



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Environmental Influences on Behavioural Responses on the Object Choice
Task in Dogs and Children: Implications for the Evolution of Human
Cognition

Hannah Clark

Thesis submitted for the Degree of Doctor of Philosophy

University of Sussex

July 2019

The thesis conforms to an ‘article format’ in which the middle chapters consist of discrete articles written in a style that is appropriate for publication in peer-reviewed journals in the field. The first and final chapters present synthetic overviews and discussions of the field and the research undertaken.

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another university for the award of any other degree. However, this thesis incorporates, to the extent indicated below, material already submitted as required coursework and/ or for the degree of:

Master of Research

Psychological Methods

which was awarded by the University of Sussex

Some of the data for Paper 1 was collected for the research internship project of the degree above but was extended on and reanalysed before being incorporated into this thesis.

Signature:

July, 2019

University of Sussex

Hannah Clark

Thesis submitted for the degree of Doctor of Philosophy

Environmental Influences on Behavioural Responses on the Object Choice Task in

Dogs and Children: Implications for the Evolution of Human Cognition

Several prominent theories claim that humans possess specialised socio-cognitive skills that emerge early in ontogeny and are attributable to selective histories not shared by other hominids. The generally poor performance of nonhuman primates on tasks designed to measure skills of joint attention, such as the Object Choice Task (OCT), is a key evidence base for these theories. Similar claims are made for domestic dogs, whose apparent proficiency in following pointing cues is taken as evidence for the idea that domestication selected for advanced socio-cognitive skills. The aim of this thesis is to extend current reviews which highlight methodological differences in OCT studies that compare across species, by first analysing the prevalence and the effects of such confounds in the existing literature, and then by conducting a series of systematic experiments to further investigate how these differences may affect behavioural responses. This thesis presents six papers, beginning with a meta-analysis of 71 published OCT studies with nonhuman primates and dogs, which finds between-groups differences in subjects' pre-experimental histories, the use of a barrier in the form of a test cage, and the types of cues presented. Effects of performance associated with these differences were also found. Following this are four papers which present the results of empirical studies manipulating the presence of a barrier in the testing environment and the spatial configuration of the test set-up distinguishing between central and peripheral versions of the task with human children and dogs. For both taxonomic groups, there were differences in the behavioural responses elicited, and some performance effects for dogs. A further paper finds no effect of manipulating food reward type in an OCT with children. This demonstrates that failing to match testing environments represents an experimental confound and challenges the validity of phylogenetic theories based on studies which use incommensurate test protocols.

For Aina: “*Tot és per a tu*”

Acknowledgements

First, I'd like to express my immense gratitude to my supervisor, Dave, for all of your continuing support over the last six years and for going above and beyond my expectations of a supervisor. You've provided invaluable guidance which has enabled me to improve my statistical, writing, and critical thinking skills and have inspired me with your commitment and dedication to scientific integrity. Thank you also to Nicola and Jessica for your support and advice and to Zoe Flack and Kassandra Giragosian for inter-rater reliability coding.

I'd also like to thank everyone who volunteered their time to participate in these studies, and Milton Village Hall for being so accommodating.

Thanks to my mum and Lloyd for supporting me in innumerable ways, to Emma, for so often making the logistics of doing a PhD with a young child much easier, to my nan and grandad and Steve and Karen for all your love and support over the years. Thanks to my friends, especially Natalie and Khaleda, for moral support.

Finally, thanks to Isma, for always sticking to your promise.

Contents

Introduction	12
Joint Attention and Social Cognition	14
The Object Choice Task and Related Theories	16
Nonhuman primates	17
Dogs	20
Criticisms of the Object Choice Task Literature	21
Pre-Experimental History	22
Barriers.....	25
Central vs. Peripheral Versions	25
Point Types	26
The Current Thesis.....	27
Summary of Thesis Chapters.....	28
Paper 1	28
Paper 2	29
Paper 3	30
Paper 4	30
Paper 5	31
Paper 6	32

Summary of Current Research.....	32
Paper 1: Ontogeny vs. Phylogeny in Primate/Canid Comparisons: A Meta-analysis of the Object Choice Task	33
Abstract	34
Introduction	36
Methods	42
Literature Selection.....	42
Subjects.....	43
Human Experience Scale	44
Cue Types	44
Individual Performance Data	45
Results	46
Human Experience.....	46
Human Experience and Performance	47
Presence of a Barrier	51
Cue Types	52
Multiple Conditions	55
Inter-Object Distance	56
Discussion	57
Table S1.....	61
Table S2.....	70
Paper 2: Changes in the Referential Problem Space of Infants and Toddlers.....	71

Abstract	72
Introduction	73
Study 1: Barrier vs. No Barrier.....	76
Method.....	76
Participants.....	76
Procedure.....	76
Materials.....	78
Results.....	80
Correct Choices.....	80
Response Latency.....	81
Response Type.....	82
Discussion.....	87
Study 2: Manipulating the Distance of the Reward	90
Method.....	91
Participants.....	91
Procedure.....	91
Results.....	93
Correct Choices.....	93
Response Latency.....	93
Response Type.....	93
Discussion.....	95
General Discussion.....	97

Paper 3: The Performance of 3-Year-Olds on Two Versions of the Object Choice

Task	100
Abstract	101
Introduction	102
Method.....	106
Participants.....	106
Procedure.....	106
Results	109
Children Tested with a Barrier.....	109
Children Tested without a Barrier.....	110
Barrier vs. No Barrier Comparisons.....	111
Discussion	114

Paper 4: Children's Responses to Different Reward Types on an Object Choice

Task	117
Abstract	118
Introduction	119
Method.....	123
Results	127
Discussion	130

Paper 5: Testing Dogs in Ape-Like Conditions: The Effect of a Barrier on Dogs'

Performance on the Object Choice Task	133
Abstract	134
Introduction	135

Study 1: Barrier vs. No Barrier Within-Subjects.....	138
Method.....	138
Results.....	141
Discussion.....	143
Study 2: Barrier vs. No Barrier Between-Subjects.....	143
Method.....	143
Results.....	144
Discussion.....	146
General Discussion.....	146
Paper 6: The Performance of Domestic Dogs (<i>Canis familiaris</i>) on Two Versions of the Object Choice Task.....	154
Abstract	155
Introduction	156
Methods	160
Subjects.....	160
Procedure	161
Materials and Set-up	162
Results	164
Dogs Tested with a Barrier	164
Dogs Tested without a Barrier	165
Barrier vs. No Barrier Comparisons	167
Discussion	170

Thesis Discussion.....	175
Summary of Papers.....	176
Pre-experimental History	178
Barriers	179
Central vs. Peripheral Versions of the OCT	180
Point Type	181
Other Findings	182
Reward Types	182
Age.....	183
Limitations.....	184
Sampling	184
Best Practices for Future Research.....	185
Future Directions	186
Theoretical Implications.....	187
References	189

Introduction

Hannah Clark

University of Sussex

Introduction

Spoken language provides us with the ability to share with others the most complex of ideas, to create vivid fantasies about imaginary beings and to communicate our inner-most thoughts. This ability is unique to humans and the driving causes for its evolution in our species but not that of our nearest living relatives, the great apes, is the subject of much interest and debate. Corballis (1999) suggested a gestural origin for the evolution of language, using as evidence nonhuman primates' superior cortical control over their hands compared with vocalisations, which, he argues, may explain why attempts to teach apes sign language have been more successful than attempts to teach them spoken language. Early hominids' move to bipedalism would have meant increased availability of the hands as a tool for communication, and life on the open savannah may have promoted the use of gestures over vocalisations in order to silently communicate about the locations of predators or food sources. Arbib, Liebal & Pika (2008) also argue that gestures preceded speech and suggest that increased manual dexterity coupled with the capacity to imitate complex goal-directed actions led to the ability to produce complex syntactical constructions, which are missing from the utterances produced by language-trained apes. For these gestural origins theories of language, the evolution of human social cognition is inextricably linked to the challenges and opportunities presented by the gestural communication of hunters and foragers in savannah environments. The extent, then, to which selection has led to the capacity to comprehend deictic cues is a widely researched area, and debate continues concerning both the abilities of nonhuman species, and the explanations of the hypothetical underlying processes that enable these abilities. Particularly contentious is the research surrounding dogs and the extent to which domestication has selected for specialised socio-cognitive skills; and also research with nonhuman primates with some

claiming that some joint attentional skills that facilitate social communication are human-unique amongst primates.

Joint attention and Social Cognition

Joint attention refers to the coordination of one's attention to another person and an object or event of interest (Leavens & Clark, 2017). It is comprised of three primary facets, Initiating Behaviour Regulation (IBR), Initiating Joint Attention (IJA) and Responding to Joint Attention (RJA). IBR refers to the use of gestures and/ or gaze to manipulate the behavioural response of a social partner to obtain a desired object or event (Mundy et al., 2007), for example, raising their arms to request being picked up (Flack & Leavens, 2017) or pointing to request an out-of-reach object of desire. This ability emerges towards the end of the first year of life in most Western cultures (Leavens & Racine, 2009). IJA is the use of gaze and/or gestures to direct the attention of another to an interesting object or event, or to themselves (Mundy, 2013) and begins to develop from around 10 months of age (Leavens & Racine, 2009). RJA is the receptive form and refers to the ability to follow a social partner's gaze or gesture to an interesting object or event (Mundy et al., 2007).

Bates, Camaoni and Volterra (1975) distinguished between two types of pointing gesture used by preverbal infants which differ in motive. Proto-imperatives involve the 'use of the adult as a means to a desired object' (p. 209), whereas proto-declaratives are 'the use of an object... as a means to obtaining adult attention' (p.209). Leavens (2012) notes that, in this original definition, both point types were described as instrumental in their nature, functioning to elicit a desired response from an adult. Specifically, he emphasises that the 'adult attention' that proto-declaratives served to elicit referred to observable, behavioural responses to the child, such as laughing,

smiling, making eye contact or commenting on the child's point of reference. Leavens (2012) describes the trajectory through which Bates et al.'s (1975) original point types classification developed such that proto-declaratives are now widely accepted to encompass a suite of advanced and human species-specific socio-cognitive skills. Evidence of a lack of declarative pointing in autistic children as well as correlations between IJA and RJA—but not IBR—with both expressive and receptive language in autism (e.g. Dawson, Toth, Abbott, Osterling & Munson, 2004) led to suggestions that underlying the use of proto-declaratives is a complex understanding of others as social agents motivated by unseen mental processes. The extent to which nonhuman primates, particularly our nearest relatives, the great apes, exhibit joint attentional skills is also the subject of wide debate. In terms of pointing production, it is now widely accepted that great apes do produce points in certain ecological environments (Leavens & Racine, 2009). Pointing comprehension is somewhat more controversial, and although there is evidence of apes successfully passing tests of RJA, debate continues as to the underlying cognitive processes that enable them to do so.

