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

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.

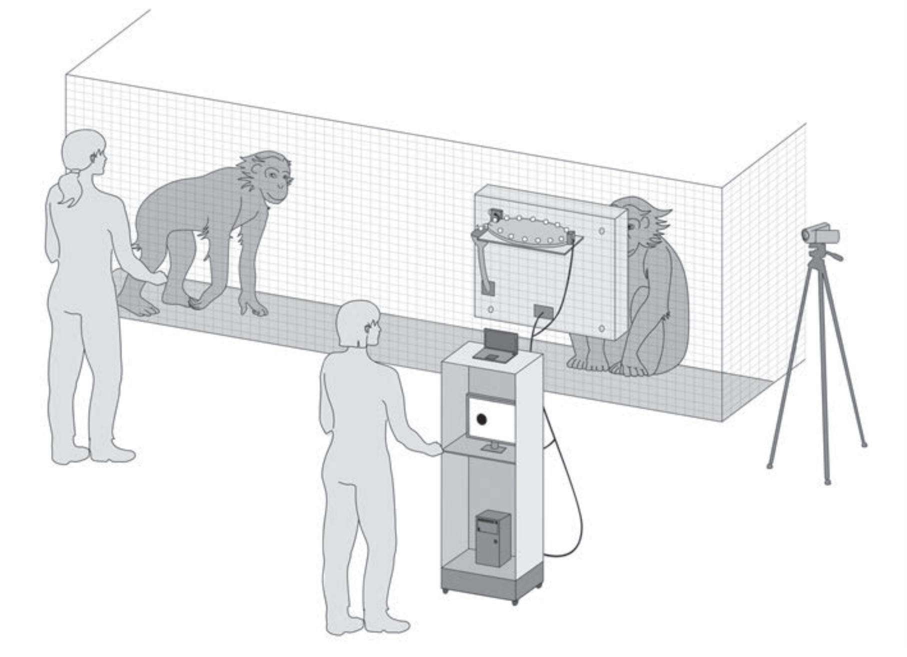


Figure 1. Abstract representation of the bonobo setup. The experimenter (right) controlled the experiment from behind the bonobo setup while a keeper (left) distracted the other bonobos. The experimenter was not visible and remained silent to the bonobos most of the time, but the experimenter would move 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).

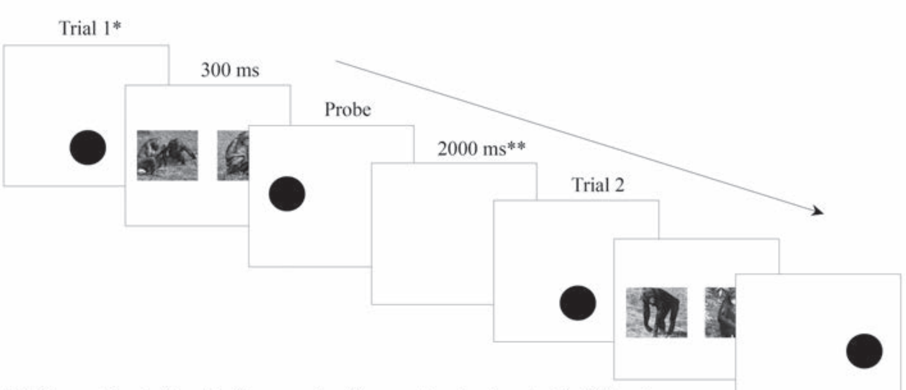
Figure 2. Examples of stimuli of all emotional categories used in Experiment 1. An emotional picture was always paired with a neutral picture. The emotional and neutral pictures were of either familiar or unfamiliar individuals.

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, $p < .001$).



* Trials contained either familiar or unfamiliar emotional and neutral individuals
**In case of a correct response, a food reward was automatically provided

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, α =.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

