Animal Cognition (2018) 21:715–728
https://doi.org/10.1007/s10071-018-1205-z

ORIGINAL PAPER

Human ostensive signals do not enhance gaze following in chimpanzees, but do enhance object-oriented attention

Fumihiro Kano[1] · Richard Moore[2] · Christopher Krupenye[3,4] · Satoshi Hirata[1] · Masaki Tomonaga[5] · Josep Call[3,4]

Received: 20 April 2018 / Revised: 13 July 2018 / Accepted: 22 July 2018 / Published online: 26 July 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

The previous studies have shown that human infants and domestic dogs follow the gaze of a human agent only when the agent has addressed them ostensively—e.g., by making eye contact, or calling their name. This evidence is interpreted as showing that they expect ostensive signals to precede referential information. The present study tested chimpanzees, one of the closest relatives to humans, in a series of eye-tracking experiments using an experimental design adapted from these previous studies. In the ostension conditions, a human actor made eye contact, called the participant's name, and then looked at one of two objects. In the control conditions, a salient cue, which differed in each experiment (a colorful object, the actor's nodding, or an eating action), attracted participants' attention to the actor's face, and then the actor looked at the object. Overall, chimpanzees followed the actor's gaze to the cued object in both ostension and control conditions, and the ostensive signals did not enhance gaze following more than the control attention-getters. However, the ostensive signals enhanced subsequent attention to both target and distractor objects (but not to the actor's face) more strongly than the control attention-getters—especially in the chimpanzees who had a close relationship with human caregivers. We interpret this as showing that chimpanzees have a simple form of communicative expectations on the basis of ostensive signals, but unlike human infants and dogs, they do not subsequently use the experimenter's gaze to infer the intended referent. These results may reflect a limitation of non-domesticated species for interpreting humans' ostensive signals in inter-species communication.

Keywords Domestication · Gaze following · Great ape · Ostensive signals · Referential communication

Introduction

Humans use 'ostensive' signals such as making eye contact and calling an intended recipient's name as a way of indicating their communicative intention to others. These signals function to alert an intended recipient to the possibility that a signaler has a message to convey (Csibra 2010; Moore 2016, 2017; Scott-Phillips 2015a, b; Sperber and Wilson 1995), and so provide the recipient with evidence that they should devote their cognitive resources to figuring out the content of that message. The ability to determine, on the basis of non-verbal cues, when others are acting with communicative intention has been argued to play a fundamentally important role in both language acquisition and cultural learning more generally (e.g. Gergely and Csibra 2006; Tomasello 1999, 2006).

Csibra and Gergely have suggested that humans possess an adaptation for 'natural pedagogy', which explains how humans efficiently transmit generic knowledge between individuals (Csibra 2010; Csibra and Gergely 2009; Gergely and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10071-018-1205-z) contains supplementary material, which is available to authorized users.

✉ Fumihiro Kano
fkanou@gmail.com

[1] Kumamoto Sanctuary, Kyoto University, 990 Misumi, Uki, Kumamoto 8693201, Japan

[2] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

[3] Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[4] School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, UK

[5] Primate Research Institute, Kyoto University, Inuyama, Japan

Csibra 2006; Gergely et al. 2007). According to their proposal, human infants have a set of perceptual and cognitive biases that make them interpret ostensive signals as indicating that an agent is trying to deliver generic information. These biases are: (1) preferential attention for the source of ostensive signals: Human infants are highly sensitive to the presence of signals (e.g. eye contact and infant-directed speech) that indicate that they are being addressed by a communicating agent. (2) Referential expectation: Following ostensive signals, human infants expect to find the intended referent of the communicating agent's message—that is, the entity about which the agent is communicating. (3) Generality: infants take ostensive communication to provide them with generic information about objects like those to which the agent referred—that is, information that is generalizable to objects in other situations.

Although the nature of the mechanisms at work in Csibra and Gergely's hypothesis remain controversial (Heyes 2016; Hoicka 2016; Moore et al. 2013, 2015b), its functional aspects are supported by compelling evidence. Related to the first feature of their proposal, it is well established that human infants have early-developing preference for attending to faces (Farroni et al. 2002) and infant-directed speech (Cooper and Aslin 1990). In addition, in relation with the second feature, Senju and Csibra (2008) found that 6-month-old humans followed an experimenter's gaze to an intended referent when it was preceded by either directed speech or ostensive eye contact, but not when a similarly salient animation was used to solicit their attention. This suggests that ostensive signals may help infants to identify an agent's referential goals, and thereby better understand referential communication. Finally, related to the third feature, Topál et al. (2009) found that 9-month-old humans (and domestic dogs, but not wolves) made more frequent search errors in the A-not-B task when an agent ostensively hid the objects compared to when the agent just hid the objects without any ostensive signals. That is, in the ostension condition, infants (and dogs) persistently searched for a hidden object at its initial hiding place even after observing it being hidden at another location. This finding is interpreted as showing that, on the basis of ostensive signals, the infants (and dogs) had formed general expectations about where the objects would be hidden, and these expectations trumped their own experience of seeing them hidden (but see Vorms (2012) for criticism of this interpretation).

While the adaptation for natural pedagogy was initially proposed to be uniquely human, more recent studies suggest that domesticated dogs are attuned to humans' ostensive signals in ways that are similar to human infants. Dogs spontaneously attend to human faces in a variety of communicative contexts (Topál et al. 2014), and they use both eye contact and (to a lesser extent) name-calling to identify that an experimenter is communicating with them (Kaminski et al. 2012). Additionally, Téglás et al. (2012) showed that, using the experimental paradigm developed for human infants (Senju and Csibra 2008), dogs followed an experimenter's gaze to a referent only if it had been preceded by ostensive eye contact and directed speech. Thus, similar to human infants, dogs may have the expectation that humans' ostensive signals precede referential information. These results support the domestication hypothesis, which postulates that, likely as a consequence of their domestication, domestic dogs have evolved adaptive responses to human referential communication, in a manner similar to human infants. This domestication hypothesis gains support particularly from comparative studies with human infants, domestic dogs and non-domesticated species such as wolves and great apes. Domestic dogs, even from a few weeks old, outperform chimpanzees and wolves at a number of tasks in which they must read human communicative signals such as gazing and pointing to locate hidden food (Hare et al. 2002; Riedel et al. 2008; Topál et al. 2014). Virányi et al. (2008) compared hand-raised 4-month-old wolf and dog puppies, finding that the dogs were both more willing to maintain eye-contact with experimenters, and better able to use the experimenters' points to find hidden food. While wolves were able to learn to respond to ostensive signals after training, the results suggest that dogs possess an early-developing responsiveness to human communication that wolves do not.

If human infants and domestic dogs follow human gaze with the expectation that ostensive signals precede referential information (Senju and Csibra 2008; Téglás et al. 2012), an interesting question is whether nonhuman great apes would do so as well. Although great apes (chimpanzee, bonobo, gorilla and orangutan) have not been domesticated by humans, they are humans' closest living relatives and thus are much more closely related to humans phylogenetically than are dogs. In addition, they are equipped with a variety of key skills that might enable them to act like human infants and dogs. Like humans (and many other animals, Bugnyar et al. 2004; Emery 2000), great apes spontaneously follow others' gaze. They do not just co-orient with others but take others' visual perspectives into consideration when following their line of sight (Bräuer et al. 2005; Okamoto-Barth et al. 2007). Moreover, as in humans, eye contact plays an important role in the natural repertoire of communicative behaviors in great apes (Gomez 1996; Liebal and Call 2012). For example, when chimpanzees and gorillas attempt to reconcile with conspecifics after fighting, they first establish eye contact before approaching their counterparts (De Waal 1990; Yamagiwa 1992). When tension arises among individuals, bonobos regulate it by making eye contact and engaging in sexual activities (De Waal 1988). Some apes even use their eye contact ostensively when requesting food from human experimenters in a laboratory (Gomez 1996).