As Leavens (2012) pointed out, Mundy (1987) equated the use of proto-declaratives in human children, both receptively and expressively, with perspective-taking; that is, the understanding that others have a different perspective from one's own. Baron-Cohen (1995; 1999) suggested that proto-declaratives are points serving to change another's mind, and Petitto (1988) suggested they were evidence of an acknowledgment of the mutuality of joint attention. Tomasello (1995) argued that the emergence of IJA and RJA behaviours is evidence of the human infant's increasing awareness and understanding of others as intentional agents. Around one year of age, infants begin to coordinate their gaze-following behaviours with looks that apparently serve to check that the adult's attention remains focused on the shared object of interest

(Butterworth, 1991). In addition, their own points are accompanied by gaze alternation between the social partner and the object, and they begin to participate in longer joint attention interactions with the caregiver. Such behaviours, suggested Tomasello (1995) are manifestations of awareness of the social partner as an individual separate from themselves with attentional states and intentions different from their own, and together with theory-of-mind skills which emerge later, contribute to the child's developing understanding of others.

A competing theoretical perspective suggests that it is not necessary to draw on such complex, hypothetical mental processes to explain the emergence of expressive and receptive pointing in human infants. Both the understanding and production of pointing gestures in human infants is a developmental process, one which occurs in an environment in which infants experience rich and prolonged human interaction. A learning-based account (e.g. Moore & Corkum, 1994) suggests that in such an environment, human infants may learn that pointing elicits a response from adults, be that a comment, a smile, or some other form of attention. This harks back to Bates et al.'s original description of proto-declaratives as being instrumental in eliciting an attentional response from the caregiver, and as Leavens (2012) pointed out, requires no complex socio-cognitive skills beyond simple learning of stimulus-response connections.

The Object Choice Task and Related Theories

The Object Choice Task (OCT) is an experimental paradigm used to assess RJA capacity and tests an individual's ability to follow an experimenter's gaze and/ or gestural cue to find a hidden reward. Typically, it involves an experimenter baiting one of two or three opaque containers and then signalling, using, for example, a declarative

pointing cue, the location of the hidden reward to the subject. The subject is classified as having made a correct choice if they approach or touch the container to which the experimenter signalled. Since its first use in the 1990s in which Anderson, Sallaberry & Barbier (1995) investigated the ability of capuchin monkeys to comprehend pointing, the OCT has been widely used with a variety of species, including elephants (Smet & Byrne, 2013), horses (Maros, Gácsi & Miklósi, 2008; Proops, Walton & McComb, 2010), dolphins (Pack & Herman, 2004), bats (Hall, Udell, Dorey, Walsh & Wynne, 2011), goats (Kaminski, Riedel, Call & Tomasello, 2011), pigs (Albiach-Serrano, Bräuer, Cacchione, Zickert & Amici, 2012; Nawroth, Ebersbach & von Borrell, 2013) and seals (Scheumann & Call, 2004). By far the most widely tested taxonomic groups are, however, nonhuman primates and domestic dogs (see Clark, Elsherif & Leavens, 2019, and Krause, Udell, Leavens & Skopos, 2018, for comprehensive histories of pointing comprehension and production, respectively, in nonhuman species).

Nonhuman primates.

Of the nonhuman primates, the most frequently studied on the OCT are the great apes, and a number of studies report that they tend to perform poorly (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Itakura, Agnetta, Hare & Tomasello, 1999; Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012; Tomasello, Call & Gluckman, 1997). On the basis on such results, several theories have been proffered regarding the differential adaptive evolutionary histories of humans and the rest of the hominid species, including Shared Intentionality, the Vygotskian Intelligence Hypothesis, and the Cultural Intelligence Hypothesis.

Tomasello and Carpenter (2007) suggested that the reason that apes demonstrate such seemingly poor performance on the OCT is because they fail to comprehend the

“joint attentional frame” or “common ground” that the task requires, that is, a shared understanding that the reward is hidden in one of the containers and that the experimenter’s cue is one which refers to the location of the reward. Tomasello (1999) suggested that this ability for shared intentionality- the sharing of psychological states with another (Tomasello & Carpenter, 2007)- is, along with the ability to understand others as intentional agents, a human-unique ability that emerges early in ontogeny. Thus, according to this theory, ape subjects see a pointing cue on the OCT and, although they are able to follow the cue’s directionality (Tomasello & Carpenter, 2007), they fail to comprehend that the point is intended to reveal the location of the hidden reward and therefore fail to make a correct choice.

Hare (2001) suggested that chimpanzee performance on socio-cognitive tasks would increase if such tasks were competitive in nature, the so-called “competitive cognition hypothesis.” In order to test this theory, Hare and Tomasello (2004) conducted a study with chimpanzees in which they compared their performance on a typical OCT (labelled a “cooperative” condition) with a “competitive” version. This latter condition involved an experimenter reaching to the location of hidden food in an attempt to obtain it for himself, as opposed to pointing to the location in an attempt to inform the subject of its whereabouts. They found increased performance in the “competitive” condition and it was therefore suggested that because chimpanzees would not, in a natural setting, experience conspecifics informing them of the location of food, they therefore do not interpret a gestural cue on the OCT as informative in nature and for this reason often fail to use it to find the hidden reward (Moll & Tomasello, 2007). Moll and Tomasello (2007) argued that human children, however, as well as being able to understand others as intentional beings in competitive contexts (as do chimpanzees) are also able to successfully participate in cooperative contexts, such as those involving

shared intentionality, for example, joint attention (see Leavens, 2012 for a critique of interpretations of humans and apes' respective abilities regarding joint attention). Once an infant begins to comprehend shared intentionality, argued Tomasello (1999), they begin a process of cultural development, in which their understanding of the world is mediated through that of others. Moll and Tomasello (2007), owing to the importance Vygotsky placed on interactions of a cooperative nature in the development of cognitive skills, termed this the Vygotskian intelligence hypothesis. This states that shared intentionality in the context of cooperation is a human-unique trait that emerged after our separation from the rest of the hominid lineage.

Extending the idea of the importance of culture to the development of human socio-cognitive skills in the Vygotskian intelligence hypothesis, proponents of the cultural intelligence hypothesis suggest that individuals should learn both universal skills, and those specific to their culture, faster through social learning, and that, as such, those with more opportunities for social learning should acquire more skills (Whiten & van Schaik, 2007; van Schaik & Burkart, 2011). Human infants, born into specific cultural groups, must quickly learn the cultural practices of that group, such as their use of artefacts, their language and social practices, in order to function within that group, and must therefore possess specialised cognitive skills that facilitate social interactions such as communication and theory of mind (Herrmann, Call, Hernandez-Lloreda, Hare & Tomasello, 2007). In a test of the cultural intelligence hypothesis, Herrmann et al. (2007) assessed 105 humans, and 138 apes (orangutans and chimpanzees) on a battery of social and physical tests, hypothesising that, early in ontogeny, before exposure to human-specific advanced training (e.g., formal education), children would have similar physical cognition to the apes, but advanced social skills due to the early emergence of human-specific socio-cognitive abilities. The OCT

formed a part of the battery of social tests and, based on superior performance by the children on this task (among others), the authors concluded that their theory was supported, and as such, humans have evolved specialised species-unique skills in understanding others.

Dogs.

Soon after the first OCT studies with apes, research with dogs on this paradigm began (Miklósi, Polgárdi, Topál & Csányi, 1998), and a very different picture with respect to performance was found. Dogs were consistently shown to be adept at following human pointing cues (Hare & Tomasello, 1999; Miklósi, Polgárdi, Topál & Csányi, 1998) and this led to the development of the domestication hypothesis (Hare, Brown, Williamson & Tomasello, 2002; Hare & Tomasello, 2005). Proponents of the domestication hypothesis suggest that through a process of convergent evolution resulting from years of domestication, dogs have evolved specialised socio-cognitive skills that endow them with the ability to comprehend human gestural cues. Studies which apparently showed that even very young puppies were able to follow human-given cues (Hare et al., 2002; Riedel, Schumann, Kaminski, Call & Tomasello, 2007) led to the suggestion that such abilities were solely a result of the species' history of artificial selection (Hare et al. 2002; Gácsi, McGreevy, Kara & Miklósi, 2009).

However, comparisons of dogs and other canids demonstrate that wolves can perform at comparable rates on the OCT when matched in pre-experimental history, specifically when human-raised and thus exposed to the intense social interaction that pet dogs experience (Udell, Dorey & Wynne, 2008; Udell, Spencer, Dorey & Wynne, 2012). In addition, dogs raised with alternative life histories to those experienced by pet dogs, for example, in shelters, research facilities, or kennels, show much poorer performance on the OCT than typical pet dogs, particularly with regard to pointing cues

that seem to be more difficult (Udell, Dorey & Wynne, 2010; D’Aniello et al., 2017; Lazarowski & Dorman, 2015). This, taken together with their rebuttal of puppy studies that claimed to demonstrate the lack of an ontogenetic role in dogs’ ability to follow points (see Udell & Wynne, 2010, for a discussion of a number of methodological and procedural issues in such research), led Udell, Dorey and Wynne (2010; 2012) to argue for the importance of both phylogeny and ontogeny in the development of socio-cognitive skills in dogs. Their two-stage hypothesis suggests that the process of domestication has increased the likelihood of acceptance of humans as social companions by dogs (phylogeny), and that acceptance of humans as social companions during a sensitive period coupled with the opportunity to learn associations between human stimuli and certain consequences of those stimuli (ontogeny) leads to an increased likelihood of successfully using human stimuli (e.g. gestural cues) on the OCT. That is, domestication may have led dogs to be more accepting of humans socially but, in order to effectively comprehend their communicative cues, there is an important role of learning that takes place through exposure to humans and human behaviour.

Criticisms of the OCT Literature

Leavens, Bard and Hopkins (2017) discussed the tendency in comparative psychology literature to attribute differences in performance by apes and humans on socio-cognitive tasks to discontinuity in the hominid lineage that has endowed humans with species-unique cognitive skills in understanding others. Such theories, they argue, are not well justified by the empirical evidence that claims to support them due to a number of procedural and methodological confounds which arise from failing to sufficiently match key selection and procedural variables across groups, and the proliferation of rich interpretations regarding hypothetical psychological processes in

explanations of human responses (see also Leavens 2012; Leavens, 2018). The OCT literature is by no means above such criticism and several recent reviews have highlighted some of the issues pertaining to this area, including lack of matching groups for pre-experimental history and differences in the experimental set-up and cue types presented across taxonomic groups.

