID (intercept)	1818	42.64		
ID (intercept)	6315	79.46		
Residual	6275	79.21		

Number of observations: 1005, Session*ID: 12, ID:4. Note that by sum coding, the levels of all factors have been mean centred. As such, the intercept reflects the grand mean of all predictors.

Table S8. Model output from the control experiment as part of Experiment 2.

Predictors	β	*SE*	*t*	*p*
(Intercept)	6.08	.009	665.1	< .001***
Congruency (*congruent)*	-.003	.002	-2.00	.046*
Familiarity (*unfamiliar*)	.000	.002	-.10	.931
Congruency* Familiarity	-.001	.002	.20	.867
Random effects	*Variance*	*SD*		
ID (intercept)	.00	.06		
Residual	.03	.16		

Number of observations: 9788, ID: 150. Note that results are given on the log (not the response) scale. For clarity, 3 decimals are reported for the beta estimate and standard error. By sum coding, the levels of all factors have been mean centred. As such, the intercept reflects the grand mean of all predictors.

Experiment 3: Humans' attentional bias towards emotions of familiar and unfamiliar conspecifics.

Appendix to participants

Table S9. Descriptives of participants and their relation to the participant on the stimuli in Experiment 3.

Relationship participant on task versus participant on photo*	Sex participant on task versus Sex participant on photos				
	Male versus Male	Male versus Female	Female versus Female	Female versus Male	Grand total
Brother/sister	25	25	32	19	101
Child	25	16	33	23	97
Parent	22	23	25	28	98
Spouse/partner	1	27	0	48	76
Niece/nephew[†]	0	1	1	0	2
Friend/Colleague	8	3	12	7	30
Grand total	81	95	103	125	404[‡]

* The relationship is seen from the viewpoint of the participant doing the task, e.g., "Child" means that the stimuli are of the child of the participant who performed the dot probe task.

[†] As we were interested in how closely bonded individuals attend to each other's emotions, we focused mainly on families and friends. We did not collect a lot of participants with a more distant family relationship (e.g., aunts/uncles, nephews/nieces, cousins), but decided not to remove the 2 participants with a niece or nephew.

[‡] Note that this number is not the same as the one reported in the main text (N=449), this is because of a technical failure, the relationship data of 45 participants was not registered.

Appendix to stimulus validation

Table S10.1. Overview of average agreement on emotionality (emotion/neutral), and average rating scores for emotional intensity and authenticity of human stimuli in Experiment 2.

Rater group	Neutral stimuli			Emotional stimuli		
	Agreement on emotionality (SD)	Intensity rating (SD)	Authenticity rating (SD)	Agreement on emotionality (SD)	Intensity rating (SD)	Authenticity rating (SD)
1	82.9% (.38)	3.16 (1.85)	5.72 (2.41)	68.3% (.47)	4.08 (1.86)	4.85 (1.87)
2	74.8% (.43)	2.17 (1.64)	5.79 (1.42)	88.3% (.32)	4.97 (1.79)	5.39 (1.65)
3	77.3% (.42)	2.69 (1.3)	5.26 (1.66)	90.5% (.29)	4.44 (1.59)	4.06 (1.88)
Grand average	81.5% (.39)	2.99 (1.83)	5.71 (1.29)	70.9% (.45)	4.18 (1.71)	5.29 (1.43)

Note. Rater group 1, N = 8; Group 2, N = 5; Group 3, N = 5. Agreement on emotion/neutral refers to the average agreement between raters on whether emotional stimuli were recognized as an emotion by them, and neutral stimuli as neutral. Agreement was not 100% on average, meaning raters sometimes rated an emotional expression as neutral and vice versa. Intensity and authenticity were rated on a scale from 1-7 (1 = low intensity/ authenticity, 7 = high intensity/authenticity). Furthermore, low intensity scores for neutral stimuli and high intensity scores for emotional stimuli are preferred, because this means there is a clear discrepancy between two simultaneously presented stimuli and as such, highly intense (emotional) stimuli can capture attention faster than low-intensity (neutral) stimuli. A generalized linear mixed model with *Emotionality* (emotion versus neutral) as fixed factor, *Rating* as target variable and *Rater Group*Rater Number* as random effect confirmed that emotional pictures are indeed rated higher in intensity than neutral pictures ($M_{emotional} = 4.17$, $SE = 0.22$, $M_{neutral} = 2.74$, $SE = 0.14$, $F(1, 30377) = 5139.47$, $p < .001$).

Table S10.2. ICC (two-way mixed, consistency) on intensity scores, emotion type, and authenticity scores per rater group in Experiment 2.

	Rater group	Intraclass Correlation	95% confidence interval		F value (df1, df2)	*p* value
Intensity scores	1	.74	.73	.76	3.90 (3078, 21546)	.000
	2	.89	.88	.91	9.34 (568, 2748)	.000
	3	.78	.74	.82	4.63 (263, 1052)	.000
Emotion type scores	1	.92	.91	.92	12.07 (3219, 22533)	.000
	2	.69	.65	.72	3.18 (687, 2748)	.000
	3	.82	.78	.85	5.49 (263, 1052)	.000
Authenticity	1	.70	.68	.71	3.29 (3078, 21546)	.000
	2	.60	.55	.64	2.47 (687, 2748)	.000
	3	.74	.69	.79	3.87 (263, 1052)	.000

Appendix to results

Table S11.1. Model output from Experiment 3.

Predictors	β	*SE*	*t*	*p*
(Intercept)	6.334	.007	934.0	< .001***
Congruency (*incongruent)*	-.002	.001	-1.88	.061
Familiarity (*unfamiliar*)	-.001	.001	-.78	.445
Congruency* Familiarity	.002	.001	-2.00	.047*
Random effects	*Variance*	*SD*		
ID (intercept)	.020	.141		

Number of observations: 16949, ID: 444. Note that results are given on the log (not the response) scale. For clarity, 3 decimals are reported for the estimate, standard error, variance, and standard deviation.

Table S11.2. Exploratory analysis: Testing for effects of sex-combinations (dot-probe participant and photo participant), familiarity and congruency on reaction times in Experiment 3.

Predictors	χ^2	*df*	*p*
Sex combination (*M-M, M-F, F-M, F-F*)	1.46	3	.691
Congruency (*congruent, incongruent)*	1.46	1	.227
Familiarity (*familiar, unfamiliar*)	.61	1	.434
Sex combination*Congruency*Familiarity	1.44	3	.696
Random effects	*Variance*	*SD*	
ID (intercept)	.00	.07	
Residual	.02	.16	

Number of observations: 15326, ID: 402. Note that results are given on the log (not the response) scale, and that some samples are missing because we lost some data relating to sex and relationship. M = male, F = female.

Appendix B

Supplementary Materials for Chapter 4

Appendix to stimuli

Table S1. Definitions of emotion categories of bonobo and human stimuli.

Picture category	bonobo	human
Distress	Aggressive displays, such as long and short charges, mutual and direct displays; submissive behaviours such as grin faces, fleeing, and crouching	Fearful displays (wide-open eyes, raised eye brows, open mouth) which may include crying and embracing
Sex	Mating, genito-genital rubbing, masturbation, prominent full swelling, penile erection	Man and woman in underwear romantically embracing and/or kissing, no genitals or penetration visible
Playing	Together or alone, with a relaxed open mouth, without an object	See bonobo
Grooming	Two or more individuals grooming with close physical contact	Two or more individuals in close physical contact, hugging, smiling, or brushing hair
Yawning	Wide open mouth with or without canine visibility	Wide open mouth with or without teeth visible
Neutral	Walking, lying down, or sitting	Walking, lying down, sitting, running, cycling

Table S2. Overview of number of individuals, adults, juveniles, males and females per picture category for bonobo stimuli

Picture category	No. of individuals		No. of adults		No. of juveniles / infants		No. of females*		No. of males*	
	M	SD	M	SD	M	SD	M	SD	M	SD
Distress	1.22	0.44	0.67	0.71	0.56	0.53	0.78	0.44	0.44	0.53
Sex	1.80	1.23	1.50	0.97	0.30	0.67	1.20	1.03	0.60	0.52
Play	1.88	0.99	0.63	0.74	1.25	0.71	1.00	0.93	0.88	0.83
Groom	2.50	0.85	2.10	0.74	0.40	0.52	1.80	0.79	0.70	0.95
Yawn	1.00	0.00	0.91	0.30	0.09	0.30	0.18	0.40	0.82	0.40
Neutral	1.75	1.21	1.23	0.81	0.52	0.90	0.98	0.84	0.77	0.86

* Since stimuli were collected from the internet, it was not always possible to deduct whether an individual was male or female (e.g., when only part of the body was visible).

Table S3. Overview of number of individuals, adults, juveniles, males and females per picture category for human stimuli.

Picture category	No. of individuals		No. of adults		No. of juveniles / infants		No. of females		No. of males	
	M	SD	M	SD	M	SD	M	SD	M	SD
Distress	6.00	3.77	4.90	4.01	1.10	3.14	4.00	3.83	3.70	3.71
Sex	2.00	0.00	2.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00
Play	2.80	1.81	1.90	2.42	0.90	1.10	1.80	1.93	1.00	0.67
Groom	3.60	2.76	3.20	3.01	0.50	0.71	2.00	1.94	1.80	1.87
Yawn	1.50	1.27	0.90	1.52	0.60	0.52	0.90	0.99	0.60	0.70
Neutral	5.37	3.68	5.26	3.72	0.29	0.64	3.61	3.52	3.80	3.68

Appendix to calibration of bonobos

During data collection, we displayed a 9-point grid to the bonobos at the beginning of each session to visually inspect calibration accuracy through Tobii Live Viewer (Figure S1). Nevertheless, after data collection ended and checking the raw data, we noticed that for two bonobos (Zuani and Monyama), there were consistent shifts in the gaze data either to the left or right relative to the position of the stimuli on the screen. To make sure our measure of interest, *Total fixation duration*, was as accurate as possible, we decided to correct the specific cases in which we could see a shift. We therefore developed a python script to establish the gaze offsets per session for Zuani and Monyama. Per session, we used K-means clustering to calculate the two centroids (i.e., the mean position of all points within a 2-dimensional space) of the gaze data on the left and right stimulus, and compared the X and Y coordinates of these centroids to the true center points of the left and right stimuli on the screen. The average offset to the left or right on the X-axis relative to the true center points was then used to calculate the adjusted regions of interest for a specific session. In Figure S2 down below you can see an example of this. For the data analyses, we did not manipulate the raw gaze data, but rather adjusted the regions of interest in Tobii Studio to accommodate the offset (see Figure S3).

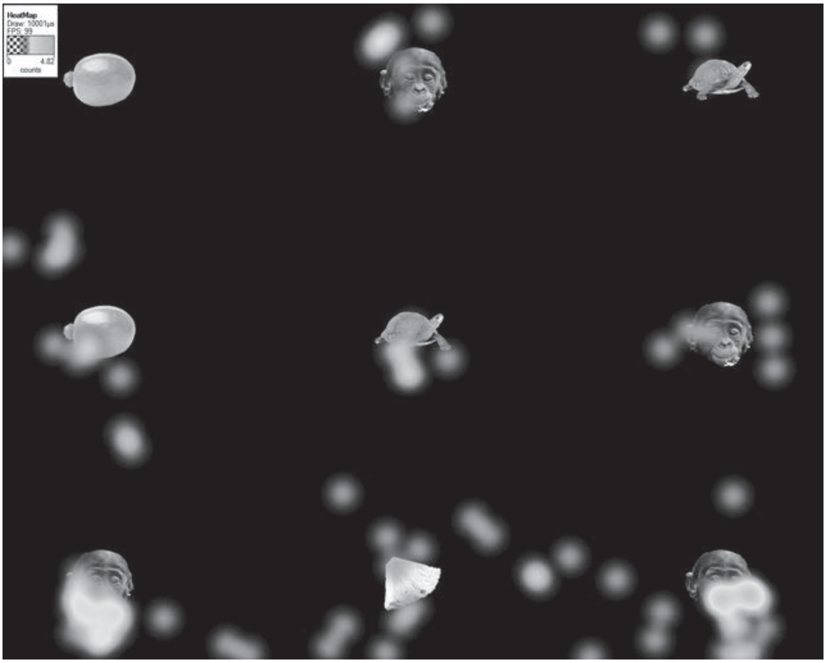

Figure S1. Example of the 9-point grid shown to the bonobos at the start of each session, here displayed with a heatmap (with green indicating lower, and red indicating higher fixation counts).

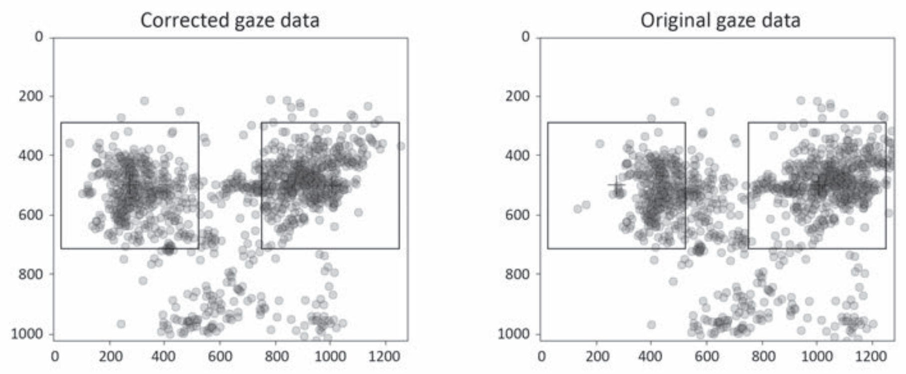

Figure S2. Example of a session in which gaze data was consistently shifted to the left relative to the positions of the stimuli (blue and orange squares) on the screen. On the right, the original gaze data of a specific session of one individual (Monyama) is plotted. For clarity, we also plotted what the corrected gaze data would look like applying K-means clustering.

Figure S3. Example of what (corrected) region of interests (ROI) could look like. On the left, the normal ROIs are displayed. On the right, the corrected ROIs for the specific session displayed in S1 are displayed.

Appendix to results experiment 1

Table S4. Overview of results of the bonobos.

	Model 1	**Model 2**	**Model 3**
Predictor	β *(89% CI)*	β *(89% CI)*	β *(89% CI)*
Intercept	-0.01 [-0.15 – 0.13]	0.06 [-0.08 – 0.20]	0.18 [0.02 – 0.33]
Species (human)		-0.16 [-0.23 – -0.09]	-0.08 [-0.23 – 0.08]
Emotion category (groom)			-0.21 [-0.35 – -0.06]
Emotion category (play)			-0.20 [-0.35 – -0.05]
Emotion category (sex)			-0.04 [-0.17 – 0.11]
Emotion category (yawn)			-0.17 [-0.32 – -0.04]
Species (human) * Emotion category (groom)			0.05 [-0.17 – 0.26]
Species (human) * Emotion category (play)			0.11 [-0.10 – 0.32]
Species (human) * Emotion category (sex)			-0.49 [-0.70 - -0.28]
Species (human) * Emotion category (yawn)			-0.06 [-0.27 – 0.15]
Random effects			
σ^2	1	1	1
τ_{00}	0.03_{ID}	0.02_{ID}	0.02_{ID}
	$0.00_{ID:Session}$	$0.00_{ID:Session}$	$0.00_{ID:Session}$
N	$134_{Session}$	$134_{Session}$	$134_{Session}$
	4_{ID}	4_{ID}	4_{ID}
Observations	1420	1420	1420

Note: All values are reported on the log scale.

Appendix to results experiment 2

Table S5. Overview of results per model, per factor.

	Model 1	Model 2	Model 3
Predictor	β *(89% CI)*	β *(89% CI)*	β *(89% CI)*
Intercept	0.13 [0.10 – 0.16]	0.06 [0.02 – 0.10]	-0.08 [-0.15 – -0.01]
Species (human)		0.13 [0.09 – 0.18]	0.32 [0.22 – 0.41]
Emotion category (groom)			0.24 [0.15 – 0.34]
Emotion category (play)			0.22 [0.13 – 0.32]
Emotion category (sex)			0.12 [0.03 – 0.22]
Emotion category (yawn)			0.11 [0.02 – 0.20]
Species (human) * Emotion category (groom)			-0.26 [-0.39 – -0.13]
Species (human) * Emotion category (play)			-0.25 [-0.39 – -0.12]
Species (human) * Emotion category (sex)			-0.13 [-0.27 – 0.00]
Species (human) * Emotion category (yawn)			-0.30 [-0.43 – -0.17]
Random effects			
σ^2	1	1	1
τ_{00}	0.02_{ID}	0.02_{ID}	0.02_{ID}
N	94_{ID}	94_{ID}	94_{ID}
Observations	2780	2780	2780

Note: All values are reported on the log scale.

Appendix C

Supplementary Materials for Chapter 5

Appendix to participants and data collection

Table S1. Additional information about the study subjects

Name	Sex	Birth year	Developmental stage	Minutes of observation	Relationship
Amos	Male	2000	Adult	1120	Father of Kawan and Baju
Baju	Male	2015	Juvenile	NA	Son of Amos and Wattana
Binti	Female	2000	Adult	1090	No kin in the group
Dayang	Female	2005	Adult	1120	Adopted by Sandy
Kawan	Male	2010	Adolescent (unflanged)	1090	Son of Wattana
Kevin	Male	1982	Adult	1110	Born in the wild, no kin in the group
Samboja	Female	2005	Adult	1080	Daughter of Sandy
Sandy	Female	1982	Adult	1110	Mother of Samboja
Silvia	Female	1965	Adult	1080	No kin in the group
Wattana	Female	1995	Adult	1110	Mother of Kawan and Baju

Table S2. Ethogram of the orangutans

	Behavior	Code	Description
Infant	Nursing	N	Infant suckling from nipple unaided (dur)
	Retrieve	R	Retrieve infant from another animal or physical structure (dur)
Socio-positive	Allogrooming	GR	One individual picking, stroking or parting of hair over any part of the body with mouth or hands of another individual (dur)
	Being groomed	BG	Focal animal is groomed by another animal (dur)
	Contact	C	Mouth-mouth contact, olfactory inspection, touch (dur)
	Contact sitting	CS	Sitting next to other individual (dur)
	Play	P	Mouth fighting or wrestling (dur)
Agonistic	Chase other	CH	Individual pursuing another using any form of locomotion (dur)
	Direct aggression	FI	Biting/hitting/grabbing (dur)
	Display	DI	Charge, shaking of climbing structure/rope (dur)
	Make way	MW	Move out of the way when another animal is approaching (pt)
Sexual	Genital contact	GC	One individual touching hand/mouth to another's genital area (dur)
	Mount	M	Mounting another animal in a copulatory position; genital-genital contact established (dur)
Food associated	Drinking	DR	Drink from drink-nipple (pt)
	Feeding	FE	Actively eating, reaching for food, processing or preparing food items (dur)
	Foraging	FO	Searching for food (dur)
	Give food	GF	Give food to another animal (pt)
	Take food	TF	Take food from another animal's mouth or hand (pt)
Locomotion	Brachiate/climbing	CL	Hand-over-hand locomotion (dur)
	Walk	W	Forward or backward locomotion either quadrupedally or bipedally (dur)
Facial expression	Funnel face	FF	Maximal pursing of the lips (pt)
	Grimace	GR	Teeth showing, mouth slightly open, corners pulled back (pt)
Vocalization	Grunt	G	Deep, belch-like vocalization (dur)
	Kiss squeak	KS	Vocalization made by the intake of air through extended lips (dur)
	Long call	LC	Deep, rumbling vocalization (dur)
Other	Auto-groom	AG	Picking, stroking or parting of own hair over any part of the body (dur)
	Auto-play	AP	Animal plays with food/items alone (dur)
	Caretaker	CA	Interacting/waiting for interactions with the caretaker (dur)
	Nest building	NB	Preparation of day/night nest (dur)
	Object manipulation	OM	Manipulating object with hand/mouth (dur)
	Out of sight	OS	Animal is out of sight (dur)
	Resting	R	Laying (in a nest), not sleeping (eyes opened) (dur)
	Self-scratching	SC	Auto-scratch (dur)
	Sleeping	SL	Resting without locomotion, eyes are closed (dur)
	Yawn	Y	Opening of mouth and lips with teeth bared or not visible (pt)

Dur = duration behavior, pt = point behavior

Appendix to results

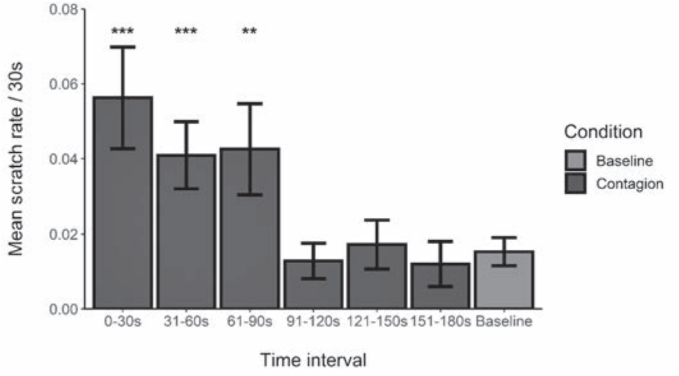

Figure S1. Mean self-scratch rates (± *SEM*) in the six 30 second intervals in the contagious condition compared to the baseline self-scratch rate. *** $p < 0.001$, ** $p < 0.01$.