In light of the pervasive role of eye contact in great ape communication, great apes should satisfy, at least at a functional level, the first feature of natural pedagogy hypothesis (Csibra and Gergely 2009): the preference for the potential source of communicative signals. The previous studies have also shown that great apes spontaneously attend to both conspecific and human faces (Kano et al. 2012; Kano and Tomonaga 2009). Like human infants (Farroni et al. 2002), chimpanzee infants preferentially attend to human faces with direct gaze, rather than those with averted gaze (Myowa-Yamakoshi et al. 2003). Chimpanzees are less accurate in distinguishing between the faces of humans and conspecifics when those faces are presented upside-down compared to when they are presented upright (Parr et al. 1998; Tomonaga 1999). Finally, for apes living in a typical zoo and research environment, human caregivers and experimenters regularly call apes' individual names (and make eye contact with them) when attempting to communicate with them. A previous study measuring a chimpanzee's event-related potentials showed that the chimpanzee became attentive immediately after they heard their own names called by a human experimenter (Ueno et al. 2010).

However, it remains unclear whether great apes have referential expectations following ostensive signals, or whether they understand others' intentions to communicate about specific referents. One study reported that, when chimpanzees saw that another individual is requesting a particular item, they could infer the item that the other was requesting, even on the basis of ambiguous gestures (Yamamoto et al. 2012). Moreover, while it was once doubted that great apes could understand others' intentions at all, recent evidence challenges this view. Great apes understand others' goals and intentions and utilize that knowledge in various social contexts (Call and Tomasello 2008; Kano and Call 2014b). A recent eye-tracking study even showed that, when apes are viewing an agent and an antagonist competing for an object, they anticipate the agent's actions according to the agent's false beliefs (Kano et al. 2017; Krupenye et al. 2016).

Nonetheless, in general, great apes do not appear to be as sensitive to the referential aspects of human communication as human infants and domestic dogs. For example, in studies, where human experimenters try to inform them of the location of hidden food with referential gestures such as gaze and pointing, both human infants (Behne et al. 2005, 2012) and domestic dogs (Hare and Tomasello 1999; Miklósi et al. 1998) excel at using such experimenters' cues to locate food. However, great apes perform comparatively poorly in similar paradigms (Hare and Tomasello 2004; Herrmann and Tomasello 2006), although enculturated apes (those reared by humans in human environments) generally perform better than unenculturated apes (Call and Tomasello 1994, 1996; Lyn et al. 2010). Previously, only one study (Moore et al. 2015a) has addressed whether orangutans respond differently to pointing gestures produced with and without ostensive signals. In this study, no effect of ostension was found. However, since comprehension was poor in all conditions, this finding is difficult to interpret.

A number of studies have tested great apes' gaze following behavior. In several previous experiments (Call et al. 2000; Tomasello et al. 2007), the human experimenter called the apes' names and made eye contact before providing a gaze cue. Such cues did seem to work, at least in drawing the apes' attention to the experimenter's face (before the gaze cue was provided). In an eye-tracking experiment presenting still images to chimpanzees, Hattori et al. (2010) showed that chimpanzees followed the gaze of a conspecific agent but not that of a human agent. In another eye-tracking experiment (Kano and Call 2014a), while bonobos, orangutans and human adults all followed the gaze of both conspecific and allospecific agents, human infants and chimpanzees only followed the gaze of conspecific agents. These findings indicate that at least chimpanzees may not be receptive to following a human experimenter's gaze when watching still pictures and movies without clear ostensive cues. None of these studies directly compared the effect of humans' ostensive signals with that of control attention-getters on gaze following in great apes.

In this study, we examined whether great apes, particularly chimpanzees, would exhibit enhanced gaze following in response to a human actor establishing ostensive eye contact and calling the participant's name. We designed our experiment based on the previous eye-tracking studies of human infants and dogs (Senju and Csibra 2008; Téglás et al. 2012). According to the domestication hypothesis and the previous evidence that apes are not good at comprehending human referential communication, they may not understand a human actor's ostensive signals in the same way as human infants and dogs do. On the other hand, based on the previous evidence that apes make eye contact with conspecifics and humans, and are familiar with humans calling their names, such signals should have some effect on apes' behavior. While it is possible that humans' ostensive signals would enhance gaze following in great apes just as in human infants and dogs, it is also possible that ostensive signals function in a more limited way—for example, by enhancing attention to the actor's face or to all objects in front of the actor (i.e. to both cued and non-cued objects).

The question we address here is whether chimpanzees would be more likely to follow the gaze of a human actor after the actor addressed them ostensively as compared with after a control attention-getter attracted their attention to the actor's face. We specifically tested (1) whether the actor's ostensive signals attracted apes' attention to the actor's face in the cueing phase (when the actor was addressing either ostensively or non-ostensively) as strongly as the control cues in the cueing phase, and (2) whether the ostensive

signals enhanced the chimpanzees' gaze following (i.e. their looking at the cued object), attention to the objects (i.e. their looking at both cued and non-cued objects in front of the actor), and/or attention to the actor's face more strongly than the control cues in the looking phase (when the actor was looking at one of the objects).

We tested chimpanzees from different rearing backgrounds: zoo-reared and institute-reared individuals. The institute-reared individuals (who were similar but technically not identical to 'enculturated' individuals; see Method) had richer early experiences of interacting with human caregivers than the zoo-reared individuals. It is well established that such individuals perform better than the other individuals in tests in which they need to locate hidden foods based on an experimenter's referential cues (Call and Tomasello 1994, 1996; Lyn et al. 2010). In addition, a previous eye-tracking study found that the institute-reared chimpanzees paid more attention to the objects manipulated by conspecifics than the zoo-reared chimpanzees (Kano et al. 2018). It was thus expected that the institute-reared chimpanzees would show greater sensitivity to human ostension than the zoo-reared chimpanzees in our test. We also tested two other closely related great ape species, bonobos and orangutans, in this study, but mainly focus on the results for the chimpanzees in this article. We do so primarily for the simplicity of analyses and reports: although we found that bonobos and orangutans were similar to chimpanzees in terms of the key results, there were differences between the three species in their basic responses to the human agent's gaze cues (but not to the ostensive and gaze cues). We discuss the bonobo and orangutan results briefly in the main text and report them in greater detail in the Supplemental Materials.

We used the actor's eye contact and calling of the participant ape's name as ostensive cues. One potential methodological issue is that it remains unclear from the previous studies what control cues are appropriate for great apes. In particular, it remains unclear to what extent different control cues attract apes' attention in comparison to the ostensive cues. The previous studies with human infants and dogs (Senju and Csibra 2008; Téglás et al. 2012) used a visually salient object (presented on the top of the actor's head) as a control for eye contact and a low-pitched, adult-directed voice as a control for a high-pitched, infant/dog-directed voice. We could not use a low-pitched voice as a control, because in their daily lives the apes we tested often hear their names in a low pitch. We were also uncertain which visual cues could be used as a control because of debates over the extent to which different control stimuli attracted covert and overt attention to the actor's face in the previous studies (see Szufnarowska et al. 2014 and also the comments by Senju and Csibra in the same journal). We thus examined the effect of control cues using several different attention-getters in Experiment 1–3 to explore to what extent the use of different control cues altered the chimpanzees' responses to the actor's ostensive and gaze cues. Experiments 1–3 used the same design but differed in the types of control cue used. Experiment 1 used the actor's head gesture with a voice (unrelated to the participant ape) as a control (partly following Szufnarowska et al. 2014). Experiment 2 used a visually salient object with an artificial sound on the actor's face as control (following Senju and Csibra 2008; Téglás et al. 2012). Experiment 3 used the actor's eating action with a crunch sound and a voice as control (a control used for the first time in this study). As a result, we could ensure that chimpanzees were similarly attentive to the human actor's ostensive and gaze cues across the experiments.

Method

Participants

Experiment 1 tested 15 chimpanzees (*Pan troglodytes*) from the Wolfgang Koehler Primate Research Center (WKPRC) in Leipzig, Germany. Experiment 2 tested 12 chimpanzees from the Kumamoto Sanctuary (KS) in Kumamoto, Japan, and the Primate Research Institute (PRI) in Inuyama, Japan. Experiment 3 tested the same 19 chimpanzees (5 from KS and 14 from WKPRC). We also tested 7 bonobos (*Pan paniscus*) and 7 orangutans (*P. troglodytes*) from WKPRC in Experiment 1, 6 bonobos from KS in Experiment 2, and the same 13 bonobos and 6 orangutans in Experiment 3 (see the Supplemental Material for the results from these species). We did not exclude any apes in this study. All apes were reared in captivity and lived with conspecifics in enriched naturalistic environments at WKPRC, KS, and PRI. All apes had some experience watching naturalistic movies for enrichment and in experiments, although they were never explicitly trained for their gaze behavior. The chimpanzees from KS (recently moved from Hayashibara Great Ape Research Institute, Okayama, Japan) and PRI had participated in numerous cognitive experiments since their youth. Consequently, they had more human interaction experience than the chimpanzees from WKPRC (they are similar to, but technically not 'enculturated' chimpanzees, as enculturated chimpanzees are typically defined as those reared by humans in human environments in literature; our chimpanzees were reared by their biological mothers or mostly by their conspecific peers in a chimpanzee group, see Table S1 for more details about each participant).