Pre-Experimental History

Discussed above were the findings that dogs with alternative life histories— that is, those not raised in human homes as pets and thus lacking the intense exposure to humans that pet dogs experience— show much lower levels of competence in following human pointing cues, compared to typical pet dogs. Similar patterns are found when researchers examine the performance shown by apes with varying pre-experimental histories. Call and Tomasello (1994), Lyn (2010), Lyn, Russell and Hopkins (2010), and Russell, Lyn, Schaeffer and Hopkins (2011) all reported that apes who have been raised in environments rich in human social interaction, that is, who are “enculturated”, are adept at using pointing cues on the OCT. In the OCT literature however (and indeed other fields in comparative psychology, see Leavens et al., 2017), there is a tendency towards taking a species’ performance at the group level without consideration for the differential life experiences that may be present within that group and generalising from the groups’ performance to the species as a whole (Leavens, 2014). Henrich, Heine and Norenzayan (2010) discussed the tendency of researchers to generalise from human samples taken from Western, industrialised societies, under the assumption that these populations are representative of the species as a whole, and Leavens, Bard and Hopkins (2010) noted similar overgeneralisations in ape samples. The majority of studies in comparative cognition, they argued, use samples of “institutionalised” apes, that is, apes living in captivity, and they noted that these samples are not representative

of the species as whole. Like humans, apes adapt to their ecological environment, and so different cognitive adaptations will be observed according to the environment the individual experiences. They highlighted differences between enculturated, institutionalised, and wild populations of apes in terms of the communicative behaviour they have been observed to display. Of particular interest is the differential instances of pointing production: Wild populations have only very rarely been observed to point, however observation of this behaviour is much more common in institutionalised individuals, and enculturated (that is, home-raised or language-trained) apes all use pointing gestures (Leavens & Bard, 2011). Thus, it appears that the ability to produce and comprehend these human gestural cues is a cognitive skill emerging as an adaptation to an ecological environment, specifically, one in which being able to comprehend human gestures is useful for individuals in their everyday lives. It does not, therefore, as Leavens et al. (2010) argued, make sense to generalise from opportunity samples of (largely) institutionalised apes and attribute their performance to such factors as genotypic make-up of the species, as a number of OCT studies have done. Leavens and Clark (2017) listed 43 individuals from great ape species who have successfully used pointing gestures on the OCT, showing that point-following is well within the cognitive repertoire of apes if given sufficient pre-experimental exposure to humans and their communicative behaviour.

Bräuer, Kaminski, Riedel, Call & Tomasello (2006) compared dogs and two species of ape (chimpanzees and bonobos) on different behavioural and communicative cues on the OCT and concluded that the dogs' superior performance provided evidence for the domestication hypothesis, and that the apes' poorer performance on such cues but superior performance using physical cues (e.g. auditory cues) was the result of species adaptations relating to foraging abilities. Their samples, however, constituted a

group of institutionalised apes compared with a group of dogs living as pets in family homes. Herrmann et al.'s (2007) study, which they claimed provided support for the cultural intelligence hypothesis, compared 2.5-year-old children with two samples of chimpanzees and a sample of orangutans. The ape samples constituted individuals living in sanctuaries, many of whom would have spent their early years in the wild, being later housed in sanctuaries as a result of the destruction of their habitat through deforestation, or the traumatic obliteration of their social group by hunters. This demonstrates a failure to match the pre-experimental opportunities for learning through exposure to human communicative cues. Such differences then, cannot be reliably attributed to genotypic differences between the species, but would be more parsimoniously explained through differential cognitive adaptations between the samples to the different ecological demands of their different rearing environments. More recently, Maclean, Herrmann, Suchindran and Hare (2017) reanalysed Herrmann et al.'s (2007) data from human children and apes and compared it with the performance of a large sample of dogs on the same battery of tests. They concluded that the dogs and the human children were comparable in their skills of cooperative communication suggesting specialised abilities in dogs in this area that are not present in the ape species. Their dog samples, however, were comprised of assistance dogs, explosive-detecting dogs and pet dogs, all of which experience intense human socialisation (even more so, it could be argued, in the case of the first two samples who not only live but work in close contact with humans) which the apes in Herrmann et al.'s study had not experienced.

Barriers

In addition to failing to match samples for pre-experimental history, Leavens et al. (2017) discussed the prevalence of incommensurate testing environments when

comparing across species. Nonhuman primates tend to take part in studies from within cages (either their home cages or special testing cages), thus inducing the necessity to both observe and respond through wire mesh caging. This imposes a partial visual barrier between the subject and the experimenter and test apparatus, through which the subject must observe the cue being presented and requires the subject to choose a container by indicating through wire mesh or by reaching through plexiglass holes. This is in stark contrast to the testing conditions experienced by representatives of other taxonomic groups, such as dogs and human children, who are tested without such a visual barrier and are freely able to interact with the testing apparatus. Leavens et al. (2017) argued that we cannot assume such a barrier does not affect individuals' responses (see Udell et al., 2008, and Kirchhofer et al., 2012, for exceptional application of barriers to the OCT in canids and the detrimental effects on performance observed).

Central vs. Peripheral Versions of the OCT

Differences, too, in the spatial configuration of the test set-up have been implicated in explanations of performance differences between different taxa. Mulcahy and Hedge (2012) reviewed published OCT studies and found a bias towards more centralised testing for apes, and more peripheral testing for dogs. The distinction between the two relates to the placement of a) the containers in relation to each other and the subject, and b) the experimenter in relation to the containers and the subject. Central versions involve the containers being placed close together, in the direct line of vision of the subject, and between the subject and the experimenter. Peripheral versions involve the containers being placed further apart, out of the direct line of vision of the subject, and peripheral to the subject and experimenter. Mulcahy and Hedge (2012) argued that the spatial configuration used could influence the salience of the cue, with decreased salience when containers are placed in between the subject and experimenter

due to the presence of the food reward distracting the subject. For example, in Bräuer et al.'s comparison of dogs and apes, the inter-object distance for the dogs was 130cm whereas, for the apes, it was 50cm. In Herrmann et al.'s (2007) study, the inter-object distance for apes was 64cm, whereas Maclean et al. (2017), who made direct comparisons with the apes in Herrmann et al. (2007), used an inter-object distance of 200cm with their dog subjects. In both cases, the dogs were tested with experimental set-ups so different from the ape subjects that the measurement between containers was more than double, yet performance differences were attributed to phylogenetic theories with no consideration of the possible effects of such mismatched testing environments.

Point Types

Miklósi and Soproni (2006), in a review of OCT studies, defined and described the wide variety of pointing cues that have been presented to representatives of various species on the OCT, and discussed the differential performance of representatives of those species according to the type of cue presented. Their descriptions distinguished between points in terms of lateral, distal and temporal properties. For lateral properties, points were categorised as being *ipsilateral*, that is, performed with the hand nearest to the container, or *contralateral*, that is, performed with the hand on the contralateral side of the body in relation to the container. They also classified points as being either *proximal* (40cm or less between fingertip and container) or *distal* (more than 40cm between fingertip and container). Finally, they distinguished between temporal features of cues such that points were categorised as either *static*, that is, the pointing gesture is enacted before the subject enters the test room and remains in position until a choice is made; *momentary*, where the pointing gesture is enacted in front of the participant and is held for around 2 seconds before the pointing hand is lowered to the side of the body; or *dynamic*, in which the pointing cue is enacted in front of the participant and is held

until a choice is made. They argued that different cue types have different levels of salience and may necessitate more demands on memory. Udell, Hall, Morrison, Dorey and Wynne (2013) found that dogs' performances were affected by point topography, which further highlights the need for ensuring that subjects are a) tested with a range of point types before conclusions are made as to their relative abilities in point following; and b) that, when comparing across species, the pointing cues are matched.

The Current Thesis

The aim of the current thesis is to address some of the above-mentioned criticisms of the OCT research, by providing an in-depth analysis of their prevalence and effects. Whilst a number of existing reviews (e.g. Lyn, 2010; Mulcahy & Hedge, 2012) have focused on one specific area of contention, or on confounds in the testing of representatives of one taxonomic group, to date there does not exist any research which synthesises these across both nonhuman primates and dogs. Here, I present the findings of a meta-analysis which identifies the prevalence and the effects on performance of the aforementioned confounds, followed by a series of empirical studies in which the effects of barriers, configuration and point type are investigated in human children and dogs. The implications of the results are discussed with reference to theories of the evolution of social cognition which use OCT studies as the evidence base.

Summary of thesis chapters

Paper 1

Reviews of the OCT literature have identified a number of inconsistencies in studies that compare across species such as systematic differences in rearing history (Lyn, 2010), differences in spatial configuration (Mulcahy & Hedge, 2012), and differences in point types (Miklósi & Soproni, 2005). This paper extends on these

reviews to provide a comprehensive meta-analysis of 71 published dog and nonhuman primate OCT studies, firstly identifying the prevalence of a number of methodological and procedural inconsistencies across these taxa, and then using individual performance data to analyse their effects. A scale of human experience was developed to effectively categorise the different rearing histories reported in the literature, and the extent to which studies fail to match for pre-experimental history was highlighted. Across representatives from both taxonomic groups, a link between human experience and performance on the OCT was demonstrated. The extent to which the testing environments of dogs and nonhuman primates differ in terms of the presence of a barrier between a) the subject, and b) the experimenter and test apparatus, is also shown. The prevalence of the presentation of different cue types was analysed and systematic differences were found in both the types of pointing cues presented to dogs and nonhuman primates and the success rates associated with different cue types across representatives of both taxonomic groups. Finally, dogs were shown to be tested with greater inter-object distances than nonhuman primates, and correlations between inter-object distance and performance were significantly positive. This is the most comprehensive review of OCT studies to date and extends current knowledge of the prevalence of failing to match experimental conditions, analysing the effects of this in nonhuman primates and dogs. It also led to the development of a series of systematic experiments in which some of these variables are manipulated in studies with children, in order to further understand their effect on responses on the OCT.