Appendix D

Supplementary Materials for Chapter 6

Appendix to participants and procedure

Table S1. Overview of test subjects, their sex, age, and relationship.

Name	Sex	Birth year	Developmental stage	Relationship
Amos	Male	2000	Adult	Father of Kawan, Baju, and Indah
Baju	Male	2015	Juvenile	Son of Amos and Wattana
Indah	Female	2017	Infant	Daughter of Amos and Samboja, granddaughter of Sandy
Kawan	Male	2010	Adolescent (unflanged)	Son of Amos and Wattana
Kevin	Male	1982	Adult	Born in the wild, no kin in group
Sandy	Female	1982	Adult	Mother of Samboja, grandmother of Indah
Samboja	Female	2005	Adult	Daughter of Sandy, mother of Indah
Wattana	Female	1995	Adult	Mother of Kawan and Baju

Note: Two adult females (Wattana and Samboja) were always housed with their offspring and sometimes with one adult male.

Test sessions were carried out in mornings and afternoons (between 10:00-15:30) in the inside enclosures (Figure 1 in Chapter 6). During testing, two observers and occasionally a zookeeper were presen. Sometimes, volunteer guides belonging to the park visited the orangutans during testing. Furthermore, nearing the end of the testing period one of the individuals had fallen ill and therefore two caretakers were present during testing almost daily. At the beginning of every session, the screen was placed in front of one of the enclosures. An enclosure usually housed one to four individuals (often mother-offspring and one adult male), and testing only started when the focal individual was in a position that allowed it to see the screen. Furthermore, if a focal individual did not see at least one full clip (yawn or control), another attempt was made by moving the screen or waiting until the focal was in a suitable position. All individuals in the enclosure were filmed with two cameras. If the focal individual moved, the observers adjusted the position of the screen as well.

Yawns that occurred during the primer or while it was clear that the subject had no direct line of sight towards the screen when the stimuli were presented, were recorded as spontaneous yawns and not taken into account in subsequent analyses. Note however, that the number of spontaneous yawns was almost exactly equal between the two conditions (18 and 17 in the yawn and control condition, respectively).

Appendix to results

Table S2.1. Overview of yawning *occurrences* per individual in yawn and control conditions and across triggers.

Trigger	Familiar		Unfamiliar		Avatar		
Individual	Yawns in Control	Yawns in Yawn	Yawns in Control	Yawns in Yawn	Yawns in Control	Yawns in Yawn	Total
Amos	2	6	0	1	2	1	12
Baju	1	2	0	0	1	2	6
Indah	0	0	0	1	0	0	1
Kawan	2	2	6	6	1	1	18
Kevin	0	2	0	1	0	0	3
Samboja	0	0	0	0	0	0	0
Sandy	0	1	0	1	0	0	2
Wattana	0	1	0	1	0	1	3
Total	5	14	6	11	4	5	45

Table S2.2. Overview of yawning *rates* per individual in yawn and control conditions and across triggers.

Trigger	Familiar		Unfamiliar		Avatar		
Individual	Yawns in Control	Yawns in Yawn	Yawns in Control	Yawns in Yawn	Yawns in Control	Yawns in Yawn	Total
Amos	2	8	0	1	2	1	14
Baju	3	8	0	0	2	4	17
Indah	0	0	0	1	0	0	1
Kawan	2	4	9	21	1	1	38
Kevin	0	2	0	3	0	0	5
Samboja	0	0	0	0	0	0	0
Sandy	0	4	0	1	0	0	5
Wattana	0	1	0	1	0	1	3
Total	7	27	9	28	5	7	83

Table S2.3. Overview of overall yawning *occurrences* and *rates* in the yawn and control condition per replication cycle

	Replication	Control	Yawn	Total
Yawn occurrences	1	7	8	15
	2	1	7	8
	3	2	6	8
	4	5	9	14
	Total	15	30	45
Yawn rates	1	7	16	23
	2	1	11	12
	3	3	17	20
	4	10	18	28
	Total	21	62	83

Table S3. Overview of number of trials per individual, per trigger and per condition.

Trigger	Familiar		Unfamiliar		Avatar		
Individual	Control	Yawn	Control	Yawn	Control	Yawn	Total
Amos	17	14	13	11	13	15	83
*Baju	9	6	9	6	4	10	44
*Indah	14	14	13	16	11	11	79
*Kawan	19	18	21	19	18	18	113
Kevin	17	14	17	15	15	16	94
Samboja	14	16	14	15	17	15	91
Sandy	15	19	16	18	14	13	95
Wattana	18	12	19	18	16	18	101
Total	123	113	122	118	108	116	700

* Baju and Kawan are the offspring of Wattana. Indah is the offspring of Samboja. Note: The numbers are not equally divided across individuals, as bystanders were present in the same enclosure as the focal were also exposed to the videos (for instance in the case of mother-offspring pairs). Furthermore, within individuals, numbers are not equal between conditions because in some cases the focal individual paid no attention to the screen, leading to no data for these particular trials. Finally, one video was accidentally presented an extra time, which means that after data collection we had 289 rather than the planned 288 sessions.

Testing the link between yawn occurrence, condition, and familiar and unfamiliar trigger (excluding avatar)

As CY seemed to be present only in response to a familiar or unfamiliar trigger, but not significantly so with the avatar trigger, we performed an extra exploratory analysis without the avatar trigger. Specifically, we first tested whether yawn occurrence is moderated by *condition* (yawn vs. control), *trigger* (familiar vs. unfamiliar), and their interaction using a binomial GLMM with *subject* nested in *trial*. Next, in those cases that at least one yawn occurred, we tested whether yawning rate is moderated by *condition, trigger,* and their interaction using a negative binomial GLMM with *subject* nested in *trial*.

In the first analysis looking at the likelihood of yawning, we re-confirmed the presence of CY in the reduced dataset: we found a main effect of *condition* ($\beta = 6.74$, $SE = 1.65$, $Z = 4.082$, $p < .001$) in which individuals were more likely to yawn in the yawn vs. control condition. We did not find evidence for an interaction between *condition*trigger* ($\beta = -3.24$, $SE = 2.05$, $Z = 1.57$, $p = .115$). Despite our reduced dataset, we still confirm the presence of CY in orangutans, but find no evidence for a familiarity effect using 'real' orangutan stimuli only.

In the second analysis in which we looked at yawning rates, we found no significant main effect of *condition* ($\beta = .37$, $SE = .46$, $Z = .81$, $p = .420$), nor an interaction effect between *condition*trigger* ($\beta = .32$, $SE = .61$, $Z = .52$, $p = .603$). Similar to the original analysis including the avatar trigger, we find no differences in strength between conditions and triggers. The model including the interaction between *condition* and *trigger* did not significantly improve the null model ($\chi^2(3) = 3.50$, $p = .321$).

Testing the link between contagious yawning, condition and familiarity using self-scratching as covariate

To control for self-scratching in our models that investigate the presence and strength of contagious yawning, we perform two Hurdle models, each containing two analyses. First, using a binomial GLMM with Subject nested in Trial as random factors and condition and self-scratch occurrence as fixed factors, we find a main effect of condition on the likelihood of yawning ($b = 3.52$, $SE = 1.05$, $Z = 3.35$, $p = .0008$), but no effect of self-scratch occurrence ($b = 1.33$, $SE = .96$, $Z = 1.39$, $p = .163$). In those cases that at least one yawn occurred, is the yawn response larger in the yawn condition versus the control condition? To answer this question, we perform a negative binomial GLMM with Subject nested in Trial as random factors and condition and self-scratch rate as fixed factors and compare this model to the null model (without fixed factors). The result shows that the alternative model cannot explain the data better than the null model: $\chi^2 2(2) = 3.32$, $p = .191$; there is no main effect of condition ($b = .47$, $SE = .26$, $Z = 1.80$, $p = .071$), nor a main effect of self-scratch rate ($b = .03$, $SE = .06$, $Z = .48$, $p = .634$) on yawn rate.

In the second hurdle model, we look at effects of familiarity on the occurrence and strength of contagious yawning. First, using a binomial GLMM with Subject nested in Trial as random factors and condition, trigger, condition*trigger, and self-scratch occurrence as fixed factors, we find a main effect of condition and condition*trigger on the likelihood of yawning, but no effect of self-scratch occurrence ($b = 1.33$, $SE = .96$, $Z = 1.39$, $p = .163$). Specifically, yawning is more likely to occur in the yawn versus control condition in case of a familiar trigger ($b = -6.50$, $SE = 1.64$, $Z = -3.97$, $p = .0001$) and an unfamiliar trigger ($b = -3.93$, $SE = 1.57$, $Z = -2.50$, $p = .012$). Next, in those cases that at least one yawn occurred, is the yawn response larger in the yawn condition versus the control condition and does familiarity affect this result? To answer these questions, we perform a negative binomial GLMM with Subject nested in Trial as random factors and condition, trigger, condition*trigger, and self-scratch rate as fixed factors and compare this model to its respective null model. The result shows that the alternative model cannot explain the data better than the null model: $\chi^2 2\ (6) = 5.30$, $p = .505$; there is no main effect of condition ($b = .18$, $SE = .62$, $Z = .28$, $p = .776$), nor a main effect of self-scratch rate ($b = .04$, $SE = .07$, $Z = .58$, $p = .560$), nor an interaction between condition*trigger ($b = .19$, $SE = .76$, $Z = .26$, $p = .798$) on yawn rate.

As such, we confirm that orangutans yawn contagiously in response to “real” orangutan stimuli, regardless of whether they are familiar or not. Furthermore, we cannot draw any conclusions on the rate of yawning. Importantly, self-scratching does not have a significant impact on the occurrence of the aforementioned results.

Appendix E

Supplementary Materials for Chapter 7

Experiment 1: PIAT in the zoo

Appendix to stimuli and procedure

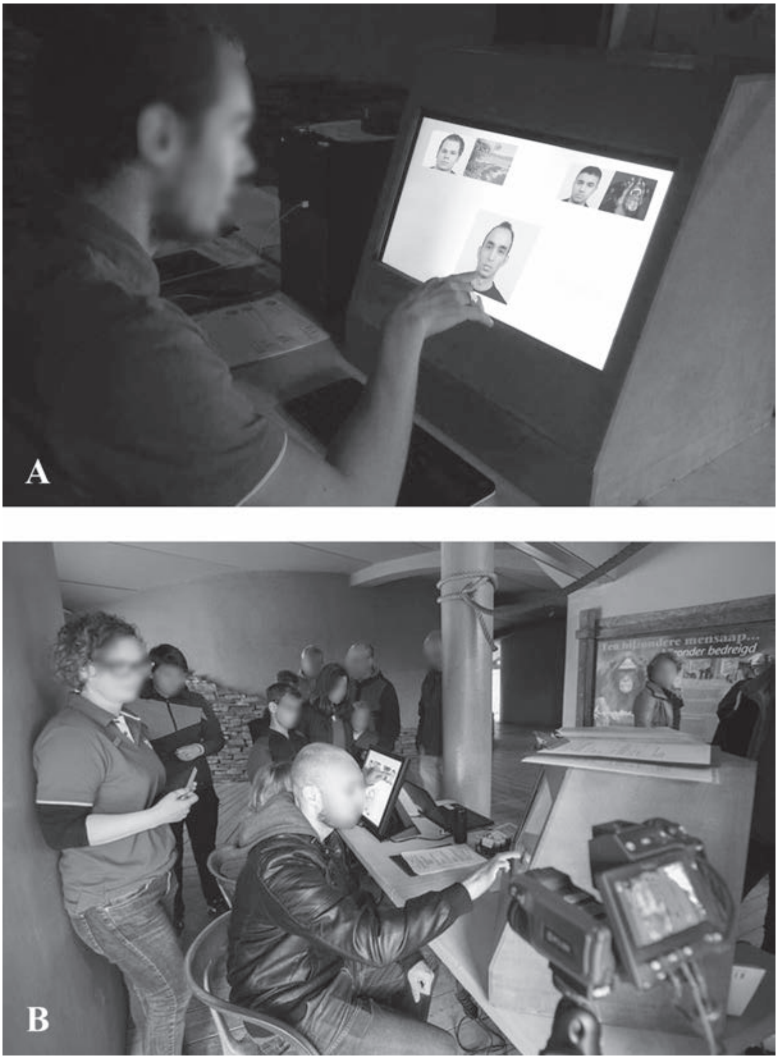

Figure S1. (A) Close-up of test setup used by the adults. (B) Photo of test environment (taken during a practice round). Photos were taken with permission.

Table S1.1. Overview of concept stimuli used in the adult and child P-IAT. The stimuli form a subset of the Radboud Faces Database (RaFD; Langer et al., 2010).

RaFD Image Name	Agreement between subjects	Intensity	Clarity	Genuineness	Valence
Rafd090_09_Dutch_male_ neutral_frontal*	87	3.30	3.74	3.65	2.87
Rafd090_10_Dutch_male_ neutral_frontal	86	3.55	3.86	3.82	2.95
Rafd090_20_Dutch_male_ neutral_frontal	68	2.82	3.32	3.55	2.68
Rafd090_24_Dutch_male_ neutral_frontal	96	3.84	4.16	4.24	3.24
Rafd090_28_Dutch_male_ neutral_frontal	88	3.36	3.76	3.92	2.88
Rafd090_36_Dutch_male_ neutral_frontal	100	3.92	3.92	4.04	3.24
Rafd090_35_Moroccan_male_ neutral_frontal	100	3.95	4.45	3.90	2.95
Rafd090_45_Moroccan_male_ neutral_frontal*	75	3.30	3.75	3.90	2.75
Rafd090_51_Moroccan_male_ neutral_frontal	89	3.68	4.11	4.05	3.05
Rafd090_52_Moroccan_male_ neutral_frontal	100	3.35	3.90	4.25	3.30
Rafd090_55_Moroccan_male_ neutral_frontal	95	3.55	3.85	4.05	2.85
Rafd090_59_Moroccan_male_ neutral_frontal	100	3.70	4.00	4.30	3.10

* These concepts are used as exemplars to indicate the "Dutch" and "Moroccan" categories.

Table S1.2. Overview of attribute stimuli used in the adult version of the P-IAT. Images form a subset of the International Affective Picture System (IAPS) (Lang, Bradley, & Cuthbert (2008).

Negative/Positive Attribute	Image Description	IAPS code	Valence (*SD*)	Arousal (*SD*)
neg1	Explosion*	9940	1.62(1.20)	7.15(2.24)
neg2	Injecting	9590	3.08(1.63)	5.41(2.23)
neg3	Snake	1120	3.79(1.93)	6.93(1.68)
neg4	Spider	1200	3.95(2.22)	6.03(2.38)
neg5	Attackdog*	1304	3.37(1.58)	6.37(1.93)
neg6	Fire*	8485	2.73(1.62)	6.46(2.10)
pos1	Seal*	1440	8.19(1.53)	6.05(2.38)
pos2	Polarbears	1441	7.97(1.28)	3.94(2.38)
pos3	Rabbit	1610	7.82(1.34)	3.08(2.19)
pos4	Sunset	5830	8.00(1.48)	4.92(2.65)
pos5	Nature*	5760	8.05(1.23)	3.22(2.39)
pos6	Icecream*	7330	7.69(1.84)	5.14(2.58)

* These attribute-images are used as exemplars to indicate the positive and negative categories.
Note. Descriptions and valence and arousal ratings are taken from Lang, Bradley, & Cuthbert (2008).

Table S1.3. Overview of valence and arousal ratings for the attribute images used in the child P-IAT.

Positive/Negative	Image Description	Valence (SD)	Arousal (SD)
Negative 1	Teenage Mutant Ninja Turtles villain	3.80 (1.92)	3.80 (2.39)
Negative 2	Spongebob villain	3.75 (2.22)	2.00 (0.82)
Negative 3	Pokémon villain	4.50 (1.91)	4.75 (1.71)
Negative 4	Avatar villain	4.20 (2.39)	4.80 (1.79)
Negative 5	Fairly Odd Parents villain	2.00 (2.45)	4.00 (2.76)
Negative 6	Totally Spies villain	2.80 (2.17)	3.40 (1.94)
Positive 1	Teenage Mutant Ninja Turtles villain	4.50 (2.50)	4.50 (2.50)
Positive 2	Spongebob hero	5.83 (0.98)	4.33 (1.75)
Positive 3	Pokémon heroes	6.00 (1.41)	5.40 (1.81)
Positive 4	Avatar heroes	7.00 (.00)	6.83 (.43)
Positive 5	Fairly Odd Parents heroes	5.83 (.75)	4.33(1.75)
Positive 6	Totally Spies heroes	4.00 (1.58)	3.00 (2.35)

Table S2. Overview of different versions of the PIAT. Attribute stimuli differ between the child and adult version, but have the same name in the table.