Ethics statement

All participants were tested in the testing rooms prepared for each species, and their daily participation in this study was voluntary. They were given regular feedings, daily

enrichment, and had ad libitum access to water. Animal husbandry and research protocol complied with international standards (the Weatherall report "The use of non-human primates in research") and institutional guidelines (KS: Wildlife Research Center "Guide for the Animal Research Ethics"; PRI: Primate Research Institute 2002 version of "The Guidelines for the Care and Use of Laboratory Primates"; WKPRC: "EAZA Minimum Standards for the Accommodation and Care of Animals in Zoos and Aquaria", "WAZA Ethical Guidelines for the Conduct of Research on Animals by Zoos and Aquariums", "Guidelines for the Treatment of Animals in Behavioral Research and Teaching" of the "Association for the Study of Animal Behavior (ASAB)").

Apparatus

An infrared head-free eye-tracker recorded their eye movements (60 Hz; X120 in WKPRC and X300 in KS and PRI; Tobii Technology AB, Stockholm, Sweden). The eye-tracker and monitor were installed outside of the testing room. Apes watched the movies on the monitor through a transparent acrylic panel (1–2 cm in thickness); we previously confirmed that this transparent acrylic panel does not interfere with recording of eye movements (Kano et al. 2011). Apes were allowed to sip diluted grape juice via a custom-made juice dispenser attached to the transparent acrylic panel (irrespective of their gaze behavior). In both facilities, the movies were presented at a viewing distance of 70 cm with a resolution of 1280 × 720 pixel on a 23-inch LCD monitor (43° × 24°) with Tobii Studio software (version 3.2.1). Two-point automated calibration was conducted for the apes by presenting a small object or movie clip on each reference point. Each time before the recording session, we manually checked the accuracy and repeated the calibration if necessary. Calibration errors were typically within a degree (Hirata et al. 2010; Kano et al. 2011).

Stimuli and procedure

Each experiment had two conditions: an ostension condition and a control condition (Fig. 1). Experiment 1–3 used the same test (ostension) condition but differed in the types of control cues. The test and control conditions of Experiment 1–2 were designed based on the previous studies with human infants and domestic dogs (Senju and Csibra 2008; Szufnarowska et al. 2014; Téglás et al. 2012). In all experiments, in the ostension condition, a human actor faced the participant and initially looked down, with two identical objects ("still phase"), one on each side of him. After 2 s, the actor looked up, made eye contact, and called the participant ape's name twice (the actor opened his mouth twice, and each participant's unique name was dubbed into the mouth movements). This "cueing phase" lasted for 2.5 s. The actor then turned his head to one of the two same objects ("target"; the other object is called "distractor") and kept still for the remaining time ("looking phase"; 5 s). The objects were supported by the actor's hand, while they rested on a table in Experiment 1–2, while they were mounted on tripods to each side of the actor at eye level in Experiment 3 (see "Results and discussion" for the reason for these changes). The control condition was the same as the ostension condition except that the cueing phase presented a non-communicative attention-getter instead of the actor's communicative cue. As a control attention-getter, in Experiment 1, the actor nodded 3 times (following Szufnarowska et al. 2014) and made a "hmm" sound (said as if to himself), during this action (Fig. 1a). The actor repeated this action and sound twice during the cueing phase. In Experiment 2, a circle with red-white color patterns (Fig. 1b) rolled 360 degrees with a chime sound twice, once clockwise, and then counterclockwise (following Senju and Csibra 2008). In Experiment 3, the actor ate an apple with crunching and "hmm" sounds (Fig. 1c).

Each participant was tested in both conditions; each participant first completed all trials in one condition over several days, before completing the other condition on subsequent days (i.e. within-subject design; one concern regarding this design is a carry-over effect, but note that a supporting analysis on the first 6 trials—which mimicked a between-subject design—yielded the same results; see Supplemental Material). The order of presentation of the conditions was counterbalanced between participants. Each condition had 6 trials; in 3 trials the actor looked at the left object in the looking phase and in the other 3 trials, the actor looked at the right object. The presentation order of the cued side was pseudo-randomized so that no more than 2 successive trials cued the same side. Each trial presented one movie file. Each trial presented a unique object as both target and distractor. Each participant received typically 2 trials a day (max. 4 trials depending on the motivation of the participant). We did not exclude any trials based on apes' performance or any other criteria.

Data analysis

The apes' eye-movement responses to each scene feature were coded automatically in the Tobii Studio software based on Area-Of-Interest (AOI). Apes' eye movement was filtered (fixations were extracted) using Tobii Fixation Filter in the same software. The AOIs were defined for the actor's face, the target object the actor gazed at and the distractor object the actor did not gaze at. The AOIs were drawn about 20% larger than the actual size of the face/object to avoid any fixation error (see Figure S1 for the defined AOIs). To check if the actor's ostensive signals equally attracted apes' attention, we examined the viewing times for the actor's face during

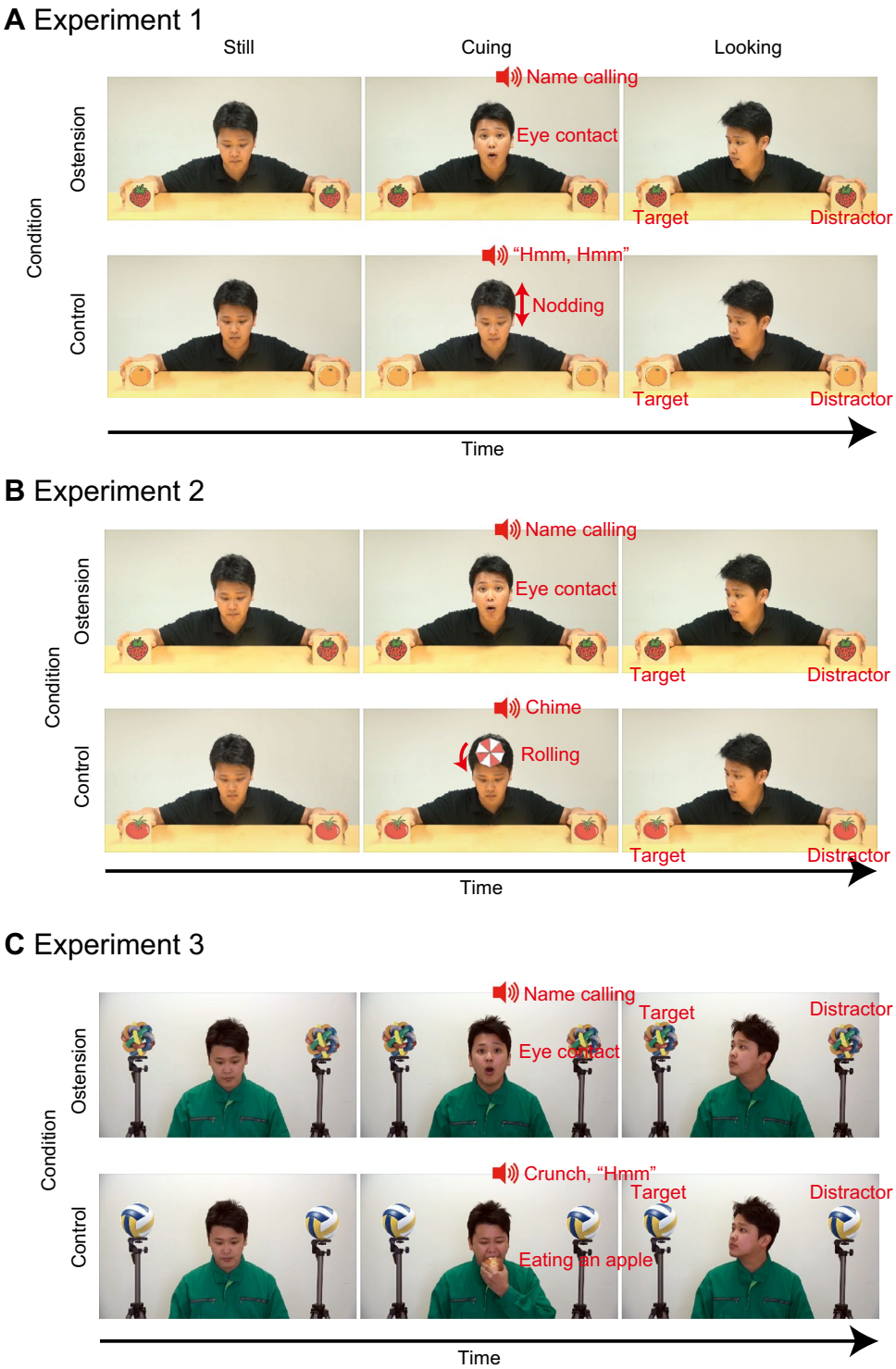

Fig. 1 Movie stimuli used in Experiment 1–3 (**a**–**c**)