Paper 2

Paper 1 showed that almost 100% of nonhuman primates are tested from within a test cage, thus imposing a barrier between the subject and both the experimenter and the test apparatus, compared with less than 1% of dogs. Analysis of any differences in

performance as a function of the presence of a barrier in Paper 1 were not possible due to a lack of data, but Kirchhofer et al. (2012) and Udell et al. (2008) both found performance differences associated with the presence of a barrier. Paper 2 presents two studies designed to investigate further the confounding effects that such a difference in testing environment may have. In Study 1, 18-month- and 36-month-old children were tested with and without a barrier on the OCT. These age groups were chosen because studies show that children effectively follow points from around 15 months, and so any behavioural differences observed could be attributed to the experimental manipulation rather than the lack of emergence of pointing comprehension skills. The study showed no effects on performance associated with a barrier, but there were differences in the children's behavioural responses, with an increase in communicative behaviour, that is, pointing to indicate the choice of container, when a barrier was present. This effect was significantly greater in the 18-month-old children. Study 2 compared 36-month-old children's behavioural responses when tested with a barrier, manipulating the distance of the containers, such that they were either within reach or not. This demonstrated that the children used more communicative behaviour when the container was out of reach, showing that they were making judgements about the accessibility of the containers. These studies are the first to test human children in a testing environment akin to that used with nonhuman primate subjects, and demonstrates that, even in human children, who are experts in the use of human communicative cues, the difference in testing environment had effects on their behavioural responses. This has important implications for studies which compare across species without matching for this possible confound.

Paper 3

Paper 1 found differences in the inter-object distances used when testing nonhuman primates and dogs, and performance differences associated with this. In

order to understand more fully the effects of this confound, Paper 3 presents a study which aimed to replicate as closely as possible Mulcahy and Call's (2009) central vs. peripheral task comparison with 36-month-old children. This was the first study to date in which human children were tested with a peripheral version of the OCT, as human children (similarly to nonhuman primates) are always tested with a central version (Clark, Elsherif & Leavens, unpublished data). Half of the children were tested with a barrier, in order to more closely match the testing conditions experienced by nonhuman primates, and half without, to provide a control to ensure that any differences found resulted from the spatial manipulation rather than the presence of a barrier. No effects were found in the central vs. peripheral analyses, but the children's behavioural responses differed as a function of the barrier, with a reduction in acts of direct prehension when the barrier was present. This provides further support for the findings in Paper 2.

Paper 4

When comparing between nonhuman primates and human children for differences in socio-cognitive skills, a number of tasks, including the OCT, involve providing rewards when an individual displays the "correct" response. One systematic (indeed, almost universal) difference in the treatment of representatives of the different species is the use of food rewards for nonhuman primates and toy or sticker rewards for human children. Paper 4 presents a study which aimed to investigate whether this difference impacts responses to the OCT by human children. 36-month-old children were tested on an OCT and were rewarded with food on half of the trials and with a sticker on the other half. No effects of reward type were found. This is the first study (as far as I am aware) which looked at the possible effect of manipulating reward type with

children and suggests that this experimental difference when testing across species may not have the same behavioural effects as other classes of confounds.

Paper 5

In Paper 1, a systematic difference in the scientific literature was found in the use of barriers between nonhuman primates and dogs, and in Paper 2, an effect of the barrier on children's behavioural responses on the OCT was demonstrated. Only one study to date (Kirchhofer et al., 2012) has looked at the effect of a barrier on dogs' performance with a pointing cue on the OCT, and found it was detrimental to success rates. The sample used in Kirchhofer's study may not have been representative however, because dogs were preselected for their tendency to retrieve (Hopkins, Russell, McIntyre & Leavens, 2013). Paper 5 presents two studies in which dogs were tested with and without a barrier on the OCT, the first using a within-subjects design, and the second a between-subjects design, with no pre-selection criteria for participation. Study 1 showed no difference in success rates as a function of the presence of a barrier, but in Study 2, those tested with a barrier failed to perform above chance. There were also differences in the types of incorrect responses, with an increase in no-choice responses associated with the barrier, whereas incorrect-choice responses were more typical in the no barrier condition. This shows that the presence of a barrier in the testing environment can affect both successful use of pointing cues, and behavioural responses to the task, and provides further support for the necessity of matching this aspect of the testing environment across species.

Paper 6

In Paper 1, significant differences in the spatial configurations used in presentations of the OCT to nonhuman primates and dogs were found, specifically with

regard to the inter-object distance, which was also found to be positively correlated with performance, extending on Mulcahy and Hedge's (2012) findings. In Paper 3, no significant effect of configuration was found between the central and the peripheral version of the task when testing human children. In Paper 6, I again investigated this manipulation, replicating as closely as possible Mulcahy and Call's (2009) ape study, this time with dogs, testing half of the dogs with a barrier (thus matching the conditions experienced by apes) and half without. For those tested without, there was no significant effect of configuration, but dogs that were tested with a barrier demonstrated decreased performance in the peripheral version but not in the central version. The presence of a barrier also had a suppressing effect on the dogs' tendency to choose a container. This shows the way in which environmental influences can interact to affect an individual's performance and behavioural responses to the OCT.

Summary of Current Research

The results presented in the current thesis demonstrate that barriers, spatial configuration and cue type affect the behavioural responses of human children, masters in the use of human deictic cues by a year and a half of age, and domestic pet dogs, who have consistently shown to be adept at using such cues also. This has implications for the wealth of studies which compare across species using incommensurate sampling and testing protocols, failing to match the testing conditions presented to representatives of different taxonomic groups, and demonstrates the need, in future research to ensure that testing protocols are matched, as closely as possible, before comparing performance across species. It also draws into question the validity of theories that appeal to species' phylogenetic histories in attempts to explain apparent differences in performance without regard for the numerous alternative explanations afforded by these confounds.

Paper 1

Ontogeny vs. Phylogeny in Primate/Canid Comparisons:

A Meta-analysis of the Object Choice Task

Hannah Clark¹, Mahmoud Elsherif ² and David A. Leavens¹

¹ University of Sussex

² University of Birmingham

A version of this paper is published in Neuroscience and Biobehavioural Reviews as:

Clark, H., Elsherif, M. M. & Leavens, D. A. (2019) Ontogeny vs. phylogeny in primate/canid comparisons: A meta-analysis of the object choice task

Parts of these data were presented at the 7th European Federation for Primatology Meeting, Strasbourg, France, 21-25 August, 2017.

Author Note

ME completed the original literature search. HC conducted the subsequent literature searches, conducted the analyses and wrote the manuscript. DAL provided advice on the analyses and comments on drafts.

Abstract

The Object Choice Task (OCT) is a widely used paradigm with which researchers measure the ability of a subject to comprehend deictic (directional) cues, such as pointing gestures and eye gaze. There is a widespread belief that nonhuman primates evince only a weak capacity to use deictic cues; in contrast, domestic dogs (*Canis familiaris*) tend to demonstrate high success rates. This pattern of canid superiority has been taken to support the Domestication Hypothesis, which posits enhancing effects of artificial selection on the socio-cognitive abilities of dogs and humans. Here we review nearly two decades of published findings, using variants of the OCT. We find systematic confounds with species classification in task-relevant preparation of the subjects, in the imposition of a barrier between reward and subject, and in the specific deictic cues used to indicate the location of hidden objects. Thus, the widespread belief that dogs outperform primates on OCTs is undermined by the systematic procedural differences in the assessments of these skills, differences that are confounded with taxonomic classification.

According to some theorists, the origins of human language may be found in gestural communication (Arbib, 2005; Corballis 2002) and there is profound interest in the effects of selection on the capacity to comprehend referential (deictic) gestures (Hare, Brown, Williamson, & Tomasello 2002). The Object Choice Task (OCT) is a widely used experimental paradigm that measures the ability of a subject to comprehend deictic (directional) cues, usually to find food (Anderson, Sallaberry, & Barbieri, 1995). There is a widespread belief that nonhuman primates have a deficient capacity to interpret deictic gestures, evidenced by their poor performances on this task (e.g., Maclean, 2016). In contrast, domestic dogs (*Canis lupus familiaris*) tend to demonstrate high success rates on the task, and the relative performances of both taxa have been taken as evidence for selective histories that facilitate social cognition in domesticated dogs and humans (i.e. the Domestication Hypothesis), while nonhuman primates are held to lack these selective histories (e.g., Hare et al., 2002). A competing theoretical perspective accounts for the performance differences on the OCT with reference to specific individual learning histories, irrespective of selective history (i.e. the Lived Experiences Model)—according to this line of reasoning, most captive great apes have impoverished social learning opportunities, relative to pet dogs, human children, enculturated primates, and their wild conspecifics (e.g., Bard & Leavens, 2014; Leavens & Bard, 2011; Racine, Leavens, Susswein & Wereha, 2008). Here we comprehensively review the OCT literature and show that experience with humans and procedural variables better explain group performance differences between dogs and nonhuman primates than do species classifications. When these mammals are matched on task-relevant pre-experimental history and on key procedural variables, species differences disappear. We find no evidence for a deficiency in social cognition, as measured by the OCT, in nonhuman primates, relative to dogs.

Gestures form an important part of human communication and the capacity to produce and comprehend them emerges in pre-verbal infants. In Western societies, the index-finger point is the predominant form of deictic gesture, used to direct another's attention to an object or event of interest (Butterworth, 2003; Masataka, 2003; but see, e.g., Cooperrider, Slotta, & Núñez, 2018, for descriptions of non-manual points). Human infants develop the ability to follow points at around 6 months of age (Butterworth, 2001) and begin to produce points at around 12- 15 months (Franco & Butterworth, 1996). The onset of pointing constitutes a significant developmental milestone (e.g., Flack & Leavens, 2018; Leavens & Clark, 2017), and its onset predicts the onset of speech (Colonnesi, Rieffe, Koops, & Perucchini, 2008; Iverson & Goldin-Meadow, 2005), and a delay in, or lack of the development of pointing has been linked to autism (Osterling & Dawson, 1994) and pervasive developmental disorders (Bernabei, Camaign, & Levi, 1998).

The OCT involves an experimenter baiting, typically, one of two or three opaque containers and then using a referential gesture, often a point or gaze cue, to indicate to the subject the container in which the bait has been placed. Human infants perform well on this task from around 12 months of age (Behne, Lizkowski, Carpenter & Tomasello, 2012) and a number of non-primate species such as horses and elephants have also been shown to be successful (Proops, Rayner, Taylor & McComb, 2013; Smet & Byrne, 2013).

Nonhuman primates, however, tend to have poor success rates on the OCT (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012). The results from studies with our nearest relatives, the great apes, have been used to propagate theories such as the Cultural Intelligence Hypothesis (Herrmann et al., 2007), the Shared Intentionality model

(Tomasello & Carpenter, 2007), and the Vygotskian Intelligence model (Moll & Tomasello, 2011), all of which share the premise that nonhuman primates are unable to comprehend the communicative significance of informative gestural cues and that, therefore, this is a human-unique ability that contributed to the emergence of verbal communication in humans. This pattern of poor comprehension of deictic cues by nonhuman primates sits uncomfortably against a large and growing literature demonstrating that great apes frequently use pointing, themselves, in captivity, typically with no explicit training to do so, along with many other animal taxa (see Krause, Udell, Leavens, & Skopos, 2018, for review).