Stim. Set	Task	Attribute		Critical Block 3			Critical Block 5		
		Left	Right	Congruency	Left	Right	Congruency	Left	Right
1	1	neg1	pos1	congruent	MLNL	DRPR	incongruent	DLNL	MRPR
1	2	neg1	pos1	incongruent	DLNL	MRPR	congruent	MLNL	DRPR
1	3	pos1	neg1	incongruent	MLPL	DRNR	congruent	DLPL	MRNR
1	4	pos1	neg1	congruent	DLPL	MRNR	incongruent	MLPL	DRNR
2	1	neg2	pos2	congruent	MLNL2	DRPR2	incongruent	DLNL2	MRPR2
2	2	neg2	pos2	incongruent	DLNL	MRPR	congruent	MLNL2	DRPR2
2	3	pos2	neg2	incongruent	MLPL2	DRNR2	congruent	DLPL2	MRNR2
2	4	pos2	neg2	congruent	DLPL2	MRNR2	incongruent	MLPL2	DRNR2
3	1	neg3	pos3	congruent	MLNL3	DRPR3	incongruent	DLNL3	MRPR3
3	2	neg3	pos3	incongruent	DLNL3	MRPR3	congruent	MLNL3	DRPR3
3	3	pos3	neg3	incongruent	MLPL3	DRNR3	congruent	DLPL3	MRNR3
3	4	pos3	neg3	congruent	DLPL3	MRNR3	incongruent	MLPL3	DRNR3

Abbreviations under columns Left and Right represent the valence and position of the image on the screen, i.e.: MLNL = Moroccan Left, Negative Left; DLNL = Dutch Left, Negative Left; MLPL = Moroccan Left, Positive Left; DLPL = Dutch Left, Positive Left. DRPR = Dutch Right, Positive Right; MRPR = Moroccan Right, Positive Right; DRNR = Dutch Right, Negative Right; MRNR = Moroccan Right, Negative Right

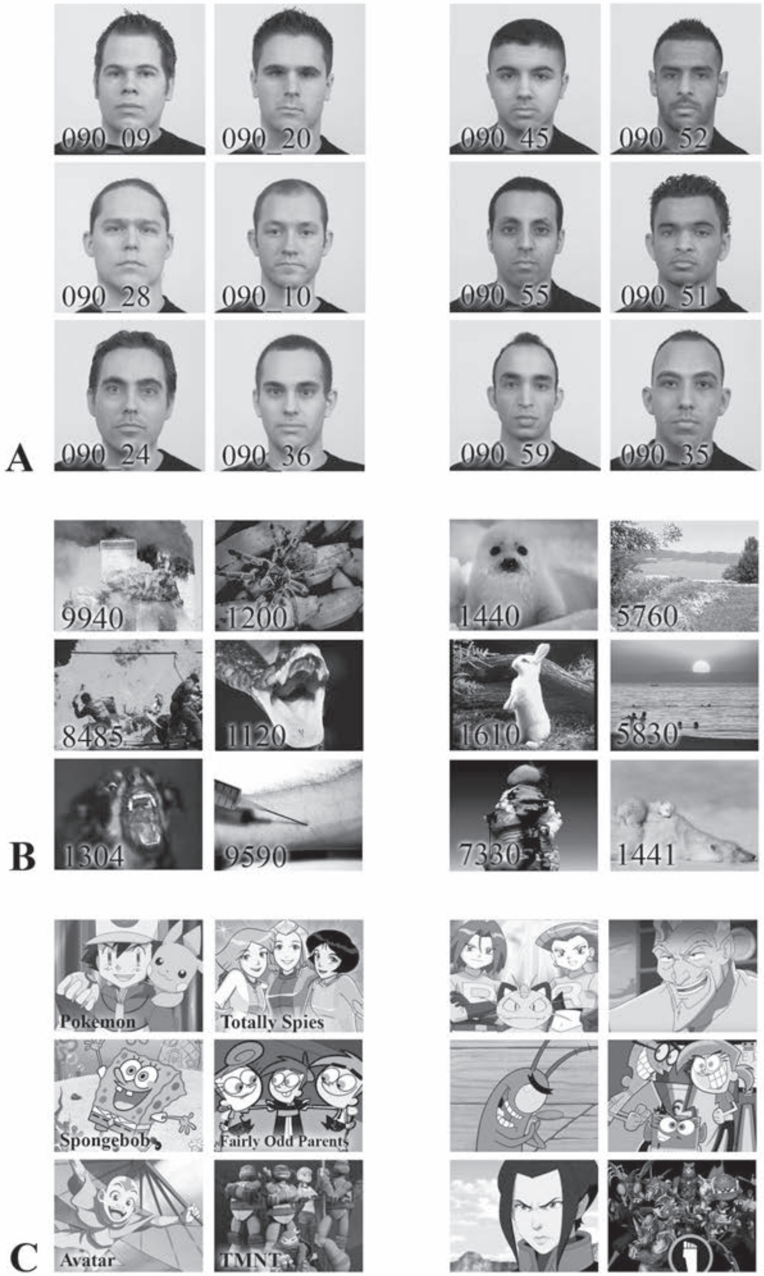

Figure S2. (A) The subset of Dutch and Moroccan faces selected from the Radboud Faces Database (Langer et al., 2010). (B) The subset of positive and negative image attributes selected from the International Affective Picture System (Lang, Bradley, & Cuthbert (2008), and used in the adult P-IAT. (C) The subset of positive and negative image attributes used in the child P-IAT. Images on the right represent the villains of the cartoons depicted on the left.

Figure S3. Overview of concept-attribute combinations for the adult (A) and child (B) P-IAT. In total, each version contains 24 unique combinations of concepts and attributes. Here we show only 12, but Dutch and Moroccan faces appeared either on the left or the right of a positive or negative attribute.

Appendix to results

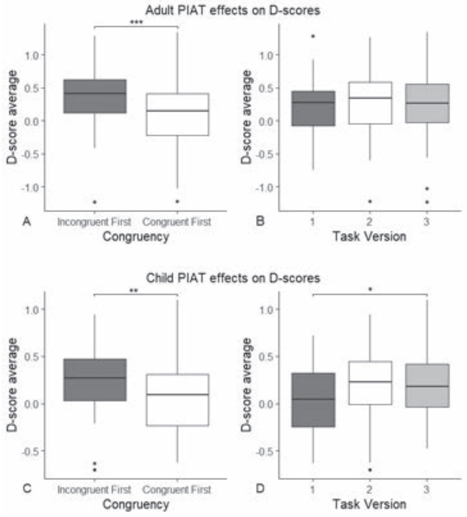

Figure S4. Box plots of effects of Congruency (A, C) and Task Version (B, D) on the Adult PIAT (A-B) and Child PIAT (C-D) D-score averages. *Task Version* represents the three different stimulus sets used in this study. D-scores are significantly higher for participants who receive incongruent trials first (A, C). Furthermore, task version 1 in the child PIAT shows a lower D-score average than version 2 and 3 (D). Outliers are visualized with a circle. * $p \leq .05$, ** $p < .01$. *** $p < .001$

Experiment 2: Online PIAT and WIAT

Appendix to procedure and results

Table S3. The randomization of starting positions for each block of the two IATs.

Block	Type	Trials	Tested	Version 1	Version 2	Version 3	Version 4
1	Training	20	Concepts only	M - D	M - D	D - M	D - M
2	Training	20	Attributes only	N - P	P - N	P - N	N – P
3	Experimental	20	Combined	MN - PD	MP - ND	DP - NM	DN - PM
4	Experimental	20	Combined	NM - DP	PM - DN	PD - MN	ND - MP
5	Training	20	Attributes only	P - N	N - P	N - P	P - N
6	Experimental	20	Combined	MP - ND	MN - PD	DN - PM	DP - NM
7	Experimental	20	Combined	PM - DN	NM - DP	ND - MP	PD - MN

Each participant only completed one randomly assigned version for each IAT. M = Moroccan ethnicity concept, C = Dutch ethnicity concept, P = positive attribute and N = negative attribute.

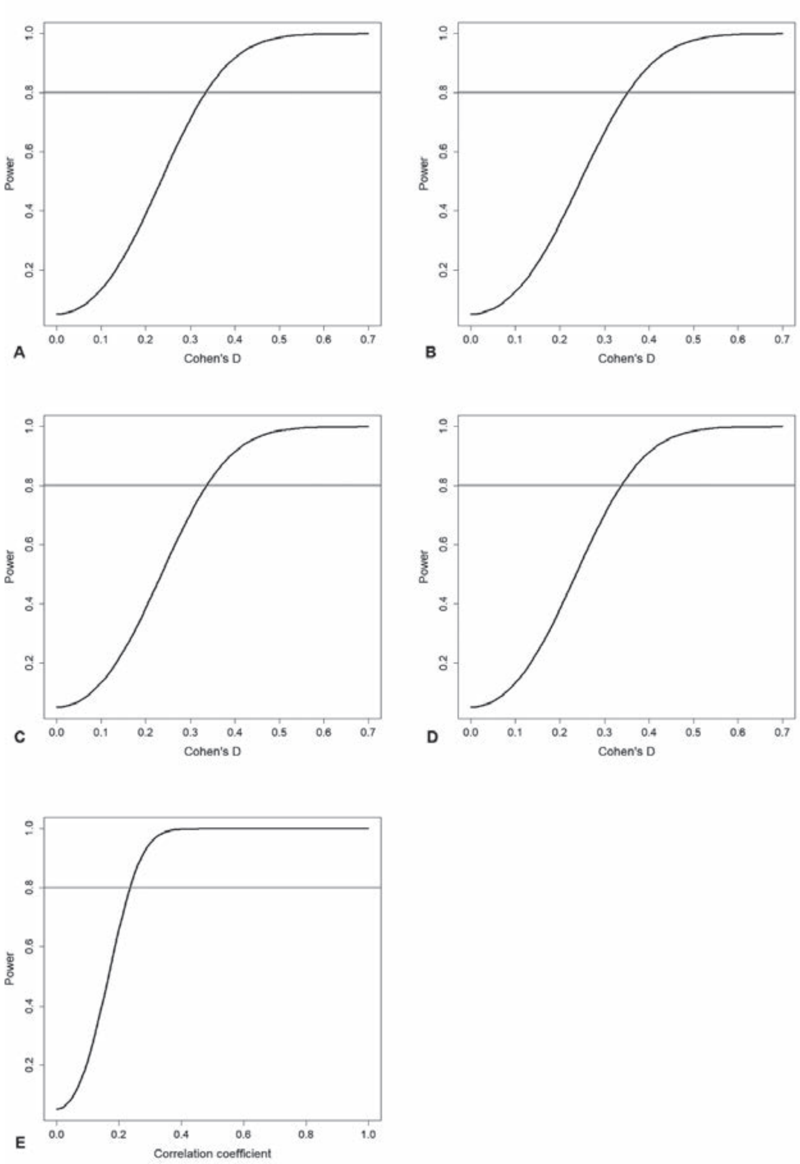

Figure S5. Curve results of the sensitivity power analyses (alpha = .05) for Experiment 1: A) Child PIAT, sample size = 142, B) Adult PIAT, sample size = 128, and Experiment 2: C) online PIAT, sample size = 140, D) online WIAT, sample size = 139, and E) correlation between the online PIAT and WIAT, sample size = 139. Given these sample sizes, we were able to detect true population effect sizes of (A): .34 (B): .35 (C): .34 (D): .34, and (E): .25 with ≥ 80% power (depicted with the horizontal red line).

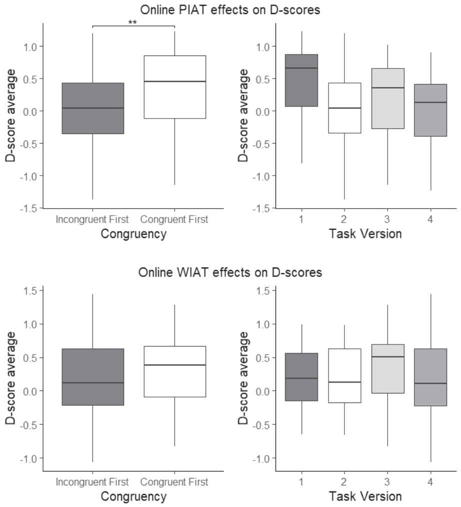

Figure S6. Box plots of the effects of Congruency and Task Version on D-scores in the online PIAT (top) and online WIAT (bottom). There was a significant effect of Congruency on D-score averages in the PIAT, with individuals receiving congruent trials first having higher D-scores averages. Task Version did not have a significant effect on scores in either the PIAT or WIAT. Finally, there was no Congruency effect in the WIAT.

References

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62. https://doi.org/10.1177/1534582302001001003

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology*, *18*(2), 166–172. https://doi.org/10.1016/j.conb.2008.06.006

Adolphs, R., Mlodinow, L., & Barrett, L. F. (2019). What is an emotion? *Current Biology*, *29*(20), R1060–R1064. https://doi.org/10.1016/j.cub.2019.09.008

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*(6507), 669–672. https://doi.org/10.1038/372669a0

Adriaense, J. E. C., Koski, S. E., Huber, L., & Lamm, C. (2020). Challenges in the comparative study of empathy and related phenomena in animals. *Neuroscience & Biobehavioral Reviews*, *112*, 62–82. https://doi.org/10.1016/j.neubiorev.2020.01.021

Adriaense, J. E. C., Martin, J. S., Schiestl, M., Lamm, C., & Bugnyar, T. (2019). Negative emotional contagion and cognitive bias in common ravens (Corvus corax). *Proceedings of the National Academy of Sciences*, *166*(23), 11547–11552. https://doi.org/10.1073/pnas.1817066116

Allritz, M., Call, J., & Borkenau, P. (2016). How chimpanzees (*Pan troglodytes*) perform in a modified emotional Stroop task. *Animal Cognition*, *19*(3), 435–449. https://doi.org/10.1007/s10071-015-0944-3

Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, *49*(3–4), 227–266. https://doi.org/10.1163/156853974X00534

Altschul, D. M., Wallace, E. K., Sonnweber, R., Tomonaga, M., & Weiss, A. (2017). Chimpanzee intellect: Personality, performance and motivation with touchscreen tasks. *Royal Society Open Science*, *4*(5). https://doi.org/10.1098/rsos.170169

Amici, F., Aureli, F., & Call, J. (2014). Response facilitation in the four great apes: Is there a role for empathy? *Primates*, *55*(1), 113–118. https://doi.org/10.1007/s10329-013-0375-1

Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, *157*(1), 187–200. https://doi.org/10.1016/j.cell.2014.03.003

Anderson, J. R., & Meno, P. (2003). Psychological Influences on Yawning in Children. *Current Psychology Letters*, *2*(11). https://doi.org/10.4000/cpl.390

Anderson, J. R., Myowa–Yamakoshi, M., & Matsuzawa, T. (2004). Contagious yawning in chimpanzees. *Biology Letters*, *271*(6), 468–470. https://doi.org/10.1098/rsbl.2004.0224

Andrews, P. (2020). Last Common Ancestor of Apes and Humans: Morphology and Environment. *Folia Primatologica*, *91*(2), 122–148. https://doi.org/10.1159/000501557

Andriessen, I., Hoegen Dijkhof, J., Van der Torre, A., Van den Berg, E., Pulles, I., Iedema, J., & De Voogd-Hamelink, M. (2020). *Ervaren discriminatie in Nederland*. https://www.scp.nl/publicaties/publicaties/2020/04/02/ervaren-discriminatie-in-nederland-ii

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review*, *32*(8), 704–723. https://doi.org/10.1016/j.cpr.2012.09.004

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography*, *24*(2), 149–163. https://doi.org/10.1007/s10548-011-0171-4

Aureli, F., Cords, M., & Van Schaik, C. P. (2002). Conflict resolution following aggression in gregarious animals: A predictive framework. *Animal Behaviour*, *64*(3), 325–343. https://doi.org/10.1006/anbe.2002.3071

Aureli, F., Cozzolino, R., Cordischi, C., & Scucchi, S. (1992). Kin-oriented redirection among Japanese macaques: an expression of a revenge system? *Animal Behaviour*, *44*(2), 283–291. https://doi.org/10.1016/0003-3472(92)90034-7

Aureli, F., & De Waal, F. B. M. (1997). Inhibition of social behavior in chimpanzees under high-density conditions. *American Journal of Primatology*, *41*(3), 213–228.

Aureli, F., Van Schaik, C. P., & Van Hooff, J. A. R. A. M. (1989). Functional aspects of reconciliation among captive long-tailed macaques (*Macaca fascicularis*). *American Journal of Primatology*, *19*(1), 39–51. https://doi.org/10.1002/ajp.1350190105

Aureli, F., & Yates, K. (2010). Distress prevention by grooming others in crested black macaques. *Biology Letters*, *6*(1), 27–29. https://doi.org/10.1098/rsbl.2009.0513

Backs, R. W., Da Silva, S. P., & Han, K. (2005). A comparison of younger and older adults' self-assessment Manikin ratings of affective pictures. *Experimental Aging Research*, *31*(4), 421–440. https://doi.org/10.1080/03610730500206808

Baenninger, R. (1987). Some comparative aspects of yawning in *Betta splendens*, *Homo sapiens*, *Panthera leo*, and *Papio sphinx*. *Journal of Comparative Psychology*, *101*(4), 349–354. https://doi.org/10.1037/0735-7036.101.4.349

Baenninger, R. (1997). On yawning and its functions. *Psychonomic Bulletin and Review*, *4*(2), 198–207. https://doi.org/10.3758/BF03209394

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1). https://doi.org/10.1037/0033-2909.133.1.1

Barbizet, J. (1958). Yawning. *Journal of Neurology, Neurosurgery, and Psychiatry*, *21*(3), 203–209.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*(1), 53–58. https://doi.org/10.1111/j.1467-9280.2005.01664.x

Bartholomew, A. J., & Cirulli, E. T. (2014). Individual variation in contagious yawning susceptibility is highly stable and largely unexplained by empathy or other known factors. *PLoS ONE*, *9*(3), e91773. https://doi.org/10.1371/journal.pone.0091773

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, *140*(4), 637–659. https://doi.org/10.1037/a0024060

Belopolsky, A. V., Devue, C., & Theeuwes, J. (2011). Angry faces hold the eyes. *Visual Cognition*, *19*(1), 27–36. https://doi.org/10.1080/13506285.2010.536186

Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology*, *9*, 1664–1078. https://doi.org/10.3389/fpsyg.2018.01647

Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence that emotion mediates social attention in rhesus macaques. *PLoS ONE*, *7*(8), e44387. https://doi.org/10.1371/journal.pone.0044387

Bijlstra, G., Holland, R. W., Dotsch, R., Hugenberg, K., & Wigboldus, D. H. J. (2014). Stereotype associations and emotion recognition. *Personality & Social Psychology Bulletin*, *40*(5), 567–577. https://doi.org/10.1177/0146167213520458

Blanco, I., Serrano-Pedraza, I., & Vázquez, C. (2017). Don't look at my teeth when i smile: Teeth visibility in smiling faces affects emotionality ratings and gaze patterns. *Emotion*, *17*(4), 640–647. https://doi.org/10.1037/emo0000260

Boone, R. T., & Buck, R. (2003). Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *Journal of Nonverbal Behavior*, *27*(3), 163–182. https://doi.org/10.1023/A:1025341931128

Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science*, *324*(5932), 1293–1298. https://doi.org/10.1126/science.1168112

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, *1*(3), 300–319. https://doi.org/10.1037/1528-3542.1.3.300

Bradshaw, H. R. (1993). Displacement activities as potential covert signals in primates. *Folia Primatologica*, *61*(3), 174–176. https://doi.org/10.1159/000156746

Brand, R., Heck, P., & Ziegler, M. (2014). Illegal performance enhancing drugs and doping in sport: A picture-based brief implicit association test for measuring athletes' attitudes. *Substance Abuse: Treatment, Prevention, and Policy*, *9*(1), 1–11. https://doi.org/10.1186/1747-597X-9-7

Bräuer, J., Hanus, D., Pika, S., Gray, R., & Uomini, N. (2020). Old and new approaches to animal cognition: There is not "one cognition." *Journal of Intelligence*, *8*(28), 1–25. https://doi.org/10.3390/jintelligence8030028

Bresciani, C., Cordoni, G., & Palagi, E. (2021). Playing together, laughing together: rapid facial mimicry and social sensitivity in lowland gorillas. *Current Zoology*. https://doi.org/10.1093/CZ/ZOAB092

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307. https://doi.org/10.1037/0033-2909.86.2.307

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB Balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear. Rapid spatial orienting toward positive emotional stimuli. *Psychological Science*, *19*(4), 362–370. https://doi.org/10.1111/j.1467-9280.2008.02094.x

Brosch, T., Scherer, K., Grandjean, D., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, *143*. https://doi.org/10.4414/smw.2013.13786

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1). https://doi.org/10.5334/joc.10

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

Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal*, *10*(1), 395–411. https://doi.org/10.32614/rj-2018-017

Burn, C. C. (2017). Bestial boredom: a biological perspective on animal boredom and suggestions for its scientific investigation. *Animal Behaviour*, *130*, 141–151. https://doi.org/10.1016/j.anbehav.2017.06.006

Burrows, A. M., Waller, B. M., Parr, L. A., & Bonar, C. J. (2006). Muscles of facial expression in the chimpanzee (Pan troglodytes): Descriptive, comparative and phylogenetic contexts. *Journal of Anatomy*, *208*(2), 153–167. https://doi.org/10.1111/j.1469-7580.2006.00523.x

Buttelmann, D., Call, J., & Tomasello, M. (2009). Do great apes use emotional expressions to infer desires? *Developmental Science*, *12*(5), 688–698. https://doi.org/10.1111/j.1467-7687.2008.00802.x

Buttner, A. P., & Strasser, R. (2014). Contagious yawning, social cognition, and arousal: An investigation of the processes underlying shelter dogs' responses to human yawns. *Animal Cognition*, *17*(1), 95–104. https://doi.org/10.1007/s10071-013-0641-z

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data (Vol. 53)* (A. C. Cameron & P. K. Trivedi (eds.)). Cambridge University Press.

Cameron, M. E., & Zeman, J. L. (2019). Development of an implicit association test for adolescent emotional expression. *Merrill-Palmer Quarterly*, *65*(3), 356–376. https://doi.org/10.13110/merrpalmquar1982.65.3.0356

Campbell, M. W., Carter, J. D., Proctor, D., Eisenberg, M. L., & De Waal, F. B. M. (2009). Computer animations stimulate contagious yawning in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1676), 4255–4259. https://doi.org/10.1098/rspb.2009.1087

Campbell, M. W., & Cox, C. R. (2019). Observational data reveal evidence and parameters of contagious yawning in the behavioral repertoire of captive-reared chimpanzees (*Pan troglodytes*). *Scientific Reports*, *9*(1), 1–13.