each phase (still, cueing, and looking) in each condition. We then examined apes' responses to the actor's gaze cues during the looking phase. We examined their first looks (i.e. their initial responses) either to the target or distractor object (the number of trials in which they first looked either to the target or distractor object) and their total viewing times (i.e. their overall responses in 5 s) to the target and distractor object (the total viewing times for the target and distractor objects). Following the previous studies with great apes (Hattori et al. 2010; Kano and Call 2014a), but unlike in the original studies with human infants, we used raw scores of looking to the objects, instead of the number of saccades from the face to the object (because apes typically do not attend to the actor's face as much as human infants do in

the gaze-following context); however, preliminary analyses confirmed that the results were the same using either of these measures.

It should be noted that the original studies used difference scores for both the first-look and viewing-time measures (the response to the target minus the response to the distractor, divided by the sum of these), while this study used raw scores of first-look and viewing time to the target and distractor. We used raw scores and a repeated-measures ANOVA (instead of difference scores), because this method can analyze participants' overall level of attention to both objects (target and distractor). This means that the main effect of AOI (target, distractor) indicates the presence of gaze following (or more looks to the target than distractor), and the interaction effect of AOI and Condition (ostension, control) indicates the difference in gaze-following responses between conditions. The main effect of Condition indicates more looks to both target and distractor (which was not measured in the difference-score analysis). Thus, the only difference between our analysis and those based on difference-scores is that ours did not involve dividing the target-distractor differences by the target-distractor sum (because we were also interested in the difference in the target-distractor sum between conditions). Our preliminary analyses, however, confirmed that this method of calculation did not critically change our results (with regard to the presence of gaze following and the difference in gaze following between conditions).

Data availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Results and discussion for Experiment 1–3

Experiment 1 (WKPRC group)

To check if the actor's ostensive signals and control attention-getter similarly attracted chimpanzees' attention to the actor's face (Fig. 2a), we conducted a repeated-measures ANOVA with Condition (Ostensive, Control) and Phase (Still, Cueing, Looking) on the chimpanzees' viewing times for the actor's face. We found a significant main effect of Phase ($F(2,28) = 33.92$, $p < 0.001$, $\eta^2 = 0.71$) but no other significant main or interaction effects. This result indicates that chimpanzees attended more strongly to the actor's face during the cueing phase than the other phases. The results also indicate that chimpanzees attended equally strongly to the actor's face during all conditions (and more strongly during the cueing than the still and looking phases). We then examined the chimpanzees' first-look and viewing-time responses toward the target and distractor objects during the looking phase (Fig. 2a). A repeated-measures ANOVA with Condition (Ostensive, Control) and AOI (Target, Distractor) revealed no significant main or interaction effects in the first-look response [condition: $F(1,14) = 3.06$, $p = 0.10$, $\eta^2 = 0.18$; AOI: $F(1,14) = 2.89$, $p = 0.11$, $\eta^2 = 0.17$; interaction effect; $F(1,14) = 0.02$, $p = 0.90$, $\eta^2 = 0.001$] or in the viewing-time response [condition: $F(1,14) = 3.98$, $p = 0.07$, $\eta^2 = 0.22$; AOI: $F(1,14) = 2.15$, $p = 0.10$, $\eta^2 = 0.13$; interaction effect; $F(1,14) = 0.094$, $p = 0.76$, $\eta^2 = 0.007$].

Experiment 2 (KS-PRI group)

In Experiment 1, chimpanzees responded to the actor's ostensive and gaze cues only weakly. This result was not surprising given that this group of chimpanzees (zoo-reared: WKPRC) did not follow a human actor's gaze in a similar eye-tracking set-up (Kano and Call 2014a). In Experiment 2, we thus tested another group of chimpanzee participants (institute-reared: KS-PRI) who had richer experiences of interacting with human experimenters/caregivers since youth. In addition, in the results, we noticed that the control attention getter in Experiment 1 elicited apes' attention during the cueing phase slightly more weakly than the actor's ostensive cues. Consequently, we used another type of control cue that has been implemented in the previous studies with infants and dogs (Senju and Csibra 2008; Téglás et al. 2012).

As in Experiment 1, we first checked if the actor's ostensive signals and control attention-getter similarly attracted chimpanzees' attention to the actor's face (Fig. 2b); a repeated-measures ANOVA with Condition (Ostensive, Control) and Phase (Still, Cueing, Looking) on the chimpanzees' viewing times for the actor's face revealed a significant main effect of Phase [$F(2,22) = 26.43$, $p < 0.001$, $\eta^2 = 0.71$], but no other significant main or interaction effects; consistent with Experiment 1. We then examined the chimpanzees' first-look and viewing-time responses during the looking phase (Fig. 2b). For the first-look responses, a repeated-measures ANOVA with Condition (Ostensive, Control) and AOI (Target, Distractor) revealed a significant main effect of AOI [$F(1,11) = 7.29$, $p = 0.02$, $\eta^2 = 0.40$], but no significant effect of Condition [$F(1,11) = 2.17$, $p = 0.17$, $\eta^2 = 0.17$] or interaction between AOI and Condition [$F(1,11) = 0.007$, $p = 0.93$, $\eta^2 = 0.001$]. That is, chimpanzees made more first looks to the target object than the distractor in both conditions. For the viewing-time responses, a repeated-measures ANOVA with condition (ostensive, control) and AOI (target, distractor) revealed a significant main effect of Condition [$F(1,11) = 15.18$, $p = 0.002$, $\eta^2 = 0.58$], but no significant main effect of AOI [$F(1,11) = 2.29$, $p = 0.16$, $\eta^2 = 0.17$] or interaction effect [$F(1,11) = 0.33$, $p = 0.58$, $\eta^2 = 0.029$]. That