However, there exist some disparities in the literature as to the relative abilities of different species, and, recently, some authors have begun to address these anomalous findings with reference to methodological and procedural factors in OCT experiments that are systematically confounded with species classification. In a review of ape OCT studies, Lyn (2010) found performance differences on the OCT as a function of rearing history in apes, such that enculturated apes, raised in an environment rich in human interaction, outperformed institutionalised apes and thus argued for greater consideration of rearing history when comparing across species. Mulcahy and Hedge (2012) reviewed 63 OCT papers and concluded that configurational differences in the testing of apes and dogs disadvantage the former due to decreased salience of, and attention, to the cue being given. Finally, in a review of gaze-following OCT studies with nonhuman primates Byrnie (2015) argues that there exists such disparity between different species in their performance on the OCT that taking one species' results as representative of their whole phylogenetic group leads to erroneous conclusions. Here, we develop and extend these findings in the most comprehensive OCT literature review to date, focusing primarily on domestic dogs and nonhuman primates, as it is evidence

from these species that has been used to support prevailing theories of human uniqueness in social cognition and of the effects of domestication on dogs' social-cognitive abilities.

The first factor that we address is that of the systematic confound between life history and species classification of subjects. Among humans, index-finger pointing is not a universally employed communicative gesture, and, in fact, in some non-Western societies, lip-pointing or nose-pointing is more predominantly used (Cooperrider et al., 2018; Enfield, 2001; Wilkins, 2003). The comprehension of pointing is a developmental process in human infants (Butterworth & Grover, 1988); it is through repeated exposure that pointing acquires its cultural and communicative significance. In fact, one of the first OCT studies conducted with nonhuman primates (Call & Tomasello, 1994) concluded that the marked difference in both pointing comprehension and production between an enculturated and an institutionalised orangutan was due to the subjects' differential experiences of human interaction. Call and Tomasello (1994) suggested that humanlike interaction in early ontogeny, therefore, was necessary for the development of an understanding of others as intentional agents and they remarked that this was possibly also the case for human infants. In spite of these speculations, the OCT literature in the ensuing 20 years comprises a multitude of studies of nonhuman primate in which the poor performance of samples of great apes raised in institutional settings with minimal exposure to humanlike interaction is attributed to a core, phylogenetic species difference in cognitive ability (e.g., Moore, Call, & Tomasello, 2015; Povinelli, Reaux, Bierschwale, Allain, & Simon, 1997; Tomasello, Call, & Gluckman, 1997; for critical analysis, see Leavens, Bard, & Hopkins, 2017; Lyn, 2010).

Bard and Leavens (2014) discussed the importance of social engagement in the development of socio-cognitive skills in human infants and there is an effect of the

amount and quality of human interaction on nonhuman primates' abilities to use human nonverbal cues in much the same way (Bard, Bakeman, Boysen & Leavens, 2014).

Studies which have compared nonhuman primates from different rearing environments on the OCT have found marked differences in their abilities as a function of their pre-experimental exposure to human interaction (Lyn, Russell & Hopkins, 2010). Hence, we argue that it is invalid to conclude that differences in experimental performance are due to species differences, without considering the experiential histories of all of the individuals tested.

This is further emphasised by the wealth of studies concluding that domestic dogs' superior performance on the OCT reflects specialised socio-cognitive skills evolved through their long history of domestication (Hare & Tomasello, 1999; Hare & Tomasello, 2005). Whilst pet dogs, who are extensively exposed to human interaction, are indeed adept at passing the OCT, the poor performance of dogs with alternative life histories, such as shelter dogs or kennel-bred research dogs (D'Aniello et al., 2017; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2010) demonstrates that this is not, in fact, an innate, evolutionarily derived adaptation. Lea and Osthaus (2018) suggest that in order to assess the extent to which dogs' cognitive skills are exceptional, it is necessary to consider their phylogenetic, ecological and anthropogenic backgrounds, comparing their abilities to representatives of taxonomic groups that share common features of these: other carnivores, other social hunters and other domestic animals, respectively. They argue that when dogs' social cognition is considered in this way, there is no evidence to show that they have unique abilities- other carnivores (e.g. sea lions, seals and dolphins) demonstrate high success rates and other domestic animals (e.g. pigs and goats) show similar abilities to follow points. With regard to social hunters, there is evidence that chimpanzees, too, can comprehend pointing cues

(Leavens & Clark, 2017). Thus, argue Lea and Osthaus (2018), when dogs' abilities are considered from these three perspectives, there is no evidence to suggest that their sociocognitive skills are exceptional.

The second factor, we propose, is the use of incommensurate testing protocols for representatives of different taxa. For example, in order to ensure the safety of the experimenter, the testing of nonhuman primates typically involves subjects participating from within a cage, thus introducing a barrier between the subject and the experimenter and testing apparatus (e.g. Bräuer et al., 2006; Herrmann et al., 2007; Hopkins, Russell, McIntyre & Leavens, 2013). Testing domestic dogs does not entail these same safety precautions, and so this barrier in the testing paradigm is generally absent. There are, however, two notable exceptions. The first (Udell, Dorey & Wynne, 2008) involved an OCT with domestic dogs in which a tapping cue was presented to subjects tested either with or without a fence separating subject and experimenter. The authors reported a significant difference between the performance levels of the two groups, with those in the barrier condition experiencing a 31% decrement in success levels. In addition, the only study to date which has involved the presentation of a pointing cue in a testing paradigm where a barrier was present for domestic dogs found significantly lower success rates in those dogs for which a barrier was present, compared to dogs for which this barrier was absent (Kirchhofer et al., 2012). We therefore argue that this difference in the testing paradigm represents a confound with taxonomic classification, which should not be ignored when comparing species' relative abilities.

A further systematic confound we consider here concerns the broad range of types of pointing cues presented in OCT experiments. Typically, researchers use an ipsilateral point where the ipsilateral arm is extended and the index finger outstretched (i.e., a point with the hand on the same side as the baited container is used). Miklósi

and Soproni (2006) differentiated between types of cues, specifically distal and proximal points, where the distances between the fingertip and the target are greater than 50cm (i.e. distal) and between 10 and 40cm (i.e. proximal), distinguished as so because the former is considered to be within reach and the latter not; and between momentary and dynamic pointing where the cue is presented for 1-2 seconds (i.e. momentary) or the cue is maintained until a choice has been made (i.e. dynamic). They reported that representatives of a number of species perform at different levels according to the type of cue presented and attribute this to the cues' differential effects on salience and memory. Udell, Hall, Morrison, Dorey and Wynne (2013) assessed dogs' performance on nine point types and found differences in levels of success as a function of the temporal and distal properties of the pointing cues.

Finally, Mulcahy and colleagues (Mulcahy & Call, 2009; Mulcahy & Hedge, 2012) argued that there is a tendency for nonhuman primates to be tested with a central version of the OCT, whereas domestic dogs are tested with a peripheral version. The distinction between the two concerns the inter-object distance between the containers, with the central version being categorised as one in which the containers are placed closer together (around 40cm. apart), whereas in the peripheral version the containers are further apart (around 2m. between containers). The authors argue that this could affect performance in several important ways. First, containers placed close together tend to be in the subject's direct line of vision, and, as such, the salience of the containers may distract subjects' attention from the cue being given. Second, retrieval of a reward from containers that are placed further apart may require increased effort, therefore increasing both attention to, and the salience of, the deictic cue. In support of this, Mulcahy and Call (2009) found that representatives of three species of great ape performed significantly better when tested with a peripheral version compared with

when tested with a distal version. In addition, the one study to date with has compared dogs' performance on the two versions of the task (Kraus, van Waveren & Huebner, 2014) found that performance was lower in the peripheral version.

In order to investigate the prevalence of these confounds, individual life history data, as classified by the original studies and performance data were collated from OCT studies published up until 2017. Given the enormous morphological variation that exists between different dog breeds (e.g., Shearin & Ostrander, 2010), we sampled the literatures for domestic dogs (an unusually and artificially diverse species) with representatives across the order Primates.

Method

Literature Selection

This study was conducted in two phases. In Phase 1 (2013-2016) we assembled every published article on animals' comprehension of deictic cues in OCT studies that we could find, across all vertebrate species, systematically noting the rearing environments where available. This phase of data collection supported the Human Experience Scale that is depicted in Figure 1 and listed in Table S2. The literature search consisted of the following components: Electronic databases (Scopus, ScienceDirect, PsychInfo and all Citation Databases included in ISI web of knowledge) from 1990 – 2015 with keywords in abstracts: “object choice task”, “object-choice”, object choice task (and) “animal”), citation search on author names, scanning reference lists, and Google scholar. We also used reference lists in the published articles to find additional reports not captured by our keyword searches. In addition, where we knew of relevant studies not captured by the above methods, we added those to the database. During a preliminary assessment of the database in 2016, it became apparent that

systematic comparisons between taxa would only be possible between canids and primates, because the database was dominated by these two taxonomic groups (a finding reported by Krause et al., 2018, Tables 1 & 2). Therefore, in Phase 2 (2016-2018), we focused on dogs and nonhuman primates, extending our cut-off date from the originally planned 2015 to 2017.

Studies were included if they involved an object choice task with at least one pointing cue condition. Initially, data were collected from 99 studies comprising 43 vertebrate species. As noted above, initially, data were collected from 99 studies comprising 43 vertebrate species. Individual rearing history, individual performance data, or both, were available for 3277 subjects. This review focuses on nonhuman primate vs. dog comparisons, for which data was available for 2534 individuals, including representatives of 16 nonhuman primate species, from 71 studies (see Table S1). Of the nonhuman primates, 82% were great apes, with 64% of the nonhuman primate subjects comprised of chimpanzees.

Subjects

Subjects' rearing histories, as classified by the original studies, were collated. Due to the variety of rearing histories of subjects, a human experience scale was created (see Table S2) which defines rearing histories in terms of the quantity and quality of experiential history with humans, such that “close” is characterised by having daily, intensive contact with humans, “occasional” as having some form of exposure, typically in the form of general husbandry, and “seldom” as having experience little or no exposure to humans. Each individual was given an ordinal grade on this scale according to the rearing history given in the original study. Figure 1 shows the human experience scale and the allocation of the original reported rearing histories to this scale.