Campbell, M. W., & De Waal, F. B. M. (2011). Ingroup-outgroup bias in contagious yawning by chimpanzees supports link to empathy. *PLoS ONE*, *6*(4), 19–22. https://doi.org/10.1371/journal.pone.0018283

Campbell, M. W., & De Waal, F. B. M. (2014). Chimpanzees empathize with group mates and humans, but not with baboons or unfamiliar chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782). https://doi.org/10.1098/rspb.2014.0013

Carpenter, T. P., Pogacar, R., Pullig, C., Kouril, M., Aguilar, S., LaBouff, J., Isenberg, N., & Chakroff, A. (2019). Survey-software implicit association tests: A methodological and empirical analysis. *Behavior Research Methods*, *51*(5), 2194–2208. https://doi.org/10.3758/s13428-019-01293-3

Case, T. I., Stevenson, R. J., Byrne, R. W., & Hobaiter, C. L. (2020). The animal origins of disgust: Reports of basic disgust in Nonhuman great apes. *Evolutionary Behavioral Sciences*, *14*(3), 231–260. https://doi.org/10.1037/ebs0000175

Casetta, G., Nolfo, A. P., & Palagi, E. (2021). Yawn contagion promotes motor synchrony in wild lions, *Panthera leo*. *Animal Behaviour*, *174*, 149–159. https://doi.org/10.1016/j.anbehav.2021.02.010

Castles, D. L., & Whiten, A. (1998). Post-conflict behaviour of wild olive baboons. II. Stress and self-directed behaviour. *Ethology*, *104*(2), 148–160. https://doi.org/10.1111/j.1439-0310.1998.tb00058.x

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1226. https://doi.org/10.1126/SCIENCE.1165565

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. https://doi.org/10.1037/0022-3514.76.6.893

Clancy, B., Finlay, B. L., Darlington, R. B., & Anand, K. J. S. (2007). Extrapolating brain development from experimental species to humans. *NeuroToxicology*, *28*(5), 931–937. https://doi.org/10.1016/j.neuro.2007.01.014

Clay, Z., & De Waal, F. B. M. (2013). Bonobos respond to distress in others: Consolation across the age spectrum. *PLoS ONE*, *8*(1), 55206. https://doi.org/10.1371/journal.pone.0055206

Clay, Z., Palagi, E., & De Waal, F. B. M. (2018). Ethological approaches to empathy in primates. In *Neuronal Correlates of Empathy: From Rodent to Human* (pp. 53–66). Elsevier. https://doi.org/10.1016/B978-0-12-805397-3.00005-X

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Cooper, N. R., Puzzo, I., Pawley, A. D., Bowes-Mulligan, R. A., Kirkpatrick, E. V., Antoniou, P. A., & Kennett, S. (2012). Bridging a yawning chasm: EEG investigations into the debate concerning the role of the human mirror neuron system in contagious yawning. *Cognitive, Affective and Behavioral Neuroscience*, *12*(2), 393–405. https://doi.org/10.3758/s13415-011-0081-7

Cordoni, G., Norscia, I., Bobbio, M., & Palagi, E. (2018). Differences in play can illuminate differences in affiliation: A comparative study on chimpanzees and gorillas. *PLoS ONE*, *13*(3), 1–23. https://doi.org/10.1371/journal.pone.0193096

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford University Press.

Cosmides, L., & Tooby, J. (2000). Evolutionary Psychology and the Emotions. *Handbook of Emotions*, *2*(2), 91–115.

Crivelli, C., & Fridlund, A. J. (2018). Facial Displays Are Tools for Social Influence. *Trends in Cognitive Sciences*, *22*(5), 388–399. https://doi.org/10.1016/j.tics.2018.02.006

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit Attitude Measures: Consistency, Stability, and Convergent Validity. *Psychological Science*, *12*(2), 163–170. https://doi.org/10.1111/1467-9280.00328

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: Evaluative processing goals shape amygdala activity: Research article. *Psychological Science*, *19*(2), 152–160. https://doi.org/10.1111/j.1467-9280.2008.02061.x

Curtis, V., Barra, M. de, & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1563), 389–401. https://doi.org/10.1098/rstb.2010.0117

Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2011). Measuring implicit attitudes of 4-year-olds: The Preschool Implicit Association Test. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2010.11.002

Damerius, L. A., Burkart, J. M., van Noordwijk, M. A., Haun, D. B. M., Kosonen, Z. K., Galdikas, B. M. F., Saraswati, Y., Kurniawan, D., & van Schaik, C. P. (2019). General cognitive abilities in orangutans (Pongo abelii and Pongo pygmaeus). *Intelligence*, *74*, 3–11. https://doi.org/10.1016/j.intell.2018.10.008

Darwin, C. (1872). *The expression of the emotions in man and animals.* John Murray. https://doi.org/10.1037/10001-000

Davila-Ross, M., Menzler, S., & Zimmermann, E. (2008). Rapid facial mimicry in orangutan play. *Biology Letters*, *4*(1), 27–30. https://doi.org/10.1098/rsbl.2007.0535

De Gelder, B., Böcker, K. B. E., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, *260*(2), 133–136. https://doi.org/10.1016/S0304-3940(98)00963-X

De Gelder, B., Meeren, H. K. M., Righart, R., Van den Stock, J., Van de Riet, W. A. C., & Tamietto, M. (2006). Beyond the face: exploring rapid influences of context on face processing. *Progress in Brain Research*, *155*, 37–48. https://doi.org/10.1016/S0079-6123(06)55003-4

De Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., & Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*, *34*(4), 513–527. https://doi.org/10.1016/j.neubiorev.2009.10.008

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. https://doi.org/10.1080/026999300378824

De Vries, J. I. P., Visser, G. H. A., & Prechtl, H. F. R. (1982). The emergence of fetal behaviour. I. Qualitative aspects. *Early Human Development*, *7*(4). https://doi.org/10.1016/0378-3782(82)90033-0

De Waal, F. B. M. (1988). The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour*, *106*(3–4), 183–251. https://doi.org/10.1163/156853988X00269

De Waal, F. B. M. (2011). What is an animal emotion? *Annals of the New York Academy of Sciences*, *1224*(1), 191–206. https://doi.org/10.1111/j.1749-6632.2010.05912.x

De Waal, F. B. M., & Pokorny, J. J. (2011). Faces and behinds: Chimpanzee sex perception. *Advanced Science Letters*, *1*(1), 99–103. https://doi.org/10.1166/asl.2008.006

De Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: Behavioural manifestations and neural basis. *Nature Reviews Neuroscience*, *18*(8), 498–509. https://doi.org/10.1038/nrn.2017.72

De Waal, F. B. M., & Van Roosmalen, A. (1979). Reconciliation and consolation among chimpanzees. *Behavioral Ecology and Sociobiology*, *5*(1), 55–66. https://doi.org/10.1007/BF00302695

Decety, J., Norman, G. J., Berntson, G. G., & Cacioppo, J. T. (2012). A neurobehavioral evolutionary perspective on the mechanisms underlying empathy. *Progress in Neurobiology*, *98*(1), 38–48. https://doi.org/10.1016/j.pneurobio.2012.05.001

Delgado, R. A., & Van Schaik, C. P. (2000). The behavioral ecology and conservation of the orangutan (*Pongo pygmaeus*): A tale of two islands. *Evolutionary Anthropology*, *9*(5), 201–218. https://doi.org/10.1002/1520-6505

Demuru, E., & Palagi, E. (2012). In bonobos yawn contagion is higher among kin and friends. *PLoS ONE*, *7*(11), e49613. https://doi.org/10.1371/journal.pone.0049613

Deng, X., An, S., & Cheng, C. (2019). Cultural differences in the implicit and explicit attitudes toward emotion regulation. *Personality and Individual Differences*, *149*, 220–222. https://doi.org/10.1016/j.paid.2019.05.057

Depaoli, S., & van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods*, *22*(2), 240–261. https://doi.org/10.1037/met0000065

Deputte, B. L. (1994). *Ethological study of yawning in primates quantitative analysis and study of causation in two species of Old World Monkeys (Cercocebus albigena and Macaca fascicularis)* [University of Rennes]. http://baillement.com/deputte-monkey.html

Dezecache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2017). Skin temperature changes in wild chimpanzees upon hearing vocalizations of conspecifics. *Royal Society Open Science*, *4*(1). https://doi.org/10.1098/rsos.160816

Diogo, R. (2018). First detailed anatomical study of bonobos reveals intra-specific variations and exposes just-so stories of human evolution, bipedalism, and tool use. *Frontiers in Ecology and Evolution*, *6*, 53. https://doi.org/10.3389/fevo.2018.00053

Duboscq, J., Agil, M., Engelhardt, A., & Thierry, B. (2014). The function of postconflict interactions: new prospects from the study of a tolerant species of primate. *Animal Behaviour*, *87*, 107–120. https://doi.org/10.1016/j.anbehav.2013.10.018

Duda, P., & Zrzavý, J. (2013). Evolution of life history and behavior in Hominidae: Towards phylogenetic reconstruction of the chimpanzee-human last common ancestor. *Journal of Human Evolution*, *65*(4), 424–446. https://doi.org/10.1016/j.jhevol.2013.07.009

Dudeney, J., Sharpe, L., & Hunt, C. (2015). Attentional bias towards threatening stimuli in children with anxiety: A meta-analysis. *Clinical Psychology Review*, *40*, 66–75. https://doi.org/10.1016/j.cpr.2015.05.007

Duka, T., & Townshend, J. M. (2004). The priming effect of alcohol pre-load on attentional bias to alcohol-related stimuli. *Psychopharmacology*, *176*(3), 353–361. https://doi.org/10.1007/S00213-004-1906-7

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., Broomhall, S., Brosch, T., Campos, J. J., Clay, Z., Clément, F., Cunningham, W. A., Damasio, A., Damasio, H., D'Arms, J., Davidson, J. W., De Gelder, B., Deonna, J., De Sousa, R., Sander, D. (2021). The rise of affectivism. *Nature Human Behaviour*, *5*(7), 816–820. https://doi.org/10.1038/s41562-021-01130-8

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

Dunbar, R. I. M., & Dunbar, P. (1975). Social dynamics of gelada baboons. *Contributions to Primatology*, *6*, 1–157. https://doi.org/10.1016/0003-3472(77)90133-6

Dunham, Y., Baron, A. S., & Banaji, M. R. (2006). From american city to japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development*, *77*(5), 1268–1281. https://doi.org/10.1111/j.1467-8624.2006.00933.x

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, *12*(7), 248–253. https://doi.org/10.1016/j.tics.2008.04.006

Edgar, J. L., & Nicol, C. J. (2018). Socially-mediated arousal and contagion within domestic chick broods. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-28923-8

Edwards, S. D., & Snowdon, C. T. (1980). Social behavior of captive, group-living Orangutans. *International Journal of Primatology*, *1*(1), 39–62. https://doi.org/10.1007/BF02692257

Egelkamp, C. L., & Ross, S. R. (2019). A review of zoo-based cognitive research using touchscreen interfaces. *Zoo Biology*, *38*(2), 220–235. https://doi.org/10.1002/zoo.21458

Eguibar, J. R., Uribe, C. A., Cortes, C., Bautista, A., & Gallup, A. C. (2017). Yawning reduces facial temperature in the high-yawning subline of Sprague-Dawley rats. *BMC Neuroscience*, *18*(1), 1–8. https://doi.org/10.1186/s12868-016-0330-3

Ekman, P. (1984). Expression and the nature of emotion. In K. Scherer & P. Ekman (Eds.), *Approaches to Emotion* (pp. 319–343). Lawrence Erlbaum.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). John Wiley & Sons, Ltd.

Elder, C. M., & Menzel, C. R. (2001). Dissociation of cortisol and behavioral indicators of stress in an orangutan (*Pongo pygmaeus*) during a computerized task. *Primates*, *42*(4), 345–357. https://doi.org/10.1007/BF02629625

Elo, H. (2011). Yawning cannot cause significant temperature decreases in humans. *Sleep Medicine*, *12*(1), 102–103. https://doi.org/10.1016/j.sleep.2010.06.005

Ember, C. R., & Ember, M. (1992). Warfare, aggression, and resource problems: Cross-cultural codes. *Cross-Cultural Research*, *26*(1–4), 169–226. https://doi.org/10.1177/106939719202600108

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*(1), 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079–1083. https://doi.org/10.1038/nature07155

Feneran, A. N., O'Donnell, R., Press, A., Yosipovitch, G., Cline, M., Dugan, G., Papoiu, A. D. P., Nattkemper, L. A., Chan, Y. H., & Shively, C. A. (2013). Monkey see, monkey do: Contagious itch in nonhuman primates. *Acta Dermato-Venereologica*, *93*(1), 27–29. https://doi.org/10.2340/00015555-1406

Ferretti, V., & Papaleo, F. (2019). Understanding others: Emotion recognition in humans and other animals. *Genes, Brain and Behavior*, *18*(1), 1–12. https://doi.org/10.1111/gbb.12544

Foroni, F., & Bel-Bahar, T. (2009). Picture -IAT versus Word -IAT: level of stimulus representation influences on the IAT. *European Journal of Social Psychology*, n/a-n/a. https://doi.org/10.1002/ejsp.626

Forsell, L. M., & Åström, J. A. (2012). Meanings of Hugging: From Greeting Behavior to Touching Implications. *Comprehensive Psychology*, *1*. https://doi.org/10.2466/02.17.21.CP.1.13

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Frijda, N. H. (2017). The laws of emotion. In *The Laws of Emotion*. Taylor & Francis Group. https://doi.org/10.4324/9781315086071

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*(2), 212–228. https://doi.org/10.1037/0022-3514.57.2.212

Fröhlich, M., Lee, K., Setia, T. M., Schuppli, C., & Van Schaik, C. P. (2019a). The loud scratch: A newly identified gesture of Sumatran orangutan mothers in the wild. *Biology Letters*, *15*(7). https://doi.org/10.1098/rsbl.2019.0209

Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & van Schaik, C. P. (2019b). 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. (2016). 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

Fruth, B., Hickey, J. R., André, C., Furuichi, T., Hart, J., Hart, T., Kuehl, H., Maisels, F., Nackoney, J., Reinartz, G., Sop, T., Thompson, J., & Williamson, E. A. (2016). *Pan paniscus* (errata version published in 2016). In *The IUCN Red List of Threatened Species* (Issues 24-03–2016). https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15932A17964305.en

Fruth, B., & Hohmann, G. (2018). Food Sharing across Borders: First Observation of Intercommunity Meat Sharing by Bonobos at LuiKotale, DRC. *Human Nature*, *29*(2), 91–103. https://doi.org/10.1007/s12110-018-9311-9

Fugate, J. M. B., Gouzoules, H., & Barrett, L. F. (2010). Reading chimpanzee faces: Evidence for the role of verbal labels in categorical perception of emotion. *Emotion*, *10*(4), 544–554. https://doi.org/10.1037/a0019017

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. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthropology*, *20*, 131–142. https://doi.org/10.1093/oso/9780198728511.003.0002

Furuichi, T., & Thompson, J. (2007). *The bonobos: Behavior, ecology, and conservation.* Springer Science & Business Media.

Galdikas, B. M. F. (1985). Orangutan sociality at Tanjung Puting. *American Journal of Primatology*, *9*(2), 101–119. https://doi.org/10.1002/ajp.1350090204

Gallo, A., Zanoli, A., Caselli, M., Palagi, E., & Norscia, I. (2021). First evidence of yawn contagion in a wild monkey species. *Scientific Reports*, *11*(1), 1–8. https://doi.org/10.1038/s41598-021-96423-3

Gallup, A. C. (2011). Why do we yawn? Primitive versus derived features. *Neuroscience & Biobehavioral Reviews*, *35*(3), 765–769. https://doi.org/10.1016/j.neubiorev.2010.09.009

Gallup, A. C. (2021). On the link between emotional contagion and contagious yawning. *Neuroscience and Biobehavioral Reviews*, *121*, 18–19. https://doi.org/10.1016/j.neubiorev.2020.11.023

Gallup, A. C., & Eldakar, O. T. (2012). The thermoregulatory theory of yawning: What we know from over 5 years of research. *Frontiers in Neuroscience*, *6*, 188. https://doi.org/10.3389/fnins.2012.00188

Gallup, A. C., & Gallup, G. G. (2007). Yawning as a Brain Cooling Mechanism: Nasal Breathing and Forehead Cooling Diminish the Incidence of Contagious Yawning. *Evolutionary Psychology*, *5*(1). https://doi.org/10.1177/147470490700500109

Gallup, A. C., & Gallup, G. G. (2008). Yawning and thermoregulation. *Physiology and Behavior*, *95*(1–2), 10–16. https://doi.org/10.1016/j.physbeh.2008.05.003

Gallup, A. C., & Massen, J. J. M. (2016). There is no difference in contagious yawning between men and women. *Royal Society Open Science*, *3*(9), 160174. https://doi.org/10.1098/rsos.160174

Gallup, A. C., Swartwood, L., Militello, J., & Sackett, S. (2015). Experimental evidence of contagious yawning in budgerigars (*Melopsittacus undulatus*). *Animal Cognition*, *18*(5), 1051–1058. https://doi.org/10.1007/s10071-015-0873-1

Gallup, G. G., & Gallup, A. C. (2010). Excessive yawning and thermoregulation: Two case histories of chronic, debilitating bouts of yawning. *Sleep and Breathing*, *14*(2), 157–159. https://doi.org/10.1007/s11325-009-0287-x

Gao, J., & Tomonaga, M. (2020). Body perception in chimpanzees (*Pan troglodytes*): The effect of body structure changes. *Journal of Comparative Psychology*, *134*(2), 222–231. https://doi.org/10.1037/com0000214

Gard, M. G., & Kring, A. M. (2007). Sex differences in the time course of emotion. *Emotion*, *7*(2), 429–437. https://doi.org/10.1037/1528-3542.7.2.429

Genty, E., Neumann, C., & Zuberbühler, K. (2015). Complex patterns of signalling to convey different social goals of sex in bonobos, *Pan paniscus*. *Scientific Reports*, *5*(1), 16135. https://doi.org/10.1038/srep16135

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*(4), 1628–1635. https://doi.org/10.1016/J.NEUROIMAGE.2004.03.049

Goldberg, T. L., & Wrangham, R. W. (1997). Genetic correlates of social behaviour in wild chimpanzees: Evidence from mitochondrial DNA. *Animal Behaviour*, *54*(3), 559–570. https://doi.org/10.1006/anbe.1996.0450

Goldsborough, Z., Van Leeuwen, E. J. C., Kolff, K. W. T., De Waal, F. B. M., & Webb, C. E. (2020). Do chimpanzees (*Pan troglodytes*) console a bereaved mother? *Primates*, *61*(1), 93–102. https://doi.org/10.1007/s10329-019-00752-x

Goodman, M., Porter, C. A., Czelusniak, J., Page, S. L., Schneider, H., Shoshani, J., Gunnell, G., & Groves, C. P. (1998). Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. *Molecular Phylogenetics and Evolution*, *9*(3), 585–598.