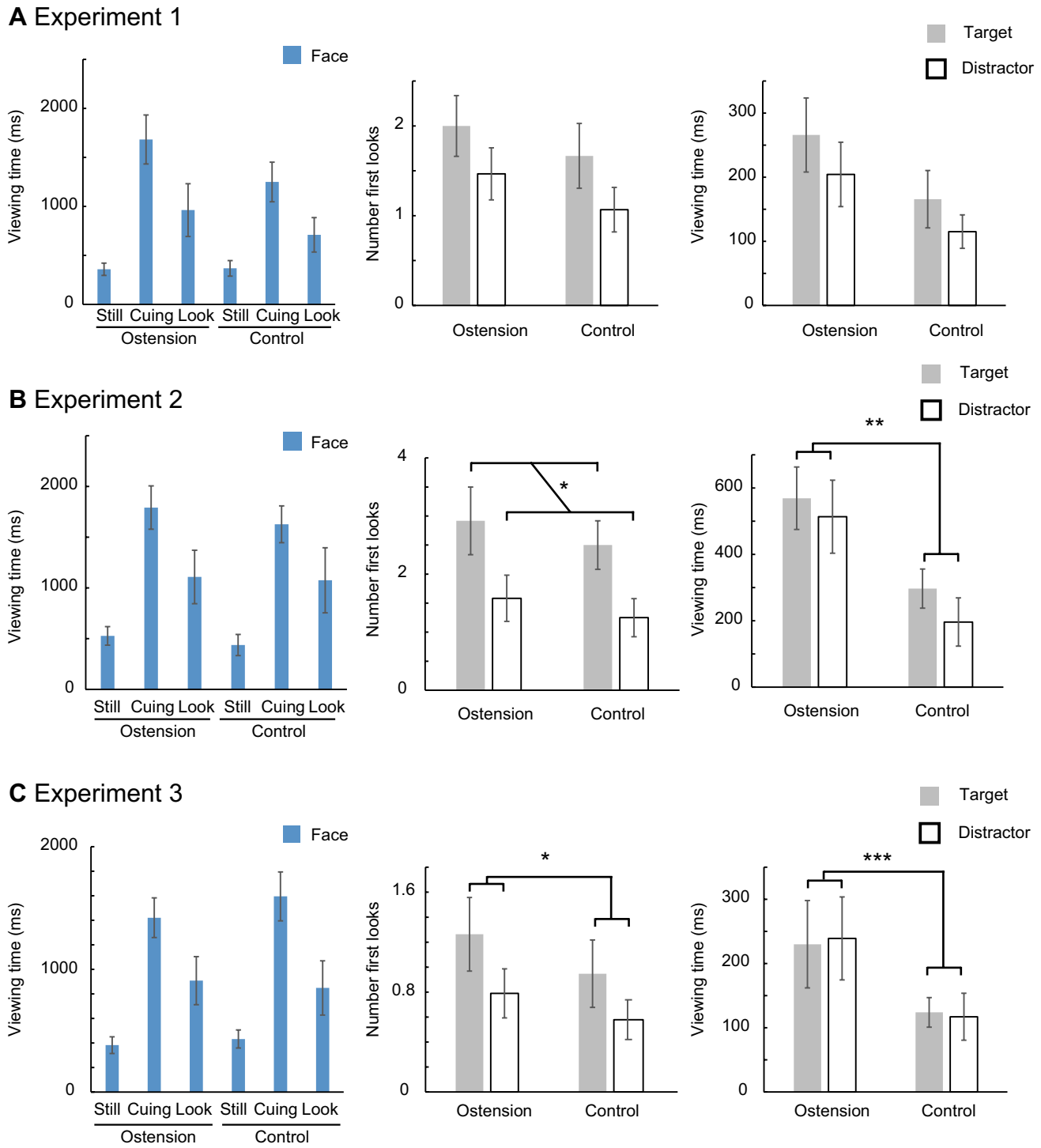

Fig. 2 Results from Experiment 1–3 (**a**–**c**). Mean viewing times for the actor's face in each phase of each condition (in 2.5 s), and number of first looks (in 6 trials) to the target or distractor object, and mean viewing times for the target or distractor object (in 5 s). Asterisks indicate the significance of the main effects (AOI-target/distractor, Condition-ostension/control). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

is, chimpanzees spent more time looking at both the target and distractor objects following the ostensive cue than following the non-ostensive cue. These results suggest that, (1) ostensive cues did not enhance those chimpanzees' gaze-following responses more than control cues, (2) although the actor's gaze cues did guide chimpanzees' looks to the target object (i.e., chimpanzees looked first to the target in both conditions). However, (3) the actor's ostensive

cues elicited greater looking to both objects (but not to the actor's face per se) as compared with control cues.

An additional analysis for Experiment 1 and 2

Experiment 1 and 2 consistently did not find significant interaction effects between Condition and AOI (i.e. a differential looking to the target and distractor between conditions), and the effect sizes were also small. In these experiments, however, the statistical results for the main effects of AOI/Condition were mixed; specifically, Experiment 1 found neither significant main effect of AOI nor that of Condition, while Experiment 2 found both (but in different eye-movement measures). Yet the result trends were similar between the experiments (Fig. 2). It thus remains unclear if the lack of significant effects simply reflects insufficient power of the statistical tests (e.g. small sample sizes) or specific differences between the experiments, such as the control cues used (the actor's shaking head vs. the animation on the actor's head) or the chimpanzee groups tested (zoo-reared vs. institute-reared).

To test this, we conducted a combined analysis on the results from Experiment 1–2 with an addition of the factor Group (WKPRC, KS-PRI). We first checked if the actor's ostensive signals and control attention-getter similarly attracted chimpanzees' attention to the actor's face; a repeated-measures ANOVA with Condition (Ostensive, Control), Phase (Still, Cueing, Looking), and Group (WK PRC, KS-PRI) on the chimpanzee's viewing times for the actor's face revealed a significant main effect of Phase [$F(2,48) = 55.85$, $p < 0.001$, $\eta^2 = 0.67$], but no other significant main or interaction effects. We then examined the chimpanzees' first-look and viewing-time responses during the looking phase. For the first-look responses, a repeated-measures ANOVA with Condition (Ostensive, Control), AOI (Target, Distractor), and Group (WK PRC, KS-PRI) revealed a significant main effect of AOI [$F(1,25) = 10.73$, $p = 0.003$, $\eta^2 = 0.30$] and Condition [$F(1,25) = 5.15$, $p = 0.032$, $\eta^2 = 0.17$], but no significant interaction effect [$F(1,11) < 0.001$, $p = 0.99$, $\eta^2 < 0.001$]. The effect of Group (either main or interaction effect) was not significant. These results indicate that chimpanzees followed the actor's gaze and looked first to the target in both conditions across experiments but also looked more often (in more trials) at both the target and distractor following the ostensive cue. For the viewing-time responses, a repeated-measures ANOVA with condition (ostensive, control), AOI (target, distractor), and group (WK PRC, KS-PRI) revealed significant main effects of AOI [$F(1,25) = 4.56$, $p = 0.043$, $\eta^2 = 0.15$] and condition [$F(1,25) = 20.57$, $p < 0.001$, $\eta^2 = 0.45$], but no significant interaction between AOI and condition [$F(1,25) = 0.18$, $p = 0.68$, $\eta^2 = 0.007$]. These results indicate that chimpanzees spent more time looking at the target across conditions but showed more looking to both the target and distractor following the ostensive cue. The main effect of group [$F(1,25) = 7.95$, $p = 0.009$, $\eta^2 = 0.24$] and the interaction effect of group and condition [$F(1,25) = 5.43$, $p = 0.028$, $\eta^2 = 0.18$] were also significant (the main effect of Condition was more evident in Experiment 2/KS-PRI group). Conducting the analysis separately for each condition (on the viewing-time responses) revealed that those two groups differed from one another in the ostensive condition [$F(1,25) = 9.31$, $p = 0.005$, $\eta^2 = 0.27$], rather than in the control condition [$F(1,25) = 2.61$, $p = 0.12$, $\eta^2 = 0.095$]. These results indicate that the KS-PRI chimpanzees responded to the actor's ostensive cues more strongly, looking more to both objects, than WKPRC chimpanzees (at least in the viewing-time responses). Overall, these combined analyses consolidated the findings from Experiment 1 and 2.

Experiment 3

Experiment 3 was conducted to further consolidate the findings from Experiment 1 and 2 with a different control attention-getter and minor changes in the stimuli. In Experiment 1 and 2, there was still a concern that those control attention-getters might have been slightly weaker than the actor's ostensive signals to attract apes' attention during the cueing phase. In Experiment 3, we used the actor's eating action as a control attention-getter, because a previous eye-tracking study confirmed that eating action strongly catches apes' attention (Kano et al. 2018). Moreover, in Experiment 1–2, there might be a concern that the actor holding both objects in his hands might confound the effect of gaze following (to the target objects) and that of attention to manual actions (to both target and distractor objects). In Experiment 3, we thus made minor changes in the configuration of the scenes in the movies so that the objects were mounted on tripods to each side of the actor at eye level instead of supported by the actor's hand, while they rested on a table. Finally, in Experiment 3, we tested the chimpanzees from both groups (WKPRC, KS-PRI) and thus included the factor Group into the analyses.