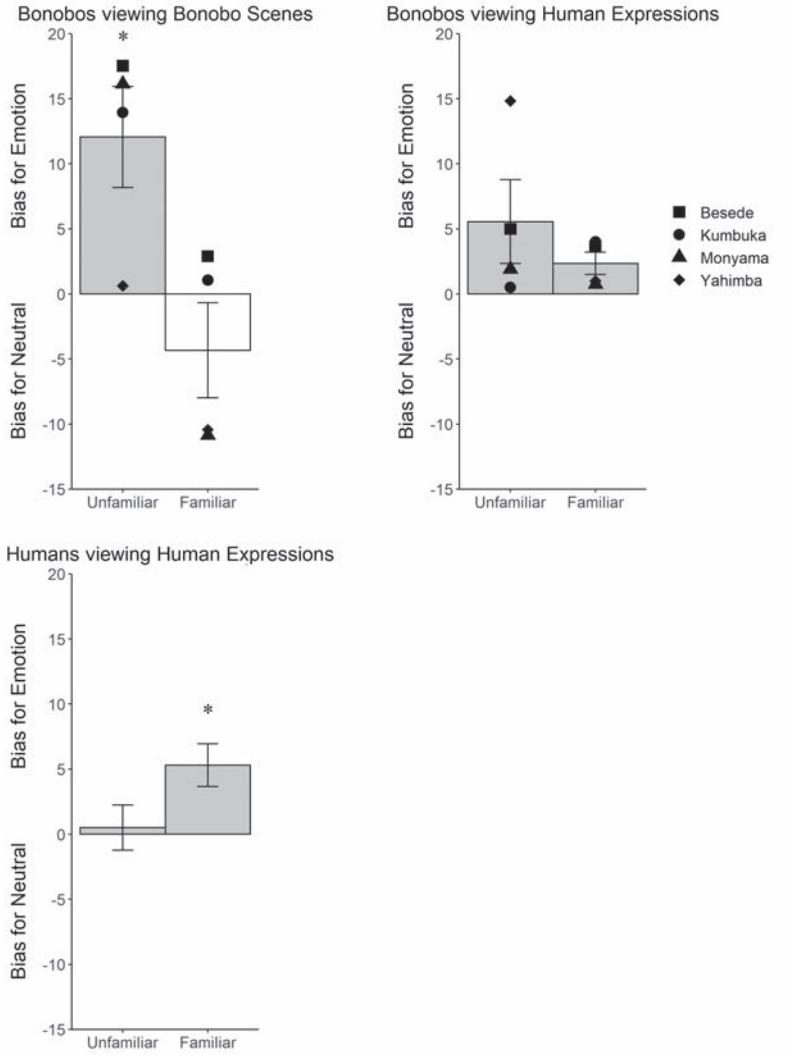


Figure 4. Experiment 1: Bonobos show an attentional bias towards emotions of unfamiliar, but not familiar conspecifics (top left). Experiment 2: Bonobos do not show an attentional bias towards emotions of familiar or unfamiliar humans (top right). Experiment 3: Humans have an attentional bias towards 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).

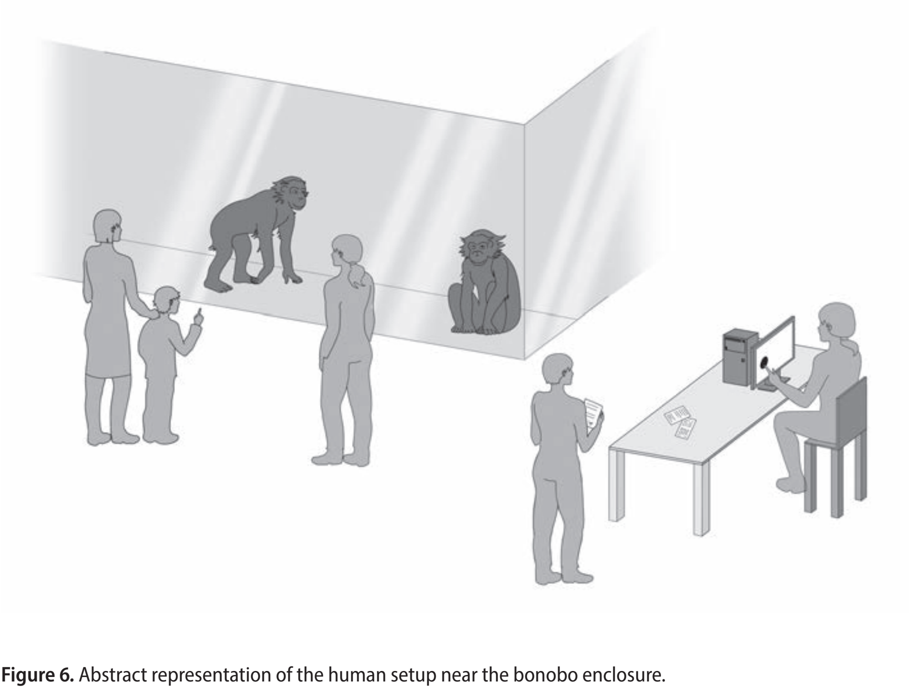


Figure 6. Abstract representation of the human setup near the bonobo enclosure.

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.