Greenwald, A. G. (2001). Health of the Implicit Association Test at age 3. *Experimental Psychology*, *48*(2), 95–93. https://doi.org/10.1026//0949-3946.48.2.85

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1). https://doi.org/10.1037//0033-295x.109.1.3

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of Personality and Social Psychology*, *85*(2). https://doi.org/10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Gruber, T., & Clay, Z. (2016). A Comparison Between Bonobos and Chimpanzees: A Review and Update. *Evolutionary Anthropology*, *25*(5), 239–252. https://doi.org/10.1002/evan.21501

Guggisberg, A. G., Mathis, J., Schnider, A., & Hess, C. W. (2010). Why do we yawn? *Neuroscience & Biobehavioral Reviews*, *34*(8), 1267–1276. https://doi.org/10.1016/j.neubiorev.2010.03.008

Guggisberg, A. G., Mathis, J., Schnider, A., & Hess, C. W. (2011). Why do we yawn? The importance of evidence for specific yawn-induced effects. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2010.12.004

Guo, K., Li, Z., Yan, Y., & Li, W. (2019). Viewing heterospecific facial expressions: an eye-tracking study of human and monkey viewers. *Experimental Brain Research*, *237*(8), 2045–2059. https://doi.org/10.1007/s00221-019-05574-3

Haker, H., Kawohl, W., Herwig, U., & Rössler, W. (2013). Mirror neuron activity during contagious yawning-an fMRI study. *Brain Imaging and Behavior*, *7*(1), 28–34. https://doi.org/10.1007/s11682-012-9189-9

Hammer, J. L., & Marsh, A. A. (2015). Why do fearful facial expressions elicit behavioral approach? Evidence from a combined approach-avoidance implicit association test. *Emotion*, *15*(2), 223–231. https://doi.org/10.1037/emo0000054

Hare, B., & Kwetuenda, S. (2010). Bonobos voluntarily share their own food with others. *Current Biology*, *20*(5), 230–231. https://doi.org/10.1016/j.cub.2009.12.038

Hare, B., Wobber, V., & Wrangham, R. W. (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

Harr, A. L., Gilbert, V. R., & Phillips, K. A. (2009). Do dogs (Canis familiaris) show contagious yawning? *Animal Cognition*, *12*(6), 833–837. https://doi.org/10.1007/s10071-009-0233-0

Harrison, M. E., & Chivers, D. J. (2007). The orang-utan mating system and the unflanged male: A product of increased food stress during the late Miocene and Pliocene? *Journal of Human Evolution*, *52*(3), 275–293. https://doi.org/10.1016/j.jhevol.2006.09.005

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional Contagion. *Current Directions in Psychological Science*, *2*(3), 96–100. https://doi.org/10.1111/1467-8721.ep10770953

Heiman, J. R. (1977). A Psychophysiological Exploration of Sexual Arousal Patterns in Females and Males. *Psychophysiology*, *14*(3), 266–274. https://doi.org/10.1111/j.1469-8986.1977.tb01173.x

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? In *Behavioral and Brain Sciences* (Vol. 33, Issues 2–3, pp. 61–83). Cambridge University Press. https://doi.org/10.1017/S0140525X0999152X

Henry, P. J., & Sears, D. O. (2002). The Symbolic Racism 2000 Scale. *Political Psychology*, *23*(2), 253–283. https://doi.org/10.1111/0162-895X.00281

Hess, U., Blaison, C., & Kafetsios, K. (2016). Judging Facial Emotion Expressions in Context: The Influence of Culture and Self-Construal Orientation. *Journal of Nonverbal Behavior*, *40*(1), 55–64. https://doi.org/10.1007/s10919-015-0223-7

Higham, T., Douka, K., Wood, R., Ramsey, C. B., Brock, F., Basell, L., Camps, M., Arrizabalaga, A., Baena, J., Barroso-Ruíz, C., Bergman, C., Boitard, C., Boscato, P., Caparrós, M., Conard, N. J., Draily, C., Froment, A., Galván, B., Gambassini, P., … Jacobi, R. (2014). The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature*, *512*(7514), 306–309. https://doi.org/10.1038/nature13621

Hirata, S., Matsuda, G., Ueno, A., Fukushima, H., Fuwa, K., Sugama, K., Kusunoki, K., Tomonaga, M., Hiraki, K., & Hasegawa, T. (2013). Brain response to affective pictures in the chimpanzee. *Scientific Reports*, *3*(1), 1–5. https://doi.org/10.1038/srep01342

Hobaiter, C. L., & 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

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369–1385. https://doi.org/10.1177/0146167205275613

Hohmann, G., Gerloff, U., Tautz, D., & Fruth, B. (1999). Social bonds and genetic ties: Kinship, association and affiliation in a community of bonobos (*Pan paniscus*). *Behaviour*, *136*(9), 1219–1235. https://doi.org/10.1163/156853999501739

Hoogenhout, M. (2013). Young Children Display Contagious Yawning When Looking at the Eyes. *Journal of Child and Adolescent Behaviour*, *1*(1), 1–4. https://doi.org/10.4172/jcalb.1000101

Hopper, L. M. (2017). Cognitive research in zoos. *Current Opinion in Behavioral Sciences*, *16*, 100–110. https://doi.org/10.1016/j.cobeha.2017.04.006

Hopper, L. M., Allritz, M., Egelkamp, C. L., Huskisson, S. M., Jacobson, S. L., Leinwand, J. G., & Ross, S. R. (2021a). A Comparative Perspective on Three Primate Species' Responses to a Pictorial Emotional Stroop Task. *Animals*, *11*(3), 588. https://doi.org/10.3390/ani11030588

Hopper, L. M., Gulli, R. A., Howard, L. H., Kano, F., Krupenye, C., Ryan, A. M., & Paukner, A. (2021b). The application of noninvasive, restraint-free eye-tracking methods for use with nonhuman primates. *Behavior Research Methods*, *53*(3), 1003–1030. https://doi.org/10.3758/s13428-020-01465-6

Howarth, E. R. I., Kemp, C., Thatcher, H. R., Szott, I. D., Farningham, D., Witham, C. L., Holmes, A., Semple, S., & Bethell, E. J. (2021). Developing and validating attention bias tools for assessing trait and state affect in animals: A worked example with *Macaca mulatta*. *Applied Animal Behaviour Science*, *234*, 105198. https://doi.org/10.1016/J.APPLANIM.2020.105198

Hrdy, S. B., & Burkart, J. M. (2020). The emergence of emotionally modern humans: Implications for language and learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803). https://doi.org/10.1098/rstb.2019.0499

Issa, H. A., Staes, N., Diggs-Galligan, S., Stimpson, C. D., Gendron-Fitzpatrick, A., Taglialatela, J. P., Hof, P. R., Hopkins, W. D., & Sherwood, C. C. (2019). Comparison of bonobo and chimpanzee brain microstructure reveals differences in socio-emotional circuits. *Brain Structure and Function*, *224*(1), 239–251. https://doi.org/10.1007/s00429-018-1751-9

James, W. (1884). What is an emotion? *Mind*, *9*(34), 188–205.

James, W. (1890). *The principles of psychology*. Henry Holt and Company.

Janssen, E., Carpenter, D., & Graham, C. A. (2003). Selecting Films for Sex Research: Gender Differences in Erotic Film Preference. *Archives of Sexual Behavior*, *32*(3), 243–251. https://doi.org/10.1023/A:1023413617648

Jeffreys, H. (1961). *Theory of probability* (3rd Editio). Clarendon Press.

Joly-Mascheroni, R. M., Senju, A., & Shepherd, A. J. (2008). Dogs catch human yawns. *Biology Letters*, *4*(5), 446–448. https://doi.org/10.1098/rsbl.2008.0333

Judge, P. G., Griffaton, N. S., & Fincke, A. M. (2006). Conflict management by hamadryas baboons (Papio hamadryas hamadryas) during crowding: A tension-reduction strategy. *American Journal of Primatology*, *68*(10), 993–1006. https://doi.org/10.1002/ajp.20290

Kaburu, S. S. K., Maclarnon, A., Majolo, B., Qarro, M., & Semple, S. (2012). Dominance rank and self-scratching among wild female Barbary macaques (*Macaca sylvanus*). *African Zoology*, *47*(1), 74–79. https://doi.org/10.3377/004.047.0111

Kano, F., Hirata, S., & Call, J. (2015). Social attention in the two species of pan: Bonobos make more eye contact than chimpanzees. *PLoS ONE*, *10*(6), 1–14. https://doi.org/10.1371/journal.pone.0129684

Kano, F., Hirata, S., Deschner, T., Behringer, V., & Call, J. (2016). Nasal temperature drop in response to a playback of conspecific fights in chimpanzees: A thermo-imaging study. *Physiology and Behavior*, *155*, 83–94. https://doi.org/10.1016/j.physbeh.2015.11.029

Kano, F., Shepherd, S. V., Hirata, S., & Call, J. (2018). Primate social attention: Species differences and effects of individual experience in humans, great apes, and macaques. *PLoS ONE*, *13*(2), e0193283. https://doi.org/10.1371/journal.pone.0193283

Kano, F., Tanaka, M., & Tomonaga, M. (2008). Enhanced recognition of emotional stimuli in the chimpanzee (Pan troglodytes). *Animal Cognition*, *11*(3), 517–524. https://doi.org/10.1007/s10071-008-0142-7

Kano, F., & Tomonaga, M. (2010a). Attention to emotional scenes including whole-body expressions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *124*(3), 287–294. https://doi.org/10.1037/a0019146

Kano, F., & Tomonaga, M. (2010b). Face scanning in chimpanzees and humans: continuity and discontinuity. *Animal Behaviour*, *79*(1), 227–235. https://doi.org/10.1016/j.anbehav.2009.11.003

Kapitány, R., & Nielsen, M. (2017). Are Yawns really Contagious? A Critique and Quantification of Yawn Contagion. *Adaptive Human Behavior and Physiology*, *3*(2), 134–155. https://doi.org/10.1007/s40750-017-0059-y

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kim, Y., Van Berlo, E., Bionda, T., & Kret, M. E. (2021). Visual perception of conspecific emotional expressions: the role of expression channels in humans and bonobos. *Manuscript in Preparation*.

King, H. M., Kurdziel, L. B., Meyer, J. S., & Lacreuse, A. (2012). Effects of testosterone on attention and memory for emotional stimuli in male rhesus monkeys. *Psychoneuroendocrinology*, *37*(3), 396–409. https://doi.org/10.1016/j.psyneuen.2011.07.010

Knott, C. D., Thompson, M. E., Stumpf, R. M., & McIntyre, M. H. (2010). Female reproductive strategies in orangutans, evidence for female choice and counterstrategies to infanticide in a species with frequent sexual coercion. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1678), 105–113. https://doi.org/10.1098/rspb.2009.1552

Kolbrink, P. (2017). *Waldo*. XYZ-Animation. http://animatio.com/waldo-page.html

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van der Vegt, B. J., Van Reenen, C. G., Hopster, H., De Jong, I. C., Ruis, M. A. W., & Blokhuis, H. J. (1999). Coping styles in animals: Current status in behavior and stress- physiology. *Neuroscience and Biobehavioral Reviews*, *23*(7), 925–935. https://doi.org/10.1016/S0149-7634(99)00026-3

Koski, S. E., & Sterck, E. H. M. (2010). Empathic chimpanzees: A proposal of the levels of emotional and cognitive processing in chimpanzee empathy. *European Journal of Developmental Psychology*, *7*(1), 38–66. https://doi.org/10.1080/17405620902986991

Koster, E. H. W., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*(10), 1183–1192. https://doi.org/10.1016/j.brat.2003.08.001

Koukounas, E., & McCabe, M. (1997). Sexual and emotional variables influencing sexual response to erotica. *Behaviour Research and Therapy*, *35*(3), 221–230. https://doi.org/10.1016/S0005-7967(96)00097-6

Kremer, L., Bus, J. D., Webb, L. E., Bokkers, E. A. M., Engel, B., Van der Werf, J. T. N., Schnabel, S. K., & Van Reenen, C. G. (2021). Housing and personality effects on judgement and attention biases in dairy cows. *Scientific Reports*, *11*(1), 1–18. https://doi.org/10.1038/s41598-021-01843-w

Kret, M. E. (2015). Emotional expressions beyond facial muscle actions. A call for studying autonomic signals and their impact on social perception. *Frontiers in Psychology*, *6*, 711. https://doi.org/10.3389/fpsyg.2015.00711

Kret, M. E., & De Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*, 169–180. https://doi.org/10.1007/S00221-010-2220-8

Kret, M. E., & De Gelder, B. (2012a). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211–1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022

Kret, M. E., & De Gelder, B. (2012b). Islamic headdress influences how emotion is recognized from the eyes. *Frontiers in Psychology*, *3*(APR), 110. https://doi.org/10.3389/fpsyg.2012.00110

Kret, M. E., & De Gelder, B. (2013). When a smile becomes a fist: The perception of facial and bodily expressions of emotion in violent offenders. *Experimental Brain Research*, *228*(4), 399–410. https://doi.org/10.1007/s00221-013-3557-6

Kret, M. E., & Fischer, A. H. (2018). Recognition of facial expressions is moderated by Islamic cues. *Cognition and Emotion*, *32*(3), 623–631. https://doi.org/10.1080/02699931.2017.1330253

Kret, M. E., Jaasma, L., Bionda, T., & Wijnen, J. G. (2016). Bonobos (*Pan paniscus*) show an attentional bias toward conspecifics' emotions. *Proceedings of the National Academy of Sciences*, *113*(14), 3761–3766. https://doi.org/10.1073/pnas.1522060113

Kret, M. E., Muramatsu, A., & Matsuzawa, T. (2018). Emotion processing across and within species: A comparison between humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *132*(4), 395–409. https://doi.org/10.1037/com0000108

Kret, M. E., Prochazkova, E., Sterck, E. H. M. M., Clay, Z., & Kret, M. E. (2020). Emotional expressions in human and non-human great apes. *Neuroscience and Biobehavioral Reviews*, *115*, 378–395. https://doi.org/10.1016/j.neubiorev.2020.01.027

Kret, M. E., Roelofs, K., Stekelenburg, J. J., & De Gelder, B. (2013a). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, *7*(DEC), 1–9. https://doi.org/10.3389/fnhum.2013.00810

Kret, M. E., Stekelenburg, J. J., Roelofs, K., & De Gelder, B. (2013b). Perception of Face and Body Expressions Using Electromyography, Pupillometry and Gaze Measures. *Frontiers in Psychology*, *4*(February), 1–12. https://doi.org/10.3389/fpsyg.2013.00028

Kret, M. E., & Straffon, L. M. (2018). Reply to Crivelli et al.: The different faces of fear and threat. Evolutionary and cultural insights. *Journal of Human Evolution*, *125*, 193–197. https://doi.org/10.1016/j.jhevol.2017.11.006

Kret, M. E., & Tomonaga, M. (2016). Getting to the bottom of face processing. species-specific inversion effects for faces and behinds in humans and chimpanzees (*Pan troglodytes*). *PLoS ONE*, *11*(11), e0165357. https://doi.org/10.1371/journal.pone.0165357

Kret, M. E., & Van Berlo, E. (2021). Attentional bias in humans toward human and bonobo expressions of emotion. *Evolutionary Psychology*, *19*(3), 14747049211032816.

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, *74*(5), 569–586. https://doi.org/10.1037/amp0000364

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387–15392. https://doi.org/10.1073/pnas.251541498

Laan, E., & Everaerd, W. (1995). Determinants of female sexual arousal: Psychophysiological theory and data. *Annual Review of Sex Research*, *6*(1), 32–76. https://doi.org/10.1080/10532528.1995.10559901

Laan, E., Everaerd, W., Van Bellen, G., & Hanewald, G. (1994). Women's sexual and emotional responses to male- and female-produced erotica. *Archives of Sexual Behavior*, *23*(2), 153–169. https://doi.org/10.1007/BF01542096

Laan, E., & Janssen, E. (2007). How do men and women feel? Determinants of subjective experience of sexual arousal. *The Psychophysiology of Sex*, 278–290.

Lacreuse, A., Schatz, K., Strazzullo, S., King, H. M., & Ready, R. (2013). Attentional biases and memory for emotional stimuli in men and male rhesus monkeys. *Animal Cognition*, *16*(6), 861–871. https://doi.org/10.1007/s10071-013-0618-y

Laidre, M. E., & Johnstone, R. A. (2013). Animal signals. *Current Biology*, *23*(18), R829–R833. https://doi.org/10.1016/j.cub.2013.07.070

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

Laméris, D. W., Van Berlo, E., Roth, T. S., & Kret, M. E. (2021). No evidence for biased attention towards emotional scenes in Bornean orangutans (Pongo pygmaeus). *Manuscript Submitted for Publication*.

Laméris, D. W., Van Berlo, E., Sterck, E. H. M., Bionda, T., & Kret, M. E. (2020). Low relationship quality predicts scratch contagion during tense situations in orangutans (*Pongo pygmaeus*). *American Journal of Primatology*, *82*(7), e23138. https://doi.org/10.1002/ajp.23138

Laméris, D. W., Verspeek, J., Eens, M., & Stevens, J. M. G. (2022). Social and nonsocial stimuli alter the performance of bonobos during a pictorial emotional Stroop task. *American Journal of Primatology*, *84*(2), e23356. https://doi.org/10.1002/ajp.23356

Lane, K. A., Banaji, M. R., Nosek, B. a., & Greenwald, A. G. (2007). Understanding and Using the Implicit Association Test: IV What We Know (So Far) about the Method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit Measures of Attitudes* (pp. 59–102). The Guilford Press. https://doi.org/10.1037/a0015575

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 29–46). Oxford University Press. https://doi.org/10.1016/B978-84-9022-437-3.00003-9

Langergraber, K. E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., Fawcett, K., Inoue, E., Inoue-Muruyama, M., Mitani, J. C., Muller, M. N., Robbins, M. M., Schubert, G., Stoinski, T. S., Viola, B., Watts, D. P., Wittig, R. M., Wrangham, R. W., Zuberbühler, K., Pääbo, S., & 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*, *109*(39), 15716–15721. https://doi.org/10.1073/pnas.1211740109

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & Van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lazarus, R. S. (2001). Relational meaning and discrete emotions. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Series in affective science. Appraisal processes in emotion: Theory, methods, research.* (pp. 37–67). Oxford University Press.

LeDoux, J. E. (1998). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon and Schuster. https://doi.org/10.1176/ajp.155.4.570

LeDoux, J. E. (2021). What emotions might be like in other animals. *Current Biology*, *31*(13), R824–R829. https://doi.org/10.1016/J.CUB.2021.05.005

Lee, C., Cafe, L. M., Robinson, S. L., Doyle, R. E., Lea, J. M., Small, A. H., & Colditz, I. G. (2018). Anxiety influences attention bias but not flight speed and crush score in beef cattle. *Applied Animal Behaviour Science*, *205*, 210–215. https://doi.org/10.1016/J.APPLANIM.2017.11.003

Lee, M. D., & Wagenmakers, E.-J. (2013). Bayesian Cognitive Modeling. In *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lehmann, H. E. (1979). Yawning. A homeostatic reflex and its psychological significance. *Bulletin of the Menninger Clinic*, *43*(2), 123.

Leone, A., Ferrari, P. F., & Palagi, E. (2014). Different yawns, different functions? Testing social hypotheses on spontaneous yawning in *Theropithecus gelada*. *Scientific Reports*, *4*(1), 1–9. https://doi.org/10.1038/srep04010

Levine, E. E., Barasch, A., Rand, D., Berman, J. Z., & Small, D. A. (2018). Signaling emotion and reason in cooperation. *Journal of Experimental Psychology: General*, *147*(5), 702–719. https://doi.org/10.1037/xge0000399

Lewis, L. S., Kano, F., Stevens, J. M. G., DuBois, J. G., Call, J., & Krupenye, C. (2021). Bonobos and chimpanzees preferentially attend to familiar members of the dominant sex. *Animal Behaviour*, *177*, 193–206. https://doi.org/10.1016/j.anbehav.2021.04.027

Liebal, K., Waller, B. M., Slocombe, K. E., & Burrows, A. M. (2014). *Primate communication: a multimodal approach*. Cambridge University Press.