As in Experiment 1 and 2, we first checked if the actor's ostensive signals and control attention-getter similarly attracted chimpanzees' attention to the actor's face (Fig. 2c); a repeated-measures ANOVA with condition (ostensive, control), Phase (still, cueing, looking), and group (WKPRC, KS-PRI) on the chimpanzees' viewing times for the actor's face revealed a significant main effect of Phase [$F(2,34) = 29.5$, $p < 0.001$, $\eta^2 = 0.64$], but no other significant main or interaction effects; consistent with Experiment 1 and 2. We then examined chimpanzees' first-look and viewing-time responses to the target and distractor objects during the looking phase (Fig. 2c). For the first-look responses, a repeated-measures ANOVA with

condition (ostensive, control), AOI (target, distractor), and group (WKPRC, KS-PRI) revealed a significant main effect of condition [$F(1,17) = 6.34$, $p = 0.022$, $\eta^2 = 0.27$], but no significant main effect of AOI [$F(1,17) = 1.43$, $p = 0.25$, $\eta^2 = 0.08$] or interaction effect [$F(1,17) = 0.054$, $p = 0.82$, $\eta^2 = 0.003$]. The effect of group (either main or interaction effect) was not significant. The main effect of condition indicates that chimpanzees looked more often (in more trials) at both the target and distractor in the ostension than control condition. For the viewing-time responses, a repeated-measures ANOVA with condition (ostensive, control), AOI (target, distractor), and group (WKPRC, KS-PRI) revealed a significant main effect of condition [$F(1,17) = 24.89$, $p < 0.001$, $\eta^2 = 0.59$], but no significant main effect of AOI [$F(1,17) = 0.069$, $p = 0.80$, $\eta^2 = 0.004$] or interaction effect [$F(1,17) = 0.071$, $p = 0.79$, $\eta^2 = 0.004$]. These results indicate that chimpanzees spent more time looking at both objects in the ostensive condition than in the non-ostensive condition. The main effect of group [$F(1,17) = 7.45$, $p = 0.014$, $\eta^2 = 0.31$] and the interaction effect of group and condition [$F(1,17) = 22.94$, $p < 0.001$, $\eta^2 = 0.57$] were also significant; the KS-PRI chimpanzees looked more to both objects than WKPRC chimpanzees. Overall, the results from Experiment 3 were consistent with Experiment 1 and 2 (although we did not find the main effect of AOI, but this is likely due to weaker power in this analysis compared with the combined analysis, rather than due to the changes in the stimulus; note that bonobos and orangutans showed the main effect of AOI; see Supplemental Material).

Summary of the results for the other species (bonobos and orangutans)

We also tested bonobos in Experiment 1–3 and orangutans in Experiments 1 and 3 using the same stimuli and procedure, see Figures S2 and S3 for the summary of the results from bonobos and orangutans, respectively. Like chimpanzees, neither species followed the actor's gaze more sensitively in the ostension than control condition (i.e. no significant interaction between Condition and AOI). Interestingly, while orangutans were somewhat similar to chimpanzees in that they viewed both target and distractor objects longer in the ostensive than control condition (i.e. a significant main effect of Condition at least in Experiment 1), bonobos were not; in all three experiments, they spent similar time looking at the objects in both conditions. Moreover, bonobos viewed the actor's face longer in the ostension than control condition during the cueing phase. Presumably, such behaviors were driven by bonobos' reflexivity in following others' gaze (Kano and Call 2014a) and their general sensitivity to eye contact with other individuals (Kano et al. 2015).

General discussion

This study tested whether humans' ostensive signals enhance gaze following in great apes, particularly in chimpanzees. We found that, although chimpanzees did follow the actor's gaze (i.e., looked first to the target object following gaze cueing), unlike infants and domestic dogs, human ostensive signals did not enhance gaze following more strongly than control attention-getters for chimpanzees (nor for bonobos or orangutans; see supplementary materials). However, chimpanzees did distinguish between the ostensive signals and the control attention-getters (as did orangutans to some extent, although not bonobos). In the ostension condition, they spent more time attending to both the target object (the actor's intended referent) and the distractor than in the control condition. Importantly, they did so even though they paid an equal level of attention to the actor's face across conditions during both the cueing and looking phases. Thus, these results showed that the ostensive signals increased apes' attention specifically to the objects but not to the actor following the actor's ostensive cues. Overall, therefore, chimpanzees seemed to expect that the actor's ostensive signals would precede information specifically about the objects (rather than about the actor). Nonetheless, this expectation seems more functionally limited than in human infants and domestic dogs, because chimpanzees did not subsequently focus their attention on the intended referent of the actor's communicative act.

The finding that humans' ostensive signals do not enhance gaze following in great apes is consistent with the idea that human infants and domestic dogs are better at understanding humans' referential signals than great apes and wolves (Hare et al. 2002; Topál et al. 2009). It thus suggests that non-domesticated species such as great apes lack one of the skills (or perceptual/cognitive biases) that would help them to understand or respond appropriately to human referential communication, while human infants and domestic dogs have acquired such skills through ontogeny and evolution (Csibra and Gergely 2009; Senju and Csibra 2008; Téglás et al. 2012). However, our results do suggest that apes understand, at least partly, that humans' ostensive signals precede referential information, as they searched the environment longer after witnessing the ostensive signals compared to equally attention-grabbing non-ostensive signals. What they clearly didn't do is attempt to specify the intended referent further, on the basis of the actor's gaze behavior after seeing the ostensive signals. This might suggest that, although they understood a basic role of humans' ostensive signals (i.e. the agent is trying to communicate something), they have failed to understand the function of ostensive and gaze

signals combined (i.e. the agent is trying to communicate something about the cued object). Such reliance on environmental cues in apes may be observed not only in communicative contexts but also in the context of apes' social referencing behavior in general—and particularly in the context of research on chimpanzee imitation and emulation. The previous studies consistently suggest that when apes watch others using tools, they preferentially attend to the features of an environment that permit certain sorts of causal affordances, while being relatively inattentive to the particular techniques used by those whom they observe (Tennie et al. 2009; Tomasello 1999). By virtue of attending to environmental affordances, apes can learn to use tools by watching others. However, because they are inattentive to the particular techniques produced by agents, they are unable to reproduce any arbitrary features of actions (this is sometimes described as a preference for emulation over imitation). In contrast, human infants tend to show an opposite preference; they sometimes even over-imitate others (Whiten et al. 2009). It may be that, when apes are seeking to gain information about their environment, they have a stronger tendency than human infants to attend to the environment rather than to social cues—even when attending to social cues might prove particularly helpful.

The differences between our results with chimpanzees and the previous results with human infants and dogs are not due to differences in the particular control stimuli used in this and the previous studies. Experiment 2 presented a similar control attention-getter (a salient pattern on the actor's head) to that used in the previous studies (Senju and Csibra 2008; Téglás et al. 2012). Although the artificial nature of this control attention-getter was considered a potential problem in other studies (Gredebäck et al. 2018; Szufnarowska et al. 2014), in our study, this control condition produced similar results to the more natural actions performed by the human actor in Experiments 1 and 3. Critically, while human infants and dogs in the previous studies (Senju and Csibra 2008; Téglás et al. 2012) followed the actor's gaze –looked more frequently and longer at the cued object, respectively, in the first-look and viewing-time measures—only in the ostension conditions, apes in this study did not show such a pattern in either measure. Instead, we found that chimpanzees followed the actor's gaze across conditions (i.e., looked first to the cued target object) but continued to search longer (i.e., looked longer at both cued and non-cued objects) in the ostension conditions than the controls.

Interestingly, the differences between our results and the previous ones with chimpanzees highlight the potential importance of attention-getters in eliciting reliable gaze following. One notable difference between this and the previous ape eye-tracking (gaze-following) studies is that chimpanzees followed the human actor's gaze in this study; in the previous studies that lacked ostensive or attention-getting signals, chimpanzees did not follow human gaze (Hattori et al. 2010; Kano and Call 2014a). Thus, this study builds on the previous work by providing evidence that the general presence of attention-getters, including both ostensive signals and non-ostensive attention-getters, may help chimpanzees to follow human gaze in this setting. This may explain why chimpanzees reliably follow the gaze of a human experimenter in a live setup, where the human experimenter typically ensures the chimpanzees' attention to the face (or at least to the body) before providing the gaze cue (e.g. Call et al. 2000; Tomasello et al. 2007). Interestingly, this same argument may apply, to some extent, to human infants. That is, while human infants (and domestic dogs) showed limited gaze-following responses after seeing non-ostensive control attention-getters in the two earliest studies (Senju and Csibra 2008; Téglás et al. 2012), in two more recent studies (Gredebäck et al. 2018; Szufnarowska et al. 2014), they followed an actor's gaze after both ostensive and non-ostensive actions (e.g. head shivering; but see Senju and Csibra's (2014) commentary on the potential problems in Szefnarowska's article). It should be also noted that, unlike chimpanzees, human infants and dogs, human adults should follow the actor's gaze in any condition, because the task demands in this and the previous studies seem too easy for them. Our preliminary tests with human adults indeed showed that human adults strongly followed the actor's gaze in both ostention and control conditions (Figure S5). Thus, although the results of our three studies consistently show that apes do not gaze follow more robustly in response to ostensive cues than various non-ostensive cues, if humans also do not distinguish between those stimuli as efficiently as previously assumed, our conclusion about a species difference between humans and the other apes must be tentative; further studies are necessary to examine to what extent and in what circumstances humans distinguish between an actor's ostensive signals and control actions in potentially communicative contexts.