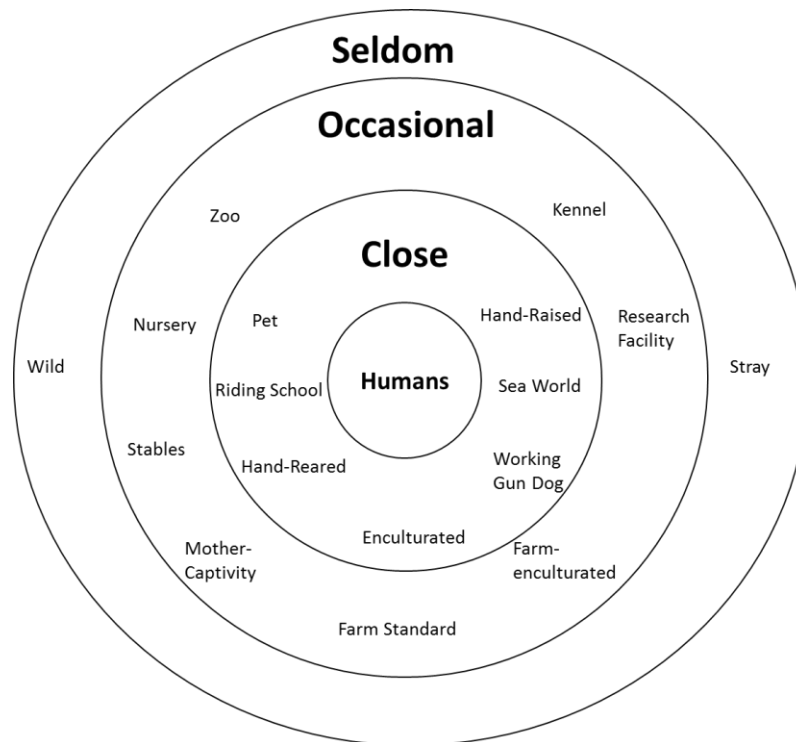


Figure 1. Rearing histories as reported in the original studies categorised according to the human experience scale. “Mother-captivity” means mother-reared in captivity.

Data were also collected regarding a number of features of the testing environment, such as the presence or absence of a barrier (most often in the form of a cage), the inter-object distances of containers and the numbers of containers used *inter alia*. Subjects who took part in multiple studies (usually nonhuman primates) were highlighted as having done so, as were those subjects who took part in multiple cue conditions.

Cue Types

For the purposes of this review, performance data were analysed only where one or more pointing cues were presented, and these cues were categorised according to Miklósi and Soproni’s (2006) definitions. The following distinctions were made:

Ipsilateral point vs. contralateral point. An ipsilateral point is where the pointing cue is presented using the hand that is on the ipsilateral side of the body in relation to the baited container. A contralateral point is one where the hand is on the opposite side of the body to the baited container is used.

Static point vs. dynamic point vs. momentary point. A point is categorised as static when the pointing hand is in place before the participant views the cue and remains so until the participant chooses one of the containers. A dynamic point is when the point is enacted once the participant is in position and is held until the participant makes a choice. A momentary point is where the point is enacted in front of the participant and is presented for 1-2 seconds before the hand returns to the resting position.

Proximal point vs. distal point. A point is said to be proximal when the distance between the fingertip and the baited container is less than 40cm. A distal point is categorised as such when the distance between the fingertip and the baited container is equal to or greater than 40cm.

Individual Performance Data

Although many studies present only group mean scores, individual performance data were obtained for 1137 individuals. Where possible, we recorded the number of trials in each pointing condition, the number of correct trials and the percentage of correct trials. Due to variation in the number of containers used in the studies, and thus, the differing chance levels of success, these scores were converted to Z-scores. A 'pass' or 'fail' was then obtained for each participant in each condition, with a 'pass' being a Z-score greater than or equal to 1.65 (one-tailed; see Rumbaugh, Washburn, & Pate, 1984, for justification).

Because lateral, temporal, and distance features of cues presented were not all systematically reported for many subjects, we analysed each of these three features separately, to maximise statistical power. Systematic confirmatory analyses were then conducted where two of these features were known, and finally, where all three properties were known, on ever-decreasing sample sizes.

For subjects participating in multiple conditions (i.e., different cue types), where these properties were shared across the cue types presented, an aggregate score was collated. For example, if a subject participated in an ipsilateral momentary distal point condition and an ipsilateral dynamic distal point condition, these scores were aggregated, such that a total number of trials and correct trials was obtained, when analyses were conducted regarding ipsilateral or distal pointing cues, but not when examining momentary or dynamic points. Performance data for subjects who participated in multiple conditions (i.e., multiple cue types) were excluded from the performance analyses, as to include their data would be to violate the assumption of independence. These data were then analysed separately, using statistical tests that allowed for within-subjects analyses. This was the case for all analyses excepting cue type distribution analyses, as the aim of this analysis was to examine the frequency of the exposure to the different types of cue across taxonomic groups.

Where participants had participated in multiple studies, their results were taken as independent data points, because studies which had taken place in different years of their lives can be viewed as independent events.

Results

Human experience

Rearing history data were available for 2534 subjects, comprised of 2064 dogs

and 470 nonhuman primates. There was a significant difference in the level of human experience between dogs and nonhuman primates (Kruskal-Wallis, $\chi^2(3, N = 2534) = 1550, p < .001$) with 91% of dogs being categorised as “close”, compared with 6% of nonhuman primates; within nonhuman primates, 87% were categorised as “occasional” and 2% as “seldom” (Figure 2). This highlights the lack of comparability between dogs and nonhuman primates with regard to human experience, with the majority of dogs having a much more enriched experiential history with humans than their nonhuman primate counterparts.

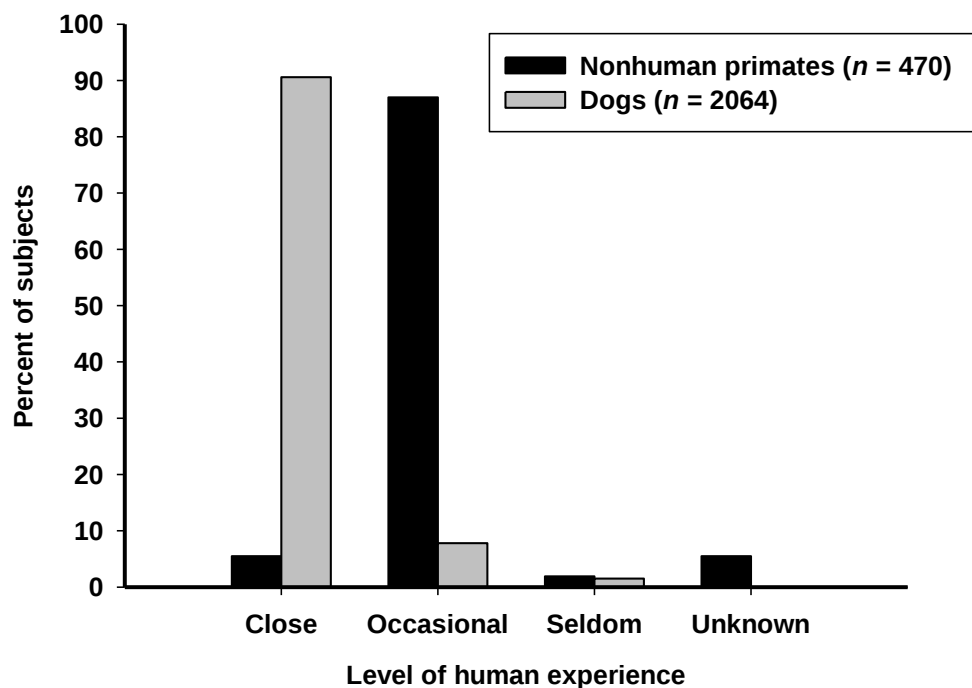


Figure 2: The distribution of levels of human experience between nonhuman primates and dogs.

Human experience and performance

Importantly, a relationship was also found between human experience and performance on the OCT on a number of pointing cues.

For *ipsilateral* pointing cues, within nonhuman primates and dogs, there was a

significant difference in performance between subjects in the different levels of the human experience scale, Kruskal-Wallis $\chi^2(2, N = 212) = 16.43, p = .001, r = .28$. Pairwise comparisons showed that subjects categorised as “close” ($N = 174, Mdn z = 1.26$) scored higher than those categorised as “occasional” ($N = 22, Mdn z = 0.00$), Mann-Whitney $U = 1209.5, p = .005$, and those categorised as “seldom” ($N = 16, Mdn z = -0.32$), Mann-Whitney $U = 731.00, p = .002$. There was no significant difference between those categorised as “occasional” ($N = 22, Mdn z = 0.00$) and “seldom” ($N = 16, Mdn z = -0.32$), Mann-Whitney $U = 135.35, p = .227$, mean $z = 0.36, SD = 0.98$. Within-nonhuman primates, sample sizes were too small (22 “occasional” subjects, 2 “seldom” subjects) for sufficient statistical power for comparisons. Within dogs, those categorised as “close” ($N = 174, Mdn z = 1.26$) scored higher than those categorised as “seldom” ($N = 14, Mdn z = -0.63$) (Mann-Whitney $U = 13.97, p < .001$).

For *contralateral* pointing cues, in contrast, within nonhuman primates and dogs, those categorised as “occasional” ($N = 95, Mdn z = 1.89$) outperformed those categorised as “close” ($N = 6, Mdn z = 0.00$), (Mann-Whitney $U = 136.5, p = .029$). Those categorised as “occasional” comprised solely nonhuman chimpanzee subjects and so further analyses were not possible. This suggests that for *contralateral* cues, there may be inherent species differences in responsiveness to cue features that explain the performance differences, or, alternatively, a more complex relationship with human experience may exist that leads to the suppression of comprehension of contralateral cues as a result of increased exposure to humans.

Where *momentary* pointing cues were presented, within nonhuman primates and dogs, subjects categorised as “close” ($N = 356, Mdn z = 0.89$) scored higher than those categorised as “seldom” ($N = 22, Mdn z = -0.63$), Mann-Whitney $U = 1235.5, p < .001$. This was a within-dog difference, because no performance data were available for

nonhuman primates on this cue, and thus replicates previous findings (D’Aniello et al., 2017; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2010) that dogs that have experienced greater exposure to humans are more capable in understanding this more difficult pointing gesture.