Liedtke, C., Kohl, W., Kret, M. E., & Koelkebeck, K. (2018). Emotion recognition from faces with in- and out-group features in patients with depression. *Journal of Affective Disorders*, *227*, 817–823. https://doi.org/10.1016/j.jad.2017.11.085

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*(August), 1–16. https://doi.org/10.3389/fpsyg.2015.01171

Luo, L., Reimert, I., De Haas, E. N., Kemp, B., & Bolhuis, J. E. (2019). Effects of early and later life environmental enrichment and personality on attention bias in pigs (Sus scrofa domesticus). *Animal Cognition*, *22*(6), 959–972. https://doi.org/10.1007/S10071-019-01287-W/FIGURES/5

Maclean, E. L. (2016). Unraveling the evolution of uniquely human cognition. *Proceedings of the National Academy of Sciences*, *113*(23), 6348–6354. https://doi.org/10.1073/pnas.1521270113

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional Bias in Emotional Disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15

Madsen, E. A., & Persson, T. (2013). Contagious yawning in domestic dog puppies (Canis lupus familiaris): The effect of ontogeny and emotional closeness on low-level imitation in dogs. *Animal Cognition*, *16*(2), 233–240. https://doi.org/10.1007/s10071-012-0568-9

Madsen, E. A., Persson, T., Sayehli, S., Lenninger, S., & Sonesson, G. (2013). Chimpanzees Show a Developmental Increase in Susceptibility to Contagious Yawning: A Test of the Effect of Ontogeny and Emotional Closeness on Yawn Contagion. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0076266

Maestripieri, D., Schino, G., Aureli, F., & Troisi, A. (1992). A modest proposal: displacement activities as an indicator of emotions in primates. *Animal Behaviour*, *44*(5), 967–979. https://doi.org/10.1016/S0003-3472(05)80592-5

Mafessoni, F., & Lachmann, M. (2019). The complexity of understanding others as the evolutionary origin of empathy and emotional contagion. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-41835-5

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019a). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in Psychology*, *10*, 2767. https://doi.org/10.3389/fpsyg.2019.02767

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019b). Understand and Describe Bayesian Models and Posterior Distributions using bayestestR. *CRAN*. https://doi.org/10.5281/zenodo.2556486

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019c). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Malavasi, R. (2014). Social modulation of yawning behavior in the domestic horse–an exploratory analysis. *Conference Paper: 48th International Conference of the International Society for Applied Ethology.*

Maréchal, L., Levy, X., Meints, K., & Majolo, B. (2017). Experience-based human perception of facial expressions in Barbary macaques (*Macaca sylvanus*). *PeerJ*, *5*, e3413. https://doi.org/10.7717/peerj.3413

Masataka, N., Koda, H., Atsumi, T., Satoh, M., & Lipp, O. V. (2018). Preferential attentional engagement drives attentional bias to snakes in Japanese macaques (*Macaca fuscata*) and humans (*Homo sapiens*). *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-36108-6

Massen, J. J. M., Church, A. M., & Gallup, A. C. (2015). Auditory Contagious Yawning in Humans: An Investigation into Affiliation and Status Effects. *Frontiers in Psychology*, *6*, 1–8. https://doi.org/10.3389/fpsyg.2015.01735

Massen, J. J. M., Dusch, K., Eldakar, O. T., & Gallup, A. C. (2014). A thermal window for yawning in humans: Yawning as a brain cooling mechanism. *Physiology and Behavior*, *130*, 145–148. https://doi.org/10.1016/j.physbeh.2014.03.032

Massen, J. J. M., & Gallup, A. C. (2017). Why contagious yawning does not (yet) equate to empathy. *Neuroscience and Biobehavioral Reviews*, *80*(July), 573–585.

Massen, J. J. M., Šlipogor, V., & Gallup, A. C. (2016). An observational investigation of behavioral contagion in common marmosets (Callithrix jacchus): Indications for contagious scent-marking. *Frontiers in Psychology*, *7*, 1190. https://doi.org/10.3389/fpsyg.2016.01190

Massen, J. J. M., Vermunt, D. A., & Sterck, E. H. M. (2012). Male yawning is more contagious than female yawning among chimpanzees (*Pan troglodytes*). *PLoS ONE*, *7*(7), 1–5. https://doi.org/10.1371/journal.pone.0040697

McEwen, B. S., & Sapolsky, R. M. (1995). Stress and cognitive function. *Current Opinion in Neurobiology*, *5*(2), 205–216. https://doi.org/10.1016/0959-4388(95)80028-X

McLellan, T., Johnston, L., Dalrymple-Alford, J., & Porter, R. (2010). Sensitivity to genuine versus posed emotion specified in facial displays. *Cognition and Emotion*, *24*(8), 1277–1292. https://doi.org/10.1080/02699930903306181

Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, *118*(3–4), 161–181. https://doi.org/10.1016/j.applanim.2009.02.023

Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1696), 2895–2904. https://doi.org/10.1098/rspb.2010.0303

Micheletta, J., Engelhardt, A., Matthews, L., Agil, M., & Waller, B. M. (2013). Multicomponent and Multimodal Lipsmacking in Crested Macaques (*Macaca nigra*). *American Journal of Primatology*, *75*(7), 763–773. https://doi.org/10.1002/AJP.22105

Mierke, J. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift Für Experimentelle Psychologie*, *48*(2), 107–122. https://doi.org/10.1026//0949-3946.48.2.107

Mierke, J., & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, *85*(6), 1180. https://doi.org/10.1037/0022-3514.85.6.1180

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the international affective picture system. *Behavior Research Methods*, *37*(4), 626–630. https://doi.org/10.3758/BF03192732

Millen, A., & Anderson, J. R. (2011). Neither infants nor toddlers catch yawns from their mothers. *Biology Letters*, *7*(3), 440–442. https://doi.org/10.1098/rsbl.2010.0966

Miller, M. L., Gallup, A. C., Vogel, A. R., & Clark, A. B. (2012a). Auditory disturbances promote temporal clustering of yawning and stretching in small groups of budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, *126*(3), 324. https://doi.org/10.1037/a0026520

Miller, M. L., Gallup, A. C., Vogel, A. R., Vicario, S. M., & Clark, A. B. (2012b). Evidence for contagious behaviors in budgerigars (Melopsittacus undulatus): An observational study of yawning and stretching. *Behavioural Processes*, *89*(3), 264–270. https://doi.org/10.1016/j.beproc.2011.12.012

Mitani, J. C., Grether, G. F., Rodman, P. S., & Priatna, D. (1991). Association among wild orang-utans: sociality, passive aggregations or chance? *Animal Behaviour*, *42*(1), 33–46. https://doi.org/10.1016/S0003-3472(05)80603-7

Mitani, J. C., Merriwether, D. A., & Zhang, C. (2000). Male affiliation, cooperation and kinship in wild chimpanzees. *Animal Behaviour*, *59*(4), 885–893. https://doi.org/10.1006/anbe.1999.1389

Mitra Setia, T., Delgado, R. A., Utami Atmoko, S. S., Singleton, I., & Van Schaik, C. P. (2009). Social organization and male-female relationships. In S. A. Wich, T. Utami Atmoko, T. Mitra Setia, & C. P. Schaik (Eds.), *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 245–254). Oxford University Press.

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

Mosher, D. L., & Maclan, P. (1994). College Men and Women Respond to X-rated Videos Intended for Male or Female Audiences: Gender and Sexual Scripts. *The Journal of Sex Research*, *31*(2), 99–113. https://doi.org/10.1080/00224499409551736

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication*, *52*(1), 1–22. https://doi.org/10.1080/10570318809389622

Moyaho, A., Rivas-Zamudio, X., Ugarte, A., Eguibar, J. R., & Valencia, J. (2015). Smell facilitates auditory contagious yawning in stranger rats. *Animal Cognition*, *18*(1), 279–290. https://doi.org/10.1007/s10071-014-0798-0

Nakayama, K. (2004). Observing Conspecifics Scratching Induces a Contagion of Scratching in Japanese Monkeys (*Macaca fuscata*). *Journal of Comparative Psychology*, *118*(1), 20–24. https://doi.org/10.1037/0735-7036.118.1.20

Neal, S. J., & Caine, N. G. (2016). Scratching under positive and negative arousal in common marmosets (Callithrix jacchus). *American Journal of Primatology*, *78*(2), 216–226. https://doi.org/10.1002/ajp.22498

Neilands, P., Claessens, S., Ren, I., Hassall, R., Bastos, A. P. M., & Taylor, A. H. (2020). Contagious yawning is not a signal of empathy: No evidence of familiarity, gender or prosociality biases in dogs. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1920), 20192236. https://doi.org/10.1098/rspb.2019.2236

Ngo, N., & Isaacowitz, D. M. (2015). Use of context in emotion perception: The role of top-down control, cue type, and perceiver's age. *Emotion*, *15*(3), 292–302. https://doi.org/10.1037/emo0000062

Nieuwburg, E. G. I., Ploeger, A., & Kret, M. E. (2021). Emotion recognition in nonhuman primates: How experimental research can contribute to a better understanding of underlying mechanisms. *Neuroscience and Biobehavioral Reviews*, *123*, 24–47. https://doi.org/10.1016/j.neubiorev.2020.11.029

Norscia, I., Coco, E., Robino, C., Chierto, E., & Cordoni, G. (2021). Yawn contagion in domestic pigs (*Sus scrofa*). *Scientific Reports*, *11*(1), 1851. https://doi.org/10.1038/s41598-020-80545-1

Norscia, I., Demuru, E., & Palagi, E. (2016a). She more than he: Gender bias supports the empathic nature of yawn contagion in Homo sapiens. *Royal Society Open Science*, *3*(2), 150459. https://doi.org/10.1098/rsos.150459

Norscia, I., Demuru, E., & Palagi, E. (2016b). Difference in contagious yawning between susceptible men and women: why not? *Royal Society Open Science*, *3*(9), 160477. https://doi.org/10.1098/rsos.160477

Norscia, I., & Palagi, E. (2011a). Yawn contagion and empathy in *Homo sapiens*. *PLoS ONE*, *6*(12), 1–5. https://doi.org/10.1371/journal.pone.0028472

Norscia, I., & Palagi, E. (2011b). When play is a family business: adult play, hierarchy, and possible stress reduction in common marmosets. *Primates*, *52*(2), 101–104. https://doi.org/10.1007/S10329-010-0228-0

Norscia, I., Zanoli, A., Gamba, M., & Palagi, E. (2020). Auditory Contagious Yawning Is Highest Between Friends and Family Members: Support to the Emotional Bias Hypothesis. *Frontiers in Psychology*, *11*, 442. https://doi.org/10.3389/fpsyg.2020.00442

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002a). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*(1), 101–115. https://doi.org/10.1037/1089-2699.6.1.101

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002b). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*(1), 44–59.

Nosek, B. A., Bar-Anan, Y., Sriram, N., Axt, J., & Greenwald, A. G. (2014). Understanding and using the brief Implicit Association Test: recommended scoring procedures. *PloS One*, *9*(12), e110938. https://doi.org/10.1371/journal.pone.0110938

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. *Personality and Social Psychology Bulletin*, *31*(2), 166–180. https://doi.org/10.1177/0146167204271418

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2013). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (pp. 265–292). Psychology Press. https://doi.org/10.4324/9780203783016-13

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/S11135-006-9018-6

O'Hara, S. J., & Reeve, A. V. (2011). A test of the yawning contagion and emotional connectedness hypothesis in dogs, Canis familiaris. *Animal Behaviour*, *81*(1), 335–340. https://doi.org/10.1016/j.anbehav.2010.11.005

Öhman, A., Carlsson, K., Lundqvist, D., & Ingvar, M. (2007). On the unconscious subcortical origin of human fear. *Physiology and Behavior*, *92*(1–2), 180–185. https://doi.org/10.1016/j.physbeh.2007.05.057

Öhman, A., Flykt, A., & Esteves, F. (2001a). Emotion drives attention: detecting the snake in the grass. *Journal of Experiemntal Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Öhman, A., Lundqvist, D., & Esteves, F. (2001b). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396. https://doi.org/10.1037/0022-3514.80.3.381

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology*, *108*(4), 562–571. https://doi.org/10.1037/pspa0000023

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., Bertini, M., Annicchiarico, G., & Cordoni, G. (2020a). Mirror replication of sexual facial expressions increases the success of sexual contacts in bonobos. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-75790-3

Palagi, E., Celeghin, A., Tamietto, M., Winkielman, P., & Norscia, I. (2020b). The neuroethology of spontaneous mimicry and emotional contagion in human and non-human animals. *Neuroscience and Biobehavioral Reviews*, *111*, 149–165. https://doi.org/10.1016/j.neubiorev.2020.01.020

Palagi, E., Leone, A., Mancini, G., & Ferrari, P. F. (2009). Contagious yawning in gelada baboons as a possible expression of empathy. *Proceedings of the National Academy of Sciences*, *106*(46), 19262–19267. https://doi.org/10.1073/pnas.0910891106

Palagi, E., & Norscia, I. (2011). Scratching around stress: Hierarchy and reconciliation make the difference in wild brown lemurs (Eulemur fulvus). *Stress*, *14*(1), 93–97. https://doi.org/10.3109/10253890.2010.505272

Palagi, E., & Norscia, I. (2013). Bonobos protect and console friends and kin. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079290

Palagi, E., & Norscia, I. (2019). The ethology of animal emotions: Investigation and interpretation. *Sistemi Intelligenti*, *31*(1), 11–32.

Palagi, E., Norscia, I., & Cordoni, G. (2019a). Lowland Gorillas (*Gorilla gorilla gorilla*) Failed to Respond to Others' Yawn: Experimental and Naturalistic Evidence. *Journal of Comparative Psychology*, *133*(3), 406. https://doi.org/10.1037/com0000175

Palagi, E., Norscia, I., & Demuru, E. (2014). Yawn contagion in humans and bonobos: Emotional affinity matters more than species. *PeerJ*, *2014*(1), 1–17. https://doi.org/10.7717/peerj.519

Palagi, E., Norscia, I., Pressi, S., & Cordoni, G. (2019b). Facial mimicry and play: A comparative study in chimpanzees and gorillas. *Emotion*, *19*(4), 665–681. https://doi.org/10.1037/emo0000476

Palagi, E., Paoli, T., & Tarli, S. B. (2004). Reconciliation and Consolation in Captive Bonobos (*Pan paniscus*). *American Journal of Primatology*, *62*(1), 15–30. https://doi.org/10.1002/ajp.20000

Palfai, T. P., Kantner, C. K., & Tahaney, K. D. (2016). The image-based alcohol-action implicit association test. *Journal of Behavior Therapy and Experimental Psychiatry*, *50*, 135–138. https://doi.org/10.1016/j.jbtep.2015.07.002

Panksepp, J. (2011). The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience and Biobehavioral Reviews*, *35*(9), 1791–1804. https://doi.org/10.1016/j.neubiorev.2011.08.003

Parr, L. A. (2001). Cognitive and physiological markers of emotional awareness in chimpanzees (*Pan troglodytes*). *Animal Cognition*, *4*(3–4), 223–229. https://doi.org/10.1007/s100710100085

Parr, L. A. (2004). Perceptual biases for multimodal cues in chimpanzee *Pan troglodytes*) affect recognition. *Animal Cognition*, *7*(3), 171–178. https://doi.org/10.1007/s10071-004-0207-1

Parr, L. A., & Hopkins, W. D. (2000). Brain temperature asymmetries and emotional perception in chimpanzees, *Pan troglodytes*. *Physiology and Behavior*, *71*(3–4), 363–371. https://doi.org/10.1016/S0031-9384(00)00349-8

Parr, L. A., Hopkins, W. D., & De Waal, 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., Modi, M., Siebert, E., & Young, L. J. (2013). Intranasal oxytocin selectively attenuates rhesus monkeys' attention to negative facial expressions. *Psychoneuroendocrinology*, *38*(9), 1748–1756. https://doi.org/10.1016/j.psyneuen.2013.02.011

Parr, L. A., Waller, B. M., & Heintz, M. (2008). Facial Expression Categorization by Chimpanzees Using Standardized Stimuli. *Emotion*, *8*(2), 216–231. https://doi.org/10.1037/1528-3542.8.2.216

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

Parr, L. A., Winslow, J. T., Hopkins, W. D., & De Waal, F. B. M. (2000). Recognizing facial cues: Individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, *114*(1), 47–60. https://doi.org/10.1037/0735-7036.114.1.47

Paukner, A., & Anderson, J. R. (2006). Video-induced yawning in stumptail macaques (*Macaca arctoides*). *Biology Letters*, *2*(1), 36–38. https://doi.org/10.1098/rsbl.2005.0411

Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience and Biobehavioral Reviews*, *29*(3), 469–491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Paul, E. S., & Mendl, M. T. (2018). Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. *Applied Animal Behaviour Science*, *205*(May 2017), 202–209. https://doi.org/10.1016/j.applanim.2018.01.008

Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion*, *35*(2), 192–201. https://doi.org/10.1007/s11031-011-9206-0

Pearson, B. L., Reeder, D. M., & Judge, P. G. (2015). Crowding increases salivary cortisol but not self-directed behavior in captive baboons. *American Journal of Primatology*, *77*(4), 462–467. https://doi.org/10.1002/ajp.22363

Peignot, P., Jankowski, F., & Anderson, J. R. (2004). On self-scratching in captive mandrills (*Mandrillus sphinx*). *Folia Primatologica*, *75*(3), 160–164. https://doi.org/10.1159/000078307

Perdue, B. M., Clay, A. W., Gaalema, D. E., Maple, T. L., & Stoinski, T. S. (2012). Technology at the zoo: the influence of a touchscreen computer on orangutans and zoo visitors. *Zoo Biology*. https://doi.org/10.1002/zoo.20378

Pérez-Manrique, A., & Gomila, A. (2022). Emotional contagion in nonhuman animals: A review. *WIREs Cognitive Science*, *13*(1), e1560. https://doi.org/10.1002/wcs.1560

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex*, *16*(3), 366–375. https://doi.org/10.1093/cercor/bhi115

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Petrova, K., Wentura, D., & Bermeitinger, C. (2013). What Happens during the Stimulus Onset Asynchrony in the Dot-Probe Task? Exploring the Role of Eye Movements in the Assessment of Attentional Biases. *PLoS ONE*, *8*(10), e76335. https://doi.org/10.1371/journal.pone.0076335

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, *48*(2), 175–187. https://doi.org/10.1016/j.neuron.2005.09.025

Pietraszewski, D. (2021). The correct way to test the hypothesis that racial categorization is a byproduct of an evolved alliance-tracking capacity. *Scientific Reports*, *11*(1), 3404. https://doi.org/10.1038/s41598-021-82975-x

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., & Kastner, S. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology*, *101*(5), 2581–2600. https://doi.org/10.1152/jn.91198.2008

Platek, S. M., Critton, S. R., Myers, T. E., & Gallup, G. G. (2003). Contagious yawning: The role of self-awareness and mental state attribution. *Cognitive Brain Research*, *17*(2), 223–227. https://doi.org/10.1016/S0926-6410(03)00109-5

Platek, S. M., Mohamed, F. B., & Gallup, G. G. (2005). Contagious yawning and the brain. *Cognitive Brain Research*, *23*(2–3), 448–452. https://doi.org/10.1016/j.cogbrainres.2004.11.011

Plavcan, J. M. (2001). Sexual dimorphism in primate evolution. *Yearbook of Physical Anthropology*, *116*(S33), 25–53. https://doi.org/10.1002/ajpa.10011

Pokorny, J. J., & De Waal, F. B. M. (2009). Monkeys recognize the faces of group mates in photographs. *Proceedings of the National Academy of Sciences*, *106*(51), 21539–21543. https://doi.org/10.1073/pnas.0912174106

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional Bias for Positive Emotional Stimuli: A Meta-Analytic Investigation. *Psychological Bulletin*, *142*(1), 79–106. https://doi.org/10.1037/bul0000026

Preston, S. D., & De Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

Preuschoft, S., & van Hooff, J. A. R. A. M. (1995). Homologizing Primate Facial Displays: A Critical Review of Methods. *Folia Primatologica*, *65*(3), 121–137. https://doi.org/10.1159/000156878

Pritsch, C., Telkemeyer, S., Mühlenbeck, C., & Liebal, K. (2017). Perception of facial expressions reveals selective affect-biased attention in humans and orangutans. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-07563-4

Prochazkova, E., & Kret, M. E. (2017). Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neuroscience and Biobehavioral Reviews*, *80*, 99–114. https://doi.org/10.1016/j.neubiorev.2017.05.013

Provine, R. R. (1986). Yawning as a Stereotyped Action Pattern and Releasing Stimulus. *Ethology*, *72*(2), 109–122. https://doi.org/10.1111/j.1439-0310.1986.tb00611.x

Provine, R. R., & Hamernik, H. B. (1986). Yawning: Effects of stimulus interest. *Bulletin of the Psychonomic Society*, *24*(6), 437–438. https://doi.org/10.3758/BF03330574

Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit Racial Biases in Preschool Children and Adults From Asia and Africa. *Child Development*. https://doi.org/10.1111/cdev.12442

R Core Team. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. https://www.r-project.org/