It is noteworthy that chimpanzees who had richer early experiences interacting with human caregivers paid more attention to the objects (but not to the actor's face) following the actor's ostensive cues than did other chimpanzees. This suggests that chimpanzees with more experience of human interaction are more sensitive to humans' ostensive signals than are the less experienced chimpanzees. This is consistent with the previous findings that enculturated chimpanzees show improved performances in tests in which they need to locate hidden foods based on an experimenter's referential cues (Call and Tomasello 1994, 1996; Lyn et al. 2010). Nonetheless, in our study, the ostensive signals did not enhance gaze following in either group of chimpanzees. These results suggest that, although chimpanzees with richer experiences with humans are generally

more sensitive to human signals, even those experienced chimpanzees do not interpret humans' ostensive signals in the same way as human infants and domestic dogs do.

An interesting question that can be addressed in future studies is whether the current findings can be extended to situations in which a conspecific ape actor, instead of a human actor, addresses participants ostensively. In general, chimpanzees (but not bonobos and orangutans) are more likely to follow conspecific than human gaze (Hattori et al. 2010; Kano and Call 2014a). While such differences may be explained by the fact that chimpanzees are more attentive to conspecific faces than human faces in the context of gaze following (Kano and Call 2014a), it remains untested whether chimpanzees would be more likely to follow the gaze of a conspecific after seeing a combination of species-typical attention-getters (hand clapping, spattering) and indexical cues (e.g. gazing and extending arms).

In conclusion, we confirmed that, unlike human infants and domestic dogs, humans' ostensive signals do not enhance gaze following in great apes. However, we also found that such signals do enhance subsequent object-related attention or search behaviors in apes, at least in chimpanzee participants who had richer early experiences of interacting with humans. Thus, they may, at least in part, expect ostensive signals to precede referential information. However, instead of fully relying on referential cues, apes may search for additional environmental cues to interpret communication. This may be a limitation (or a lack of human-like perceptual/cognitive bias) of non-domesticated species for interpreting humans' ostensive signals in inter-species communication.

Acknowledgements We thank Naruki Morimura, Yutaro Sato, Ruiting Song and the staff at Kumamoto Sanctuary, Primate Research Institute, Kyoto University Institute for Advanced Study, and Wolfgang Kohler Primate Research Center for their assistance in conducting the series of experiments. We also thank two anonymous reviewers for their helpful comments. Financial support came from Ministry of Education, Culture, Sports, Science and Technology [K-CONNEX to FK], Japan Society for Promotion of Science KAKENHI 26885040, 16K21108, and 18H05072 to FK, 26245069, 16H06301, 16H06283, and 18H05524 to SH, 15H05709, 16H06238, and JSPS-CCSN to MT, and JSPS-LGP-U04 and Great Ape Information Network to SH and MT, and the European Research Council [SOMICS 609819 to JC].

References

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

Behne T, Liszkowski U, Carpenter M, Tomasello M (2012) Twelve-month-olds' comprehension and production of pointing. Br J Dev Psychol 30:359–375. https://doi.org/10.1111/j.2044-835X.2011.02043.x

Bräuer J, Call J, Tomasello M (2005) All great ape species follow gaze to distant locations and around barriers. J Comp Psychol 119:145–154. https://doi.org/10.1037/0735-7036.119.2.145

Bugnyar T, Stöwe M, Heinrich B (2004) Ravens, Corvus corax, follow gaze direction of humans around obstacles. Proc R Soc Lond Ser B Biol Sci 271:1331–1336. https://doi.org/10.1098/rspb.2004.2738

Call J, Tomasello M (1994) Production and comprehension of referential pointing by orangutans (Pongo pygmaeus). J Comp Psychol 108:307. https://doi.org/10.1037//0735-7036.108.4.307

Call J, Tomasello M (1996) The effect of humans on the cognitive development of apes. In: Russon AE, Bard KA, Parker ST (eds) Reaching into thought: the minds of the great apes. Cambridge University Press, New York, pp 371–403

Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later. Trends Cogn Sci 12:187–192. https://doi.org/10.1016/j.tics.2008.02.010

Call J, Agnetta B, Tomasello M (2000) Cues that chimpanzees do and do not use to find hidden objects. Anim Cogn 3:23–34. https://doi.org/10.1007/s100710050047

Cooper RP, Aslin RN (1990) Preference for infant directed speech in the first month after birth. Child Dev 61:1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Csibra G (2010) Recognizing communicative intentions in infancy. Mind Lang 25:141–168. https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra G, Gergely G (2009) Natural pedagogy. Trends Cogn Sci 13:148–153. https://doi.org/10.1016/j.tics.2009.01.005

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

De Waal FBM (1990) Peacemaking among primates. Harvard University Press, Cambridge

Emery NJ (2000) The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci Biobehav Rev 24:581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

Farroni T, Csibra G, Simion F, Johnson MH (2002) Eye contact detection in humans from birth. Proc Nat Acad Sci 99:9602. https://doi.org/10.1073/pnas.152159999

Gergely G, Csibra G (2006) Sylvia's recipe: the role of imitation and pedagogy in the transmission of cultural knowledge. In: Enfield N, Levenson SC (eds) Roots of human sociality: culture, cognition, and human interaction. Berg, Oxford, pp 229–255

Gergely G, Egyed K, Király I (2007) On pedagogy. Dev Sci 10:139–146. https://doi.org/10.1111/j.1467-7687.2007.00576.x

Gomez JC (1996) Ostensive behavior in great apes: the role of eye contact. In: Russon AE, Bard KA, Parker ST (eds) Reaching into thought: the minds of the great apes. Cambridge University Press, New York, pp 131–151

Gredebäck G, Astor K, Fawcett C (2018) Gaze following is not dependent on ostensive cues: a critical test of natural pedagogy. Child dev. https://doi.org/10.1111/cdev.13026 **(in press)**

Hare B, Tomasello M (1999) Domestic dogs (Canis familiaris) use human and conspecific social cues to locate hidden food. J Comp Psychol 113:173. https://doi.org/10.1037//0735-7036.113.2.173

Hare B, Tomasello M (2004) Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. Anim Behav 68:571–581. https://doi.org/10.1016/j.anbehav.2003.11.011

Hare B, Brown M, Williamson C, Tomasello M (2002) The domestication of social cognition in dogs. Science 298:1634–1636. https://doi.org/10.1126/science.1072702

Hattori Y, Kano F, Tomonaga M (2010) Differential sensitivity to conspecific and allospecific cues in chimpanzees and humans: a comparative eye-tracking study. Biology Lett 6:610–613. https://doi.org/10.1098/rsbl.2010.0120

Herrmann E, Tomasello M (2006) Apes' and children's understanding of cooperative and competitive motives in a communicative situation. Dev Sci 9:518–529. https://doi.org/10.1111/j.1467-7687.2006.00519.x

Heyes C (2016) Born pupils? Natural pedagogy and cultural pedagogy. Perspect Psychol Sci 11:280–295. https://doi.org/10.1177/1745691615621276

Hirata S, Fuwa K, Sugama K, Kusunoki K, Fujita S (2010) Facial perception of conspecifics: chimpanzees (*Pan troglodytes*) preferentially attend to proper orientation and open eyes. Anim Cogn 13:679–688. https://doi.org/10.1007/s10071-010-0316-y