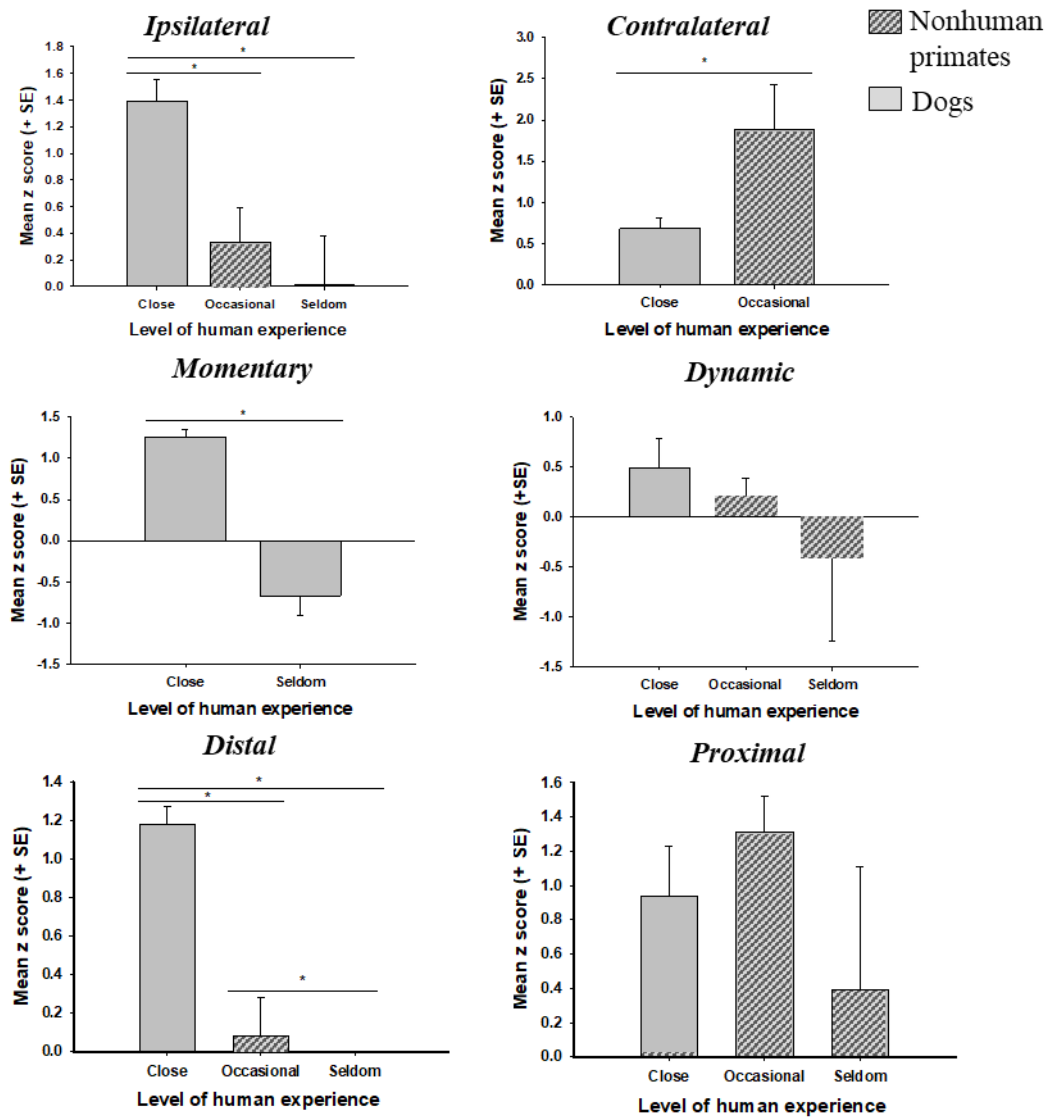


Figure 3. The mean standardised z scores (and standard errors) of nonhuman primates and dogs and the proportion of each specie/ taxonomic group contributing to those means, categorised according to level of human experience on six different pointing cues. * denotes $p < .05$.

Where *dynamic* cues were presented, there were no differences in performance

between the categorisations of level of human experience within nonhuman primates and dogs, Kruskal-Wallis $\chi^2(2, N = 82) = 1.84, p = .398$, nor within nonhuman primates alone, Kruskal-Wallis $\chi^2(2, N = 36) = 2.81, p = .246$. Dogs for which there were performance data available were all categorised as “close”, so within-species analyses were not possible. This shows that level of human experience may be of less importance in the comprehension of this easier pointing cue than for those more-difficult-to-follow cues.

Where *distal* cues were presented, within nonhuman primates and dogs, there was a significant effect of level of human experience on performance, (Kruskal-Wallis $\chi^2(2, N = 395) = 35.27, p < .001$. Pairwise comparisons showed that those categorised as “close” ($N = 353, \text{Mdn } z = 0.89$) scored higher than those categorised as “occasional” ($N = 20, \text{Mdn } z = 0.00$), Mann-Whitney $U = 2066.00, p = .002$, and those categorised as “seldom” ($N = 22, \text{Mdn } z = -0.63$), Mann-Whitney $U = 1334.50, p < .001$ (this was a within-dog comparison). “Occasional” ($N = 20, \text{Mdn } z = 0.00$) subjects also scored higher than “seldom” subjects ($N = 22, \text{Mdn } z = -0.63$), Mann-Whitney $U = 126.00, p = .017$. Within-nonhuman primate comparisons were not possible because all subjects for which there were data were categorised as “occasional”. This shows that level of human experience may better explain performance differences than taxonomic group affiliation.

Where *proximal* cues were presented, there was no significant effect of level of human experience on performance, (Kruskal-Wallis $\chi^2(2, N = 94) = 2.12, p < .346$. Within nonhuman primates only, there was no significant effect of human experience on performance, (Kruskal-Wallis $\chi^2(2, N = 51) = 2.31, p = .315$). All 43 of the dog subjects were categorised as “close” so within-dog analyses were not possible. This shows that, for *proximal* cues, intense exposure to humans may not have as important a role in

facilitating comprehension as for more difficult distal cues. Figure 3 shows the comparisons in performance between subjects with different levels of human experience across the different point types.

Presence of a barrier

Comparisons of the presence of a barrier in the testing environment between nonhuman primates and dogs showed that less than 1% of dogs were tested with a barrier, compared with 99% of nonhuman primates, $\chi^2(1, N=2534) = 2411.77, p < .001$, as shown in Figure 4. This highlights the systematic inconsistencies present in testing environments across the two taxonomic groups, and, therefore, the reduced validity of interpreting group differences as phylogenetic traces of differences in selective histories. Comparisons of performance between dogs tested with a barrier and those tested without were not possible because among the only two studies to introduce barriers to the testing protocol, Kirchhofer et al. (2012) did not specify which individuals participated in the barrier condition and Udell et al.'s (2008) dogs tested with a barrier did not take part in a pointing cue condition. Comparisons between nonhuman primates tested with and without a barrier were not possible because only 3 infant chimpanzees were tested without a barrier present (Okamoto-Barth, Tomonaga, Tanaka & Matsuzawa, 2008), thus sufficient statistical power was lacking.

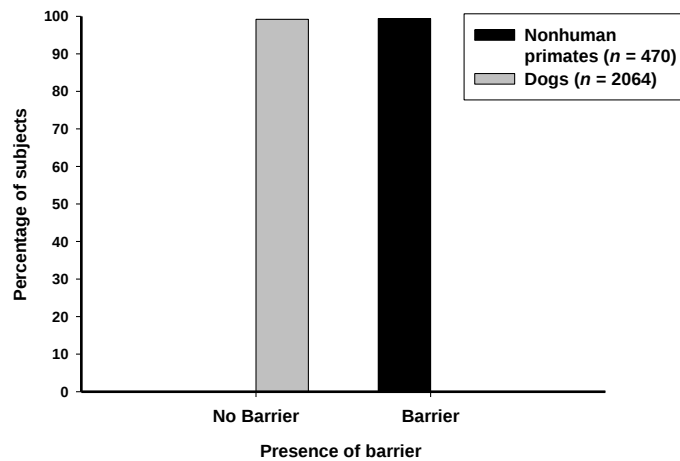


Figure 4. The percentage of nonhuman primates and dogs tested with and without a barrier.

Cue types

Among those subjects tested with only one type of pointing cue, differences were also found regarding the types of cues presented to the two taxonomic groups. There was a significant difference between lateral properties (i.e. whether ipsilateral or contralateral hand was used to point) of cues presented to nonhuman primates and dogs, with 26% of points to nonhuman primates being *ipsilateral* and 74% *contralateral*, compared with 82% *ipsilateral* and 18% *contralateral* for dogs, $\chi^2(1, N = 1777) = 328.59, p < .001$, as shown in Figure 5a. For nonhuman primates, there was no significant difference in performance between *ipsilateral* ($N = 24, \text{Mdn } z = 0.25$) and *contralateral* ($N = 6, \text{Mdn } z = 1.89$), Mann-Whitney $U = 34.5, p = .05$, shown in Figure 6a. Dogs scored significantly higher on *ipsilateral* ($N = 188, \text{Mdn } z = 1.00$) than on *contralateral* ($N = 95, \text{Mdn } z = 0.00$) pointing cues, shown in Figure 6b. This shows that lateral cue features can differentially affect different species' performance.

Figure 5b shows the percentage of the different temporal cue types presented to nonhuman primates and dogs. There was a significant difference across taxa in temporal

cue properties, with 4% of cues presented to nonhuman primates being static, 7% momentary and 90% dynamic, compared with 1% static for dogs, 45% momentary and 42% dynamic, Kruskal-Wallis $\chi^2 (2, N = 2105) = 195.48, p < .001, r = .30$. This demonstrates that there are also systematic differences in the temporal properties of cue types presented to nonhuman primates and dogs, with a bias towards dynamic pointing for nonhuman primates. There were insufficient data to analyse performance differences between the three point types for nonhuman primates (see Figure 6a), and there was no significant difference in performance on *dynamic* ($N = 46, Mdn z = 0.57$) and *momentary* ($N = 378, Mdn z = 0.89$) pointing cues within dogs, Mann-Whitney $U = 7376.0, p = .092$ (Figure 6b).

There was a significant difference in the distance properties, with 16% of cues presented to nonhuman primates being distal cues, and 84% proximal, compared with 74% distal for dogs and 26% proximal, $\chi^2 (1, N = 1621) = 387.86, p < .001$ (Figure 5c). This shows that there are marked differences in the distance properties of cue types presented between nonhuman primates and dogs, with a bias towards distal pointing cues for dogs and towards proximal pointing cues for nonhuman primates. Within nonhuman primates, subjects scored higher on *proximal* ($N = 54, Mdn z = 0.90$) than on *distal* ($N = 20, Mdn z = 0.00$) pointing cues, Mann-Whitney $U = 309.50, p = .005$ (Figure 6a). Within dogs, there was no significant difference in performance between *distal* ($N = 375, Mdn z = 1.07$) and *proximal* ($N = 43, Mdn z = 0.63$) pointing cues, Mann-Whitney $U = 7441.00, p = .406$ (Figure 6b). This demonstrates that, for nonhuman primates, there are performance differences associated with the distance properties of the cue being presented.

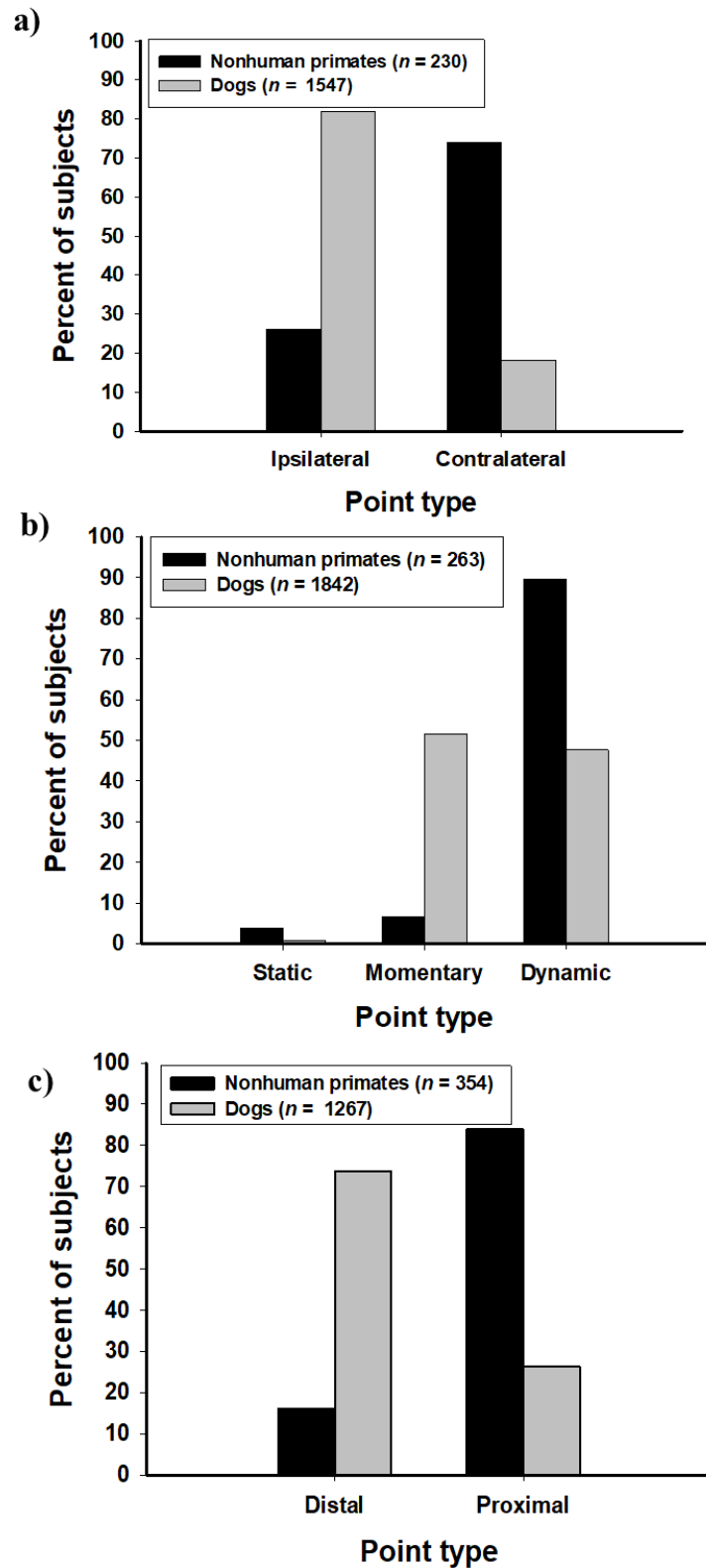


Figure 5. The percentage of nonhuman primate and dog subjects presented with each point type according to a) lateral features b) temporal features and c) distance features of the cue.

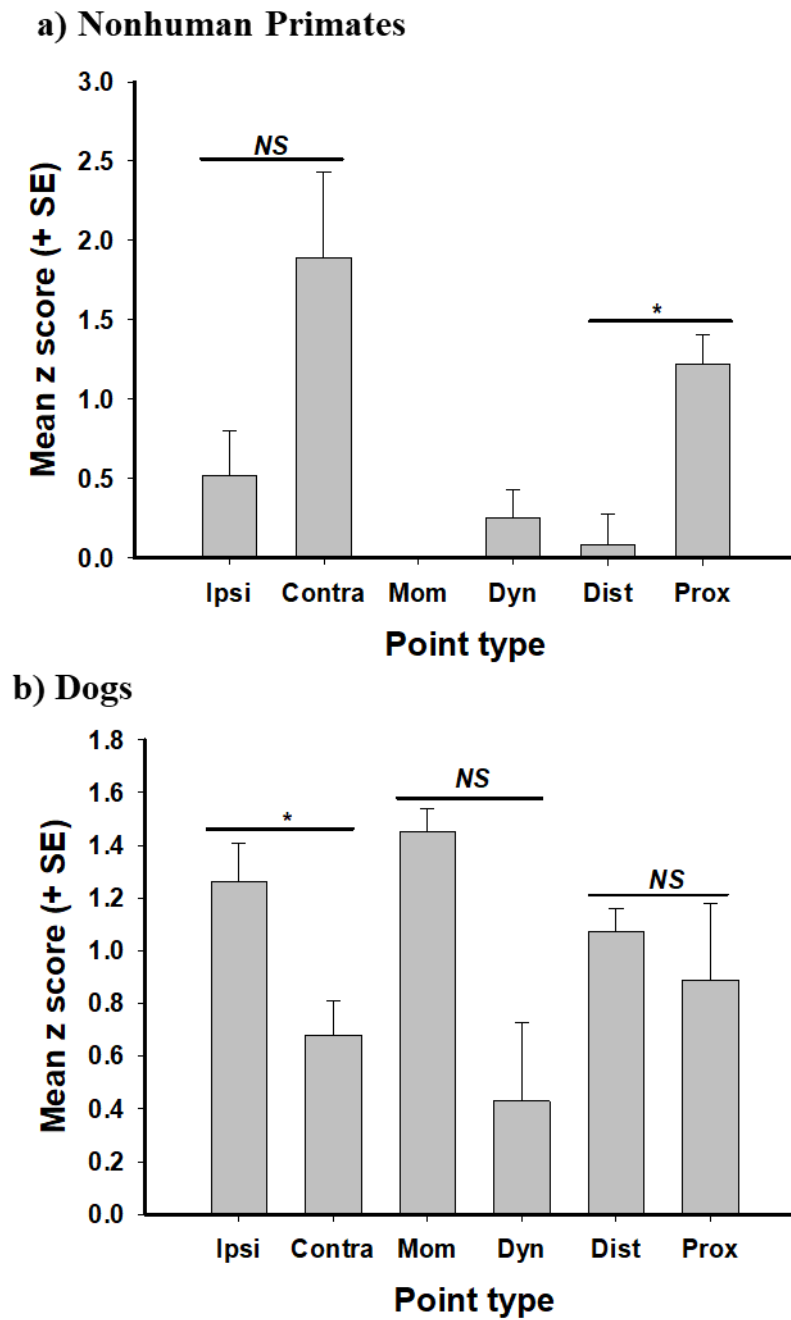


Figure 6. The mean z scores and standard errors for a) nonhuman primates and b) dogs on the different pointing cue types. Ipsi = *ipsilateral*; Contra = *contralateral*. * denotes significant at $p < .05$. NS = not significant.

Multiple conditions

Four hundred and two subjects took part in studies in which they were presented

with multiple cue types. Individual performance data were available for 210 of these subjects. There were insufficient data to conduct statistically robust comparisons of performance according to level of human experience. Analyses were possible, however, for comparisons of performance according to cue type. Nonhuman primates scored higher when tested with *distal* cues (mean $z = 2.47$, $SD = 1.62$) than with *proximal* cues (mean $z = 0.75$, $SD = 1.62$), ($Z = -3.01$, $p = .003$). Dogs, in contrast, performed better when tested with *proximal* (mean $z = 3.90$, $SD = 0.62$) rather than *distal* cues (mean $z = 0.51$, $SD = 2.59$), ($Z = -2.37$, $p = .018$). With regard to temporal properties of cues, nonhuman primates scored higher when tested with *dynamic* (mean $z = 1.07$, $SD = 0.93$) rather than *momentary* (mean $z = 0.05$, $SD = 1.09$) cues, ($Z = -2.58$, $p = .010$). This was also the case for dogs (*dynamic* mean $z = 2.33$, $SD = 1.50$; *momentary* mean $z = 0.80$, $SD = 1.75$), ($Z = -2.94$, $p = .003$). This shows that both temporal and distance properties of pointing cues may affect individual performances, and that there may be different processes at play in terms of their effects depending on taxonomic group. There were insufficient data to analyse performance for *static*, *ipsilateral* and *contralateral* pointing cues.

Inter-object distance

Further procedural differences were found with respect to the inter-object distance between containers. Dogs ($Mdn = 155.0$ cm) were tested with significantly greater inter-object distances than nonhuman primates ($Mdn = 58.0$ cm), (Mann-Whitney $U = 4917.5$, $z = -27.99$, $p < .001$), demonstrating a bias towards greater distances between the containers for dogs than for nonhuman primates, congruent with Mulcahy and Hedge's (2012) findings. Analyses of the relationship between inter-object distance and performance found significant correlations between these two variables, with greater inter-object distances found to be associated with increased performance on

ipsilateral ($r_s(179) = .20, p = .007$), *contralateral* ($r_s(25) = .41, p = .040$), *distal* ($r_s(319) = .35, p < .001$), and *momentary* ($r_s(269) = .23, p < .001$), pointing cues. There were no significant correlations with performance for *proximal* ($r_s(51) = -.142, p = .340$), or *dynamic* ($r_s(73) = -.010, p = .930$), points. This shows that, for a number of pointing cues, overall, greater inter-object distances seem to foster increased performances and therefore current testing protocols may be disadvantaging nonhuman primates.

Discussion

Our results highlight the procedural and methodological factors that can influence a subject's performance on the OCT and demonstrate that the trend in the existing literature to compare across these two taxonomic groups without considering these factors greatly reduces the legitimacy of findings. First, it is clear from the results that experiential history with humans can influence an individual's ability in the