Ramon, M., & Gobbini, M. I. (2018). Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition*, *26*(3), 179–195. https://doi.org/10.1080/13506285.2017.1405134

Rauchbauer, B., Majdandžic, J., Stieger, S., & Lamm, C. (2016). The modulation of mimicry by ethnic group-membership and emotional expressions. *PLoS ONE*, *11*(8), e0161064. https://doi.org/10.1371/journal.pone.0161064

Reddy, R. B., Krupenye, C., MacLean, E. L., & Hare, B. (2016). No evidence for contagious yawning in lemurs. *Animal Cognition*, *19*(5), 889–898. https://doi.org/10.1007/s10071-016-0986-1

Richetin, J., Costantini, G., Perugini, M., & Schönbrodt, F. (2015). Should We Stop Looking for a Better Scoring Algorithm for Handling Implicit Association Test Data? Test of the Role of Errors, Extreme Latencies Treatment, Scoring Formula, and Practice Trials on Reliability and Validity. *PLoS ONE*, *10*(6), e0129601. https://doi.org/10.1371/journal.pone.0129601

Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social Cognitive and Affective Neuroscience*, *3*(3), 270–278. https://doi.org/10.1093/scan/nsn021

Righart, R., & De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective and Behavioral Neuroscience*, *8*(3), 264–272. https://doi.org/10.3758/CABN.8.3.264

Rilling, J. K., Scholz, J., Preuss, T. M., Glasser, M. F., Errangi, B. K., & Behrens, T. E. (2012). Differences between chimpanzees and bonobos in neural systems supporting social cognition. *Social Cognitive and Affective Neuroscience*, *7*(4), 369–379. https://doi.org/10.1093/scan/nsr017

Robbins, M. M., Bermejo, M., Cipolletta, C., Magliocca, F., Parnell, R. J., & Stokes, E. (2004). Social structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*). *American Journal of Primatology*, *64*(2), 145–159. https://doi.org/10.1002/AJP.20069

Romero, T., Ito, M., Saito, A., & Hasegawa, T. (2014). Social Modulation of Contagious Yawning in Wolves. *PLoS ONE*, *9*(8), e105963. https://doi.org/10.1371/journal.pone.0105963

Romero, T., Konno, A., & Hasegawa, T. (2013). Familiarity Bias and Physiological Responses in Contagious Yawning by Dogs Support Link to Empathy. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071365

Roth, T. S., Rianti, P., Fredriksson, G. M., Wich, S. A., & Nowak, M. G. (2020). Grouping behavior of Sumatran orangutans (*Pongo abelii*) and Tapanuli orangutans (*Pongo tapanuliensis*) living in forest with low fruit abundance. *American Journal of Primatology*, *82*(5), e23123. https://doi.org/10.1002/ajp.23123

Roth, T. S., Samara, I., Tan, J., Prochazkova, E., & Kret, M. E. (2021). A comparative framework of inter-individual coordination and pair-bonding. In *Current Opinion in Behavioral Sciences* (Vol. 39, pp. 98–105). Elsevier. https://doi.org/10.1016/j.cobeha.2021.03.005

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/H0077714

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. https://doi.org/10.1037/0022-3514.76.5.805

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development*. https://doi.org/10.1111/j.1467-8624.2005.00856.x

Saito, Y., Yuki, S., Seki, Y., Kagawa, H., & Okanoya, K. (2016). Cognitive bias in rats evoked by ultrasonic vocalizations suggests emotional contagion. *Behavioural Processes*, *132*, 5–11. https://doi.org/10.1016/j.beproc.2016.08.005

Schino, G., Carducci, P., & Truppa, V. (2020). Attention to social stimuli is modulated by sex and exposure time in tufted capuchin monkeys. *Animal Behaviour*, *161*, 39–47. https://doi.org/10.1016/j.anbehav.2019.12.019

Schino, G., Perretta, G., Taglioni, A. M., Monaco, V., & Troisi, A. (1996). Primate displacement activities as an ethopharmacological model of anxiety. *Anxiety*, *2*, 186–191.

Schino, G., Troisi, A., Perretta, G., & Monaco, V. (1991). Measuring anxiety in nonhuman primates: Effect of lorazepam on macaque scratching. *Pharmacology, Biochemistry and Behavior*, *38*(4), 889–891. https://doi.org/10.1016/0091-3057(91)90258-4

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. In *Trends in Cognitive Sciences* (Vol. 21, Issue 3, pp. 216–228). Elsevier Current Trends. https://doi.org/10.1016/j.tics.2017.01.001

Schroepfer-Walker, K., Wobber, V., & Hare, B. (2015). Experimental evidence that grooming and play are social currency in bonobos and chimpanzees. In B. Hare & S. Yamamoto (Eds.), *Bonobo Cognition and Behaviour* (pp. 299–316). Brill.

Schut, C., Grossman, S., Gieler, U., Kupfer, J., & Yosipovitch, G. (2015). Contagious itch: what we know and what we would like to know. *Frontiers in Human Neuroscience*, *9*, 57. https://doi.org/10.3389/fnhum.2015.00057

Schwartz, J. H. (1987). *The red ape: Orangutans and human origins*. Elm Tree Books, London.

Senju, A., Maeda, M., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2007). Absence of contagious yawning in children with autism spectrum disorder. *Biology Letters*. https://doi.org/10.1098/rsbl.2007.0337

Setia, T. M., & van Schaik, C. P. (2007). The Response of Adult Orang-Utans to Flanged Male Long Calls: Inferences about Their Function. *Folia Primatologica*, *78*(4), 215–226. https://doi.org/10.1159/000102317

Setoh, P., Lee, K. J. J., Zhang, L., Qian, M. K., Quinn, P. C., Heyman, G. D., & Lee, K. (2019). Racial Categorization Predicts Implicit Racial Bias in Preschool Children. *Child Development*, *90*(1), 162–179. https://doi.org/10.1111/cdev.12851

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., Leibenluft, E., & Pine, D. S. (2012). Attention biases, anxiety, and development: Toward or away from threats or rewards? *Depression and Anxiety*, *29*(4), 282–294. https://doi.org/10.1002/da.20914

Signe, P., & Van Hooff, J. A. R. A. M. (2018). The Social Function of "Smile" and "Laughter": Variations Across Primate Species and Societies 1. In U. C. Segerstråle & P. Molnár (Eds.), *Nonverbal Communication* (pp. 171–190). Lawrence Erlbaum Associates, Inc. https://doi.org/10.4324/9781351243131-12

Silk, J. B., Alberts, S. C., & Altmann, J. (2004). Patterns of coalition formation by adult female baboons in Amboseli, Kenya. *Animal Behaviour*, *67*(3), 573–582. https://doi.org/10.1016/j.anbehav.2003.07.001

Silk, J. B., Altmann, J., & Alberts, S. C. (2006). Social relationships among adult female baboons (papio cynocephalus) I. Variation in the strength of social bonds. *Behavioral Ecology and Sociobiology 2006 61:2*, *61*(2), 183–195. https://doi.org/10.1007/S00265-006-0249-2

Silva, K., Bessa, J., & De Sousa, L. (2012). Auditory contagious yawning in domestic dogs (Canis familiaris): First evidence for social modulation. *Animal Cognition*, *15*(4), 721–724. https://doi.org/10.1007/s10071-012-0473-2

Singleton, I., Knott, C. D., Morrogh-Bernard, H. C., Wich, S. A., & Van Schaik, C. P. (2009). Ranging behavior of orangutan females and social organization. In S. A. Wich, S. S. Utami Atmoko, T. Mitra Setia, & C. P. van Schaik (Eds.), *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 205–2014). Oxford University Press.

Singleton, I., & Van Schaik, C. P. (2002). The social organisation of a population of Sumatran orang-utans. *Folia Primatologica*, *73*(1), 1–20. https://doi.org/10.1159/000060415

Slabbinck, H., De Houwer, J., & Van Kenhove, P. (2011). A pictorial Attitude IAT as a Measure of Implicit Motives. *European Journal of Personality*, *25*(1), 76–86. https://doi.org/10.1002/per.778

Steckenfinger, S. A., & Ghazanfar, A. A. (2009). Monkey visual behavior falls into the uncanny valley. *Proceedings of the National Academy of Sciences*, *106*(43), 18362–18366. https://doi.org/10.1073/pnas.0910063106

Stevens, J. (2020). *EAZA Best Practice Guidelines for Bonobos (*Pan paniscus*)*.

Stimpson, C. D., Barger, N., Taglialatela, J. P., Gendron-Fitzpatrick, A., Hof, P. R., Hopkins, W. D., & Sherwood, C. C. (2016). Differential serotonergic innervation of the amygdala in bonobos and chimpanzees. *Social Cognitive and Affective Neuroscience*, *11*(3), 413–422. https://doi.org/10.1093/scan/nsv128

Stone, A. A., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (2000). *The science of self-report*. Psychology Press.

Swithenbank, S., Cowdell, F., & Holle, H. (2016). The Role of Auditory Itch Contagion in Psoriasis. *Acta Dermato-Venereologica*, *96*(6), 728–731. https://doi.org/10.2340/00015555-2320

Tajima, T., & Kurotori, H. (2010). Nonaggressive interventions by third parties in conflicts among captive Bornean orangutans (*Pongo pygmaeus*). *Primates*, *51*(2), 179–182. https://doi.org/10.1007/s10329-009-0180-z

Talbot, C. F., Mayo, L., Stoinski, T., & Brosnan, S. F. (2015). Face discriminations by orangutans (*Pongo spp.*) vary as a function of familiarity. *Evolutionary Psychological Science*, *1*(3), 172–182. https://doi.org/10.1007/s40806-015-0019-3

Tan, J., Ariely, D., & Hare, B. (2017). Bonobos respond prosocially toward members of other groups. *Scientific Reports*, *7*(1), 1–11.

Tan, J., & Hare, B. (2013). Bonobos Share with Strangers. *PLoS ONE*, *8*(1), e51922. https://doi.org/10.1371/journal.pone.0051922

Tan, J., & Hare, B. (2017). Prosociality among non-kin in bonobos and chimpanzees compared. In B. Hare & S. Yamamoto (Eds.), *Bonobos: Unique in Mind, Brain, and Behavior* (pp. 140–154). Oxford University Press.

Te Broekhorst, I. J. A., Schürmann, C. L., & Sugardjito, J. (1990). Residential status and seasonal movements of wild orang-utans in the Gunung Leuser Reserve (Sumatera, Indonesia). *Animal Behaviour*, *39*(6), 1098–1109. https://doi.org/10.1016/S0003-3472(05)80782-1

Thomas, S., Smith, R. B., & Ball, P. (2007). Implicit attitudes in very young children: An adaptation of the IAT. *Current Research in Social Psychology*, *13*(7), 75–85.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, *20*(4), 410–433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends in Cognitive Sciences*, *16*(7), 365–372. https://doi.org/10.1016/j.tics.2012.06.003

Tokuyama, N., Toda, K., Poiret, M.-L., Iyokango, B., Bakaa, B., & Ishizuka, S. (2021). Two wild female bonobos adopted infants from a different social group at Wamba. *Scientific Reports*, *11*, 4967. https://doi.org/10.1038/s41598-021-83667-2

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), 673–692. https://doi.org/10.1086/668207

Tottenham, N., Phuong, J., Flannery, J., Gabard-Durnam, L., & Goff, B. (2013). A negativity bias for ambiguous facial-expression valence during childhood: Converging evidence from behavior and facial corrugator muscle responses. *Emotion*, *13*(1), 92–103. https://doi.org/10.1037/a0029431

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Townshend, J. M., & Duka, T. (2001). Attentional bias associated with alcohol cues: Differences between heavy and occasional social drinkers. *Psychopharmacology*, *157*(1), 67–74. https://doi.org/10.1007/s002130100764

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. In *Current Opinion in Neurobiology* (Vol. 13, Issue 4, pp. 428–432). Elsevier Ltd. https://doi.org/10.1016/S0959-4388(03)00105-3

Troisi, A. (1999). Ethological research in clinical psychiatry: The study of nonverbal behavior during interviews. *Neuroscience and Biobehavioral Reviews*, *23*(7), 905–913. https://doi.org/10.1016/S0149-7634(99)00024-X

Troisi, A. (2002). Displacement activities as a behavioral measure of stress in nonhuman primates and human subjects. *Stress*, *5*(1), 47–54. https://doi.org/10.1080/102538902900012378

Troisi, A., Schino, G., D'Antoni, M., Pandolfi, N., Aureli, F., & D'Amato, F. R. (1991). Scratching as a behavioral index of anxiety in macaque mothers. *Behavioral and Neural Biology*, *56*(3), 307–313. https://doi.org/10.1016/0163-1047(91)90469-7

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, *105*(49), 19514–19519. https://doi.org/10.1073/pnas.0809662105

Usui, S., Senju, A., Kikuchi, Y., Akechi, H., Tojo, Y., Osanai, H., & Hasegawa, T. (2013). Presence of Contagious Yawning in Children with Autism Spectrum Disorder. *Autism Research and Treatment*. https://doi.org/10.1155/2013/971686

Van Berlo, E., Bionda, T., & Kret, M. E. (2020a). Attention Towards Emotions is Modulated by Familiarity with the Expressor. A Comparison Between Bonobos and Humans. *BioRxiv*. https://doi.org/10.1101/2020.05.11.089813

Van Berlo, E., Díaz-Loyo, A. P., Juárez-Mora, O. E. O. E., Kret, M. E., & Massen, J. J. M. M. (2020b). Experimental evidence for yawn contagion in orangutans (*Pongo pygmaeus*). *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-79160-x

Van den stock, J., Vandenbulcke, M., Sinke, C. B. A., Goebel, R., & De Gelder, B. (2014). How affective information from faces and scenes interacts in the brain. *Social Cognitive and Affective Neuroscience*, *9*(10), 1481–1488. https://doi.org/10.1093/scan/nst138

Van Hooff, J. A. R. A. M. (1972). A comparative approach to the phylogeny of laughter and smiling. In R. A. Hinde (Ed.), *Non-verbal communication*. Cambridge University Press.

Van Noordwijk, M. A., Arora, N., Willems, E. P., Dunkel, L. P., Amda, R. N., Mardianah, N., Ackermann, C., Krützen, M., & van Schaik, C. P. (2012). Female philopatry and its social benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology*, *66*(6), 823–834. https://doi.org/10.1007/s00265-012-1330-7

Van Renswoude, D. R., Visser, I., Raijmakers, M. E. J., Tsang, T., & Johnson, S. P. (2019). Real-world scene perception in infants: What factors guide attention allocation? *Infancy*, *24*(5), 693–717. https://doi.org/10.1111/infa.12308

Van Rooijen, R., Ploeger, A., & Kret, M. E. (2017). The dot-probe task to measure emotional attention: A suitable measure in comparative studies? *Psychonomic Bulletin and Review*, *24*(6), 1686–1717. https://doi.org/10.3758/s13423-016-1224-1

Van Schaik, C. P. (1999). The socioecology of fission-fusion sociality in Orangutans. *Primates*, *40*(1), 69–86. https://doi.org/10.1007/BF02557703

Van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B. M. F., Knott, C. D., Singleton, I., Suzuki, A., Suci Utami, S., & Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*, *299*(5603), 102–105. https://doi.org/10.1126/science.1078004

Vick, S. J., & Paukner, A. (2010). Variation and context of yawns in captive chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, *72*(3), 262–269. https://doi.org/10.1002/ajp.20781

Victeur, Q., Huguet, P., & Silvert, L. (2020). Attentional allocation to task-irrelevant fearful faces is not automatic: experimental evidence for the conditional hypothesis of emotional selection. *Cognition and Emotion*, *34*(2), 288–301. https://doi.org/10.1080/02699931.2019.1622512

Vonk, J., & Hamilton, J. (2014). Orangutans (*Pongo abelii*) and a gorilla (*Gorilla gorilla gorilla*) match features in familiar and unfamiliar individuals. *Animal Cognition*, *17*(5), 1089–1105.

Vrooman, C., Gijsberts, M., & Boelhouwer, J. (2014). *Verschil in Nederland*. https://www.narcis.nl/publication/RecordID/oai:scp.nl:f6a4f47b-7261-4bd1-94b9-11719e3bff1b

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Waller, B. M., Julle-Daniere, E., & Micheletta, J. (2020). Measuring the evolution of facial 'expression' using multi-species FACS. In *Neuroscience and Biobehavioral Reviews* (Vol. 113, pp. 1–11). Pergamon. https://doi.org/10.1016/j.neubiorev.2020.02.031

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, *7*(6), 177–188. https://doi.org/10.1111/1467-8721.ep10836912

Whitehouse, J., Micheletta, J., Kaminski, J., & Waller, B. M. (2016). Macaques attend to scratching in others. *Animal Behaviour*, *122*, 169–175. https://doi.org/10.1016/j.anbehav.2016.10.020

Whitehouse, J., Micheletta, J., & Waller, B. M. (2017). Stress behaviours buffer macaques from aggression. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-10754-8

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. In *Emotion Review* (Vol. 5, Issue 1, pp. 72–77). SAGE PublicationsSage UK: London, England. https://doi.org/10.1177/1754073912451492

Wilkinson, A., Sebanz, N., Mandl, I., & Huber, L. (2011). No evidence of contagious yawning in the red-footed tortoise Geochelone carbonaria. *Current Zoology*, *57*(4), 477–484. https://doi.org/10.1093/czoolo/57.4.477

Williams, L. A., Brosnan, S. F., & Clay, Z. (2020). Anthropomorphism in comparative affective science: Advocating a mindful approach. *Neuroscience and Biobehavioral Reviews*, *115*, 299–307.

Wilson, D. A., & Tomonaga, M. (2018). Exploring attentional bias towards threatening faces in chimpanzees using the dot probe task. *PLoS ONE*, *13*(11), 1–17. https://doi.org/10.1371/journal.pone.0207378

Wilson, M. L. (2021). Insights into human evolution from 60 years of research on chimpanzees at Gombe. *Evolutionary Human Sciences*, *3*. https://doi.org/10.1017/EHS.2021.2

Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., Hobaiter, C. L., Hohmann, G., Itoh, N., Koops, K., Lloyd, J. N., Matsuzawa, T., Mitani, J. C., Mjungu, D. C., Morgan, D., Muller, M. N., Mundry, R., Nakamura, M., Pruetz, J., … Wrangham, R. W. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature*, *513*(7518), 414–417. https://doi.org/10.1038/nature13727

Wilson, M. L., & Wrangham, R. W. (2003). Intergroup relations in chimpanzees. *Annual Review of Anthropology*, *32*, 363–392. https://doi.org/10.1146/annurev.anthro.32.061002.120046

Wilson, V. A. D., Kade, C., Moeller, S., Treue, S., Kagan, I., & Fischer, J. (2020). Macaque Gaze Responses to the Primatar: A Virtual Macaque Head for Social Cognition Research. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01645

Wirth, B. E., & Wentura, D. (2019). Attentional bias towards angry faces is moderated by the activation of a social processing mode in the general population. *Cognition and Emotion*, *33*(7), 1317–1329. https://doi.org/10.1080/02699931.2018.1561423

Wojczulanis-Jakubas, K., Plenzler, J., & Jakubas, D. (2019). Indications of contagious behaviours in the southern elephant seal: An observational study. *Behaviour*, *156*(1), 59–77. https://doi.org/10.1163/1568539X-00003530

Wrangham, R. W. (1999). Evolution of coalitionary killing. *American Journal of Physical Anthropology*, *110*(S29), 1–30. https://doi.org/10.1002/(sici)1096-8644(1999)110:29+<1::aid-ajpa2>3.3.co;2-5

Wrangham, R. W., & Glowacki, L. (2012). Intergroup aggression in chimpanzees and war in nomadic hunter-gatherers. *Human Nature*, *23*(1), 5–29. https://doi.org/10.1007/s12110-012-9132-1

Yonezawa, T., Sato, K., Uchida, M., Matsuki, N., & Yamazaki, A. (2017). Presence of contagious yawning in sheep. *Animal Science Journal*, *88*(1), 195–200. https://doi.org/10.1111/asj.12681

Yoon, J. M. D., & Tennie, C. (2010). Contagious yawning: a reflection of empathy, mimicry, or contagion? *Animal Behaviour*, *79*(5), 2007–2009. https://doi.org/10.1016/j.anbehav.2010.02.011

Zentall, T. R. (2003). Imitation by Animals: How Do They Do It? *Current Directions in Psychological Science*, *12*(3), 91–95. https://doi.org/10.1111/1467-8721.01237

Zucker, E. L. (1987). Control of intragroup aggression by a captive male orangutan. *Zoo Biology*, *6*(3), 219–223. https://doi.org/10.1002/zoo.1430060304

Zvielli, A., Bernstein, A., & Koster, E. H. W. (2014). Dynamics of attentional bias to threat in anxious adults: Bias towards and/or away? *PLoS ONE*, *9*(8), e104025. https://doi.org/10.1371/journal.pone.0104025

Zych, A. D., & Gogolla, N. (2021). Expressions of emotions across species. *Current Opinion in Neurobiology*, *68*, 57–66. https://doi.org/10.1016/J.CONB.2021.01.003

Samenvatting

Samenvatting

Emoties gezien door de ogen van onze 'naaste verwanten': Een onderzoek naar aandachts- en gedragsmechanismen

Emoties vormen de toegangspoort tot onze innerlijke wereld en zijn fundamenteel voor het (voort)bestaan van onze soort. Van een huilende baby die de ouders activeert om hem of haar gerust te stellen tot aan een grote menigte die juicht bij de opkomst van hun favoriete band op het podium; het zijn onze emoties die ons met elkaar verbinden. Wij mensen zijn echter niet de enige diersoort die emoties kunnen ervaren en uitdrukken. Inmiddels weten we dat emoties diep geworteld zitten in de evolutie van de mens en dat er waarschijnlijk veel andere diersoorten zijn die emoties op een of andere wijze kunnen ervaren en uitdrukken. Met de komst van verbeterde observationele methodes en ultramoderne technologie kunnen we nieuwe inzichten krijgen in de mentale wereld van zowel mensen als dieren.

Om dit te bereiken is het een eerste logische stap om te kijken naar onze naaste verwanten: de mensapen. De familie van mensapen bestaat uit de kleine mensapen (gibbons) en de grote mensapen (ook wel *great apes* genoemd in het Engels: orang-oetans, gorilla's, chimpansees, bonobo's en mensen). Van de grote mensapensoorten is de orang-oetan onze verste neef (Figuur 2 in de introductie). Met hem delen we een gemeenschappelijke apen-voorouder die zo'n 14 miljoen jaar geleden leefde. Later ontstond een aftakking die leidde tot de gorilla's. De apen die het dichtst bij ons staan in de evolutionaire boom zijn de chimpansee en de bonobo. Met hen delen we de meest 'recente' voorouder, die zo'n 7 tot 8 miljoen jaar geleden leefde. Als we nu overeenkomstige eigenschappen vinden in deze dieren dan is het goed mogelijk dat onze gemeenschappelijke voorouder deze eigenschappen ook al bezat. Door mensapen te bestuderen kunnen we meer inzicht krijgen in onze eigen evolutie. Dit wordt echter overschaduwd door stroperij en ontbossing; die hebben er voor gezorgd dat mensapen met uitsterven bedreigd worden. Ik vind het daarom heel belangrijk dat er meer onderzoek gedaan wordt naar de cognitieve en emotionele capaciteiten van mensapen voordat deze mogelijkheid wellicht voor altijd verloren gaat.

In dit proefschrift richt ik mij op drie elementen die een rol spelen in het waarnemen, verwerken, en uiteindelijk begrijpen van emoties: *aandacht*, *spontane mimicry* (oftewel het spontaan nabootsen van anderen) en *impliciete associaties*. Op hun betekenis en hun rol in emoties kom ik later terug. Ik onderzoek de drie elementen in een aantal studies om zo een vergelijking te kunnen maken tussen mensen, bonobo’s en orang-oetans. De bonobo is een zeldzame apensoort die in groepen leeft. Hoewel bonobo's sociale banden vormen met hun groepsgenoten, associëren ze zich ook graag ook met onbekende soortgenoten. Dit is uniek onder de mensapen en wordt ook wel xenofilie genoemd, oftewel het aangetrokken voelen tot vreemden. Recentelijk onderzoek heeft laten zien dat bonobo’s sterk ontwikkelde breinstructuren hebben die een rol spelen in het verwerken van emotionele prikkels en reguleren van emoties. Deze eigenschappen vallen op in vergelijking met andere mensapen en daarom zijn bonobo’s interessant voor vergelijkend onderzoek naar emotieperceptie. Orang-oetans, ook wel ‘*man of the forest*’ genoemd hebben een unieke sociale structuur in vergelijking met de andere mensapen. Zij zijn semi-solitair, dit betekent, dat ze hun dagen voornamelijk alleen doorbrengen en af en toe samen komen om bijvoorbeeld te paren of om hun jongen met elkaar te laten spelen. Op dit moment weten we zeer weinig over de emotionele capaciteiten van orang-oetans, maar hun unieke leefstijl maakt hun een interessante soort om meer inzicht te krijgen in de evolutie van hun emotieperceptie.

Allereerst heb ik de link onderzocht tussen aandacht en emoties in mensapen. Aandacht fungeert als een poortwachter die bepaalt welke informatie onze hersenen zullen prioriteren en welke informatie niet belangrijk is. Dit selectieve proces is nodig omdat onze hersenen niet álles tegelijkertijd kunnen verwerken. Uit onderzoek blijkt dat onze aandacht onmiddellijk uit gaat naar biologisch relevante signalen, bijvoorbeeld emotionele expressies. Dit gebeurt buiten ons bewustzijn en het stelt ons er toe in staat om in een razendsnel tempo hoofd- en bijzaken te onderscheiden om zo onze overleving te waarborgen. Een kernvraag hierbij is: delen wij de aandachtsmechanismen voor emotieperceptie met andere dieren zoals bonobo’s?

De bevindingen in dit proefschrift laten zien dat de aandacht van mensen en bonobo’s onmiddellijk uit gaat naar emotionele prikkels van soortgenoten. Emotionele prikkels zijn bijvoorbeeld het zien van spelende soortgenoten, of soortgenoten die erg bang zijn. Bij bonobo’s is dit voornamelijk het geval voor emotionele prikkels van onbekenden, terwijl mensen gevoeliger zijn voor emotionele prikkels van bekenden zoals familie en vrienden (**hoofdstuk 2**). Daarnaast zien we, dat mensen foto’s van emotionele scènes van bonobo’s en mensen vergelijkbaar beoordelen op basis van valentie (positiviteit en negativiteit) én intensiteit van de getoonde emotie. Ook trekken emotionele scènes van bonobo’s sneller de aandacht dan neutrale scenes, net als emotionele mensenscènes (**hoofdstuk 3**). Tot slot hebben we door middel van *eye-tracking* (een techniek om nauwkeurig te kunnen meten waar individuen naar kijken) vast kunnen stellen dat specifieke emotionele scènes langer de aandacht vasthouden dan neutrale scènes. Mensen kijken over het algemeen langer naar emotionele scènes van andere mensen, maar óók langer naar scènes waarin bonobo's vlooien of spelen. Bonobo's kijken het langst naar scènes met gestreste of bange soortgenoten, of scènes met seks (**hoofdstuk 4**). Belangrijk is dat deze resultaten laten zien dat de aandachtsmechanismen die ten grondslag liggen aan emotieperceptie gedeeld worden met onze naaste verwanten, de bonobo's. Daarnaast zien we, dat sociale kenmerken van beide soorten een rol spelen in hóé aandacht uit gaat naar emoties. Bonobo's zijn van nature erg geïnteresseerd in onbekende soortgenoten, hetgeen een mogelijke verklaring is voor het resultaat dat hun aandacht vooral getrokken wordt door emotionele prikkels van onbekenden. Mensen zijn gevoelig voor de emoties van andere mensen, maar focussen zich over het algemeen op bescherming van hun eigen groep, waaronder familie en vrienden. Bij mensen leidt deze tendens waarschijnlijk tot onmiddellijke aandacht voor emoties van bekenden.

Naast aandacht is er een rol weggelegd voor spontane mimicry in het verwerken van emotionele prikkels. In de psychologie wordt mimicry beschreven als het automatisch en onbewust kopiëren of imiteren van gezichtsexpressies en gedragingen (bijvoorbeeld lachen en gapen). Mimicry faciliteert daarmee het waarnemen van en communiceren over emoties door middel van emotionele aanstekelijkheid. Hierbij kan gedacht worden aan het meevoelen van verdriet als je iemand ziet huilen. In de laatste 10 jaar wordt er veel onderzoek gedaan naar de mimicry van gapen, oftewel de aanstekelijkheid van gapen. Het is onduidelijk welke emotie ten grondslag ligt aan gapen, maar wel blijkt de aanstekelijkheid van gapen sterker te zijn tussen individuen met een goede sociale band. Omdat gapen aanstekelijk is, zou het een indicator kunnen zijn voor emotionele aanstekelijkheid, oftewel het voelen wat een ander voelt. Er is echter nog veel kritiek op dit idee, omdat de link tussen emotionele aanstekelijkheid en de aanstekelijkheid van gapen nooit direct is aangetoond. Een andere variant van mimicry de aanstekelijkheid van zelfkrabben. Als individuen zichzelf krabben, dan is dat een indicator van stress en uit een kleine hoeveelheid eerdere studies blijkt dat zelfkrabben mogelijk aanstekelijk is.

In een tweetal onderzoeken heb ik onderzocht of de aanstekelijkheid van krabben (**hoofdstuk 5**) en gapen (**hoofdstuk 6**) voorkomen bij orang-oetans, omdat deze twee fenomenen tot dusver alleen aangetroffen zijn in sociale dieren die in groepen leven, bijvoorbeeld in mensen en bonobo's. De resultaten tonen aan dat orang-oetans inderdaad meer aan zelfkrabben doen als ze een ander individu zien krabben. De aanstekelijkheid van zelfkrabben was het sterkst tussen individuen met een minder sterke sociale band, wat op het bestaan van *negatieve* emotionele aanstekelijkheid kan duiden. In een vervolgstudie met gaapvideo's hebben ik laten zien dat orang-oetans ook gevoelig zijn voor de aanstekelijkheid van gapen, iets dat nog niet eerder is aangetoond. Opvallend hierbij is dat er geen verschil was in of orang-oetans een bekende of een onbekende soortgenoot zagen gapen. Dit contrasteert met eerdere bevindingen die laten zien, dat de kans op aanstekelijk gapen toeneemt naarmate de relatie tussen de individuen sterker is. Over het algemeen tonen de resultaten aan dat orang-oetans mimicry vertonen ondanks hun semi-solitaire (d.w.z. minder sociale) bestaan.

Tot slot kan een sensitiviteit voor emoties in mensen en andere dieren ook indirect gemeten worden via *impliciete associaties.* Een voorbeeld van impliciete associaties is het automatisch indelen van anderen in categorieën zoals 'prettig', 'onprettig', of 'bekende' en 'vreemde'. Onze hersenen blijken vaak onbewust dit soort categorisaties toe te passen om zo onze sociale wereld overzichtelijk te maken. Emoties spelen een cruciale rol in het reguleren van deze impliciete sociale evaluaties en versterken of verzwakken ze op basis van eerdere ervaringen. Of apen, net als mensen, dit soort impliciete associaties hebben met soortgenoten en emoties, is nog niet bekend. In **hoofdstuk 7** onderzocht ik of een bekende taak in de psychologie, de *Implicit Association Test (IAT)*, aangepast kon worden om zo onderzoek met apen (of andere dieren die een touchscreen kunnen bedienen) mogelijk te maken. In de originele IAT wordt gebruik gemaakt van woorden, waardoor deze test niet geschikt is voor onderzoek met dieren. Ik heb daarom onderzocht of een pictoriale variant van de IAT (de P-IAT) even sterk is in het meten van impliciete associaties als zijn woord-tegenhanger (W-IAT). De bevindingen laten zien dat de P-IAT inderdaad impliciete associaties kan blootleggen bij mensen en dat de taak nagenoeg even goed werkt als een W-IAT. Hoewel ik niet heb kunnen testen of de P-IAT ook bij apen werkt, hoop ik dat andere onderzoekers de P-IAT bruikbaar zullen vinden voor vergelijkend onderzoek tussen soorten.

In mijn onderzoeken heb ik me toegelegd op het verkrijgen van nieuwe inzichten in de emotionele belevingswereld van mensapen. Helaas worden mensapen ernstig met uitsterven bedreigd en verdwijnt de unieke kans om ons eigen evolutionaire verleden te ontdekken in een razendsnel tempo. Het is daarom van belang dat er snel meer onderzoek gedaan wordt naar de emotionele capaciteiten van mensapen. Op basis van deze drijfveer heb ik in dit proefschrift de overeenkomsten en verschillen in emotieperceptie bij bonobo's, orang-oetans en mensen bestudeerd. In het bijzonder heb ik de aandachts- en gedragsmechanismen onderzocht die ten grondslag liggen aan emotieperceptie en de bouwstenen vormen voor complexe cognitieve processen zoals empathie. De resultaten in dit proefschrift duiden op een gedeelde evolutionaire oorsprong voor aandachtsmechanismen betrokken bij emotieperceptie en voor mimicry in de mensapen. Belangrijk hierbij is dat soort-specifieke eigenschappen zoals xenofilie in bonobo's en het semi-solitaire bestaan van orang-oetans de aandachts- en gedragsmechanismen voor emotieperceptie aansturen. Mijn werk heeft nieuwe, cruciale vragen opgeroepen die in toekomstig onderzoek verder onderzocht moeten worden. Zo weten we op dit moment nog niet welke emotionele uitingen het meest in het oog springen of relevant zijn voor de mensapen. Daarnaast is het belangrijk dat we beter leren begrijpen hoe mensapen (en andere dieren) emotionele uitingen waarnemen als het gaat om emotionele valentie en intensiteit. We kunnen ontzettend veel leren van onze naaste verwanten en van de unieke eigenschappen van andere dieren. Met dit proefschrift hoop ik een opstap te kunnen bieden voor verder onderzoek naar de aandachts- en gedragsmechanismen die ten grondslag liggen aan emoties.

Curriculum Vitae

Curriculum Vitae

Evy van Berlo was born in 1990 in Maastricht and graduated from high school Porta Mosana in Maastricht in 2008. She completed her Bachelor of Science degree in Psychology at the Erasmus University Rotterdam (2009 – 2012), after completing the *Brain and Cognition* track. She continued with a Research Master of Science degree in Cognitive Neuroscience at Utrecht University (2012 – 2016). It was during this time she discovered the work of Dutch primatologist Prof. dr. Liesbeth Sterck and subsequently Prof. dr. Frans de Waal, which inspired her to pursue a career in comparative psychology and specialize in primate behavior and cognition.

For her first Master's project, Evy investigated the social dynamics within a group of long-tailed macaques during the introduction of a new alpha male[1]. During this time, she became familiar with behavioral observation methods commonly used in Biology, and learned a great deal about primate cognition and behavior. For her second Master's project, she studied the function of yawning in bonobos housed at primate park Apenheul[2]. To conduct this project, she acquired funding from the European Human Behaviour and Evolution Association (EHBEA). For her final Master's project, she wrote a research proposal on measuring an attentional bias for emotions of familiar and unfamiliar others in bonobos[3]. With the invaluable help from Prof. dr. Mariska E. Kret, Evy could continue her research work in pursuance of a Doctoral degree in comparative psychology.

Evy conducted her Ph.D. research at the Cognitive Psychology Unit of the Institute of Psychology (Leiden University, 2016 – 2021) and Apenheul under the supervision of Prof. dr. Kret, and later also Dr. Yena Kim. During this time, she completed eight studies of which six are part of her Doctoral degree and described in this dissertation. In collaboration with Prof. dr. Kret, Evy set up two new research facilities at Apenheul, introducing eye-tracking to bonobos and touchscreen-based research to orangutans[4]. To make comparisons with humans possible and engage the general public in research, she also realized a research setup to test (human) visitors of Apenheul. Furthermore, half-way through her first year of the Ph.D., she conducted a 3-month study on stress and (contagious) yawning in wild chimpanzees in Rwanda, Africa[5]. Moreover, Evy was involved in teaching several courses at Leiden University and supervised many Bachelor and Master students who helped her collect data with the visitors of Apenheul. Due to health problems, Evy was required to take a hiatus in 2019. With the support of her fantastic colleagues and friends, she was able to hand in the dissertation you now have before you.

After her Ph.D., Evy brought together an interdisciplinary team of scientists to study social decision-making in chimpanzees and humans using VR technology as part of a postdoctoral project. Together with her team[6], she acquired funding from the University of Amsterdam (Amsterdam Brain and Cognition) and the Dr. J. C. Dobberke Stichting voor Vergelijkende Psychologie to make this project a reality.

---

[1] Supervision: Prof. Dr. Liesbeth H. M. Sterck, dr. Marie-José H. M. Duchateau, and Lisette van den Berg, M.Sc., Utrecht University

[2] Supervision: Prof. Dr. Mariska E. Kret, University of Amsterdam and Leiden University; Dr. Jorg J. M. Massen, University of Vienna.

[3] Supervision: Prof. dr. Kret and Prof. dr. Frans B. M. de Waal, Emory University and Utrecht University.

[4] To do so, she acquired funding from the Dr. J. C. Dobberke Stichting voor Vergelijkende Psychologie, and the Stichting Elize Mathilde Fonds (Leiden University Fund, in collaboration with Prof. Dr. Kret).

[5] Made possible with the invaluable help of two research assistants (Daan W. Laméris, M.Sc. and Berta Roura Torres, M.Sc.) and funding from the KNAW Ecology grant (Kret & van Berlo) and Stichting Het Kronendak (Laméris & van Berlo). In collaboration with Prof. dr. Kret, Prof. Dr. Karline R. L. Janmaat (Max Planck Institute for Evolutionary Anthropology), Dr. Aaron Rundus, and Dr. Rebecca Chancellor (West Chester University)

[6] Team: Prof. Dr. Janmaat, Dr. Jan Engelmann and Dr. Jan Hausfeld, and Dr. Lucas Molleman (University of Amsterdam)

List of publications

List of publications

Published work

Mendoza-Straffon, L. M., Agnew, G., Desch-Bailey, C., Van Berlo, E., Goclowska, G., & Kret, M. E. (2022). Visual attention bias for self-made artworks.
Psychology of Aesthetics, Creativity, and the Arts (*in press*).

Kret, M. E., & Van Berlo, E. (2021). Attentional Bias in Humans Toward Human and Bonobo Expressions of Emotion. *Evolutionary Psychology, 19*(3), 14747049211032816.

Van Berlo, E., Díaz-Loyo, A. P., Juárez-Mora, O. E., Kret, M. E., & Massen, J. J. (2020). Experimental evidence for yawn contagion in orangutans (*Pongo pygmaeus*).
Scientific reports, 10(1), 1-11.

Laméris, D. W., Van Berlo, E., Sterck, E. H. M., Bionda, T., & Kret, M. E. (2020). Low relationship quality predicts scratch contagion during tense situations in orangutans (*Pongo pygmaeus*).
American Journal of Primatology, e23138.

Work submitted for publication

Van Berlo, E.*, Laméris, D. W.*, Roth, T. S., & Kret, M. E. (2022). No evidence for biased attention towards emotional scenes in Bornean orangutans (*Pongo pygmaeus*).
Manuscript submitted to Affective Science.
** Authors equally contributed to the paper*

Van Berlo, E., Kim, Y., & Kret, M. E. (2022). Attentional selectivity for emotions: humans and bonobos compared.
Manuscript submitted to Proceedings of the Royal Society B: Biological Sciences.

Van Berlo, E., Bionda, T., & Kret. M. E. (2020). Attention towards emotions is modulated by familiarity with the expressor. A comparison between bonobos and humans.
Manuscript submitted to Emotion.

Van Berlo, E., Otten, M., Binnekamp, J., Roth, T. S., Van der Ven, E., Kret, M. E. (2020). Measuring implicit associations with a Pictorial IAT: Implications for comparative research.
Manuscript submitted to Humanities and Social Sciences Communications.

Work in preparation

Van Berlo, E., Roth, T. S., Samara, I., Kret, M. E., & Massen, J. J. M. (2022). Yawn and self-scratch contagion in bonobos.

Cover design: Evy van Berlo