Hoicka E (2016) Parents and toddlers distinguish joke, pretend and literal intentional contexts through communicative and referential cues. J Pragmat 95:137–155. https://doi.org/10.1016/j.pragma.2015.10.010

Kaminski J, Schulz L, Tomasello M (2012) How dogs know when communication is intended for them. Dev Sci 15:222–232. https://doi.org/10.1111/j.1467-7687.2011.01120.x

Kano F, Call J (2014a) Cross-species variation of gaze following and conspecific preference among great apes human infants adults. Anim Behav 91:137–150. https://doi.org/10.1016/j.anbehav.2014.03.011

Kano F, Call J (2014b) Great apes generate goal-based action predictions: an eye-tracking study. Psychol Sci 25:1691–1698. https://doi.org/10.1177/0956797614536402

Kano F, Tomonaga M (2009) How chimpanzees look at pictures: a comparative eye-tracking study. Proc Roy Soc B 276:1949–1955. https://doi.org/10.1098/rspb.2008.1811

Kano F, Hirata S, Call J, Tomonaga M (2011) The visual strategy specific to humans among hominids: a study using the gap-overlap paradigm. Vis Res 51:2348–2355. https://doi.org/10.1016/j.visres.2011.09.006

Kano F, Call J, Tomonaga M (2012) Face and eye scanning in gorillas, orangutans, and humans: unique eye-viewing patterns in humans among hominids. J Comp Psychol 126:388–398. https://doi.org/10.1037/a0029615

Kano F, Hirata S, Call J (2015) Social attention in Pan: bonobos exhibit more eye contacts than chimpanzees. PLoS One 10:e0129684. https://doi.org/10.1371/journal.pone.0129684

Kano F, Krupenye C, Hirata S, Call J (2017) Eye tracking uncovered great apes' ability to anticipate that other individuals will act according to false beliefs. Commun Integr Biol 10:e1299836. https://doi.org/10.1080/19420889.2017.1299836

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

Krupenye C, Kano F, Hirata S, Call J, Tomasello M (2016) Great apes anticipate that other individuals will act according to false beliefs. Science 354:110–114. https://doi.org/10.1126/science.aaf8110

Liebal K, Call J (2012) The origins of non-human primates' manual gestures. Philos Trans R Soc Lond B Biol Sci 367:118–128. https://doi.org/10.1098/rstb.2011.0044

Lyn H, Russell JL, Hopkins WD (2010) The impact of environment on the comprehension of declarative communication in apes. Psychol Sci 21:360–365. https://doi.org/10.1177/0956797610362218

Miklósi Á, Polgárdi R, Topál J, Csányi V (1998) Use of experimenter-given cues in dogs. Anim Cogn 1:113–121. https://doi.org/10.1007/s100710050016

Moore R (2016) Meaning and ostension in great ape gestural communication. Anim Cogn 19:223–231. https://doi.org/10.1007/s10071-015-0905-x

Moore R (2017) Gricean communication and cognitive development. Philos Q 67:303–326. https://doi.org/10.1093/pq/pqw049

Moore R, Liebal K, Tomasello M (2013) Three-year-olds understand communicative intentions without language, gestures or gaze. Interact Stud 14:62–80. https://doi.org/10.1075/is.14.1.05moo

Moore R, Call J, Tomasello M (2015a) Production and comprehension of gestures between orang-utans (Pongo pygmaeus) in a referential communication game. PloS One 10:e0129726. https://doi.org/10.1371/journal.pone.0129726

Moore R, Mueller B, Kaminski J, Tomasello M (2015b) Two-year-old children but not domestic dogs understand communicative intentions without language, gestures or gaze. Dev Sci 18:232–242. https://doi.org/10.1111/desc.12206

Myowa-Yamakoshi M, Tomonaga M, Tanaka M, Matsuzawa T (2003) Preference for human direct gaze in infant chimpanzees (*Pan troglodytes*). Cognition 89:113–124. https://doi.org/10.1016/S0010-0277(03)00071-4

Okamoto-Barth S, Call J, Tomasello M (2007) Great apes' understanding of other individuals' line of sight. Psychol Sci 18:462–468. https://doi.org/10.1111/j.1467-9280.2007.01922.x

Parr LA, Dove T, Hopkins WD (1998) Why faces may be special: evidence of the inversion effect in chimpanzees. J Cogn Neurosci 10:615–622. https://doi.org/10.1162/089892998563013

Riedel J, Schumann K, Kaminski J, Call J, Tomasello M (2008) The early ontogeny of human–dog communication. Anim Behav 75:1003–1014. https://doi.org/10.1016/j.anbehav.2007.08.010

Scott-Phillips TC (2015a) Meaning in animal and human communication. Anim Cogn 18:801–805. https://doi.org/10.1007/s10071-015-0845-5

Scott-Phillips TC (2015b) Nonhuman primate communication, pragmatics, and the origins of language. Curr Anthropol 56:56–80. https://doi.org/10.1086/679674

Senju A, Csibra G (2008) Gaze following in human infants depends on communicative signals. Curr Biol 18:668–671. https://doi.org/10.1016/j.cub.2008.03.059

Senju A, Csibra G (2014) A commentary by Csibra, Senju, et al. on gaze following. http://cognitionandculture.net/blog/icci-blog/for-the-record-a-commentary-by-csibra-senju-et-al-on-gaze-following. Retrieved 1 Jul 2018

Sperber D, Wilson D (1995) Relevance: communication and cognition, 2nd edn. Blackwell, Oxford

Szufnarowska J, Rohlfing KJ, Fawcett C, Gredebäck G (2014) Is ostension any more than attention? Sci Rep 4:5304. https://doi.org/10.1038/srep05304

Téglás E, Gergely A, Kupán K, Miklósi Á, Topál J (2012) Dogs' gaze following is tuned to human communicative signals. Curr Biol 22:209–212. https://doi.org/10.1016/j.cub.2011.12.018

Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. Philos Trans R Soc B Biol Sci 364:2405–2415. https://doi.org/10.1098/rstb.2009.0052

Tomasello M (1999) The cultural origins of human cognition. Harvard University Press, Cambridge

Tomasello M (2006) Why don't apes point? In: Enfield N, Levinson SC (eds) Roots of human sociality: culture, cognition, and interaction, vol 197. Berg, London, pp 506–524

Tomasello M, Hare B, Lehmann H, Call J (2007) Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. J Hum Evol 52:314–320. https://doi.org/10.1016/j.jhevol.2006.10.001

Tomonaga M (1999) Inversion effect in perception of human faces in a chimpanzee (*Pan troglodytes*). Primates 40:417–438. https://doi.org/10.1007/BF02557579

Topál J, Gergely G, Erdőhegyi Á, Csibra G, Miklósi Á (2009) Differential sensitivity to human communication in dogs wolves human infants. Science 325:1269–1272. https://doi.org/10.1126/science.1176960

Topál J, Kis A, Oláh K (2014) Dogs' sensitivity to human ostensive cues: a unique adaptation. In: Kaminski J, Marshall-Pescini S

(eds) The social dog: behavior and cognition. Elsevier, New York, pp 319–346

Ueno A et al (2010) Brain activity in an awake chimpanzee in response to the sound of her own name. Biol Lett 6:311–313. https://doi.org/10.1098/rsbl.2009.0864

Virányi Z, Gácsi M, Kubinyi E, Topál J, Belényi B, Ujfalussy D, Miklósi Á (2008) Comprehension of human pointing gestures in young human-reared wolves (*Canis lupus*) and dogs (*Canis familiaris*). Anim Cogn 11:373. https://doi.org/10.1007/s10071-007-0127-y

Vorms M (2012) A-not-B errors: testing the limits of natural pedagogy theory. Rev Philos Psychol 3:525–545. https://doi.org/10.1007/s13164-012-0113-4

Whiten A, McGuigan N, Marshall-Pescini S, Hopper LM (2009) Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. Philos Trans R Soc Lond B Biol Sci 364:2417–2428. https://doi.org/10.1016/j.anbehav.2007.08.010

Yamagiwa J (1992) Functioinal analysis of social staring behavior in an all-male group of mountain gorillas. Primates 33:523–544. https://doi.org/10.1007/BF02381153

Yamamoto S, Humle T, Tanaka M (2012) Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. Proc Natl Acad Sci 109:3588–3592. https://doi.org/10.1073/pnas.1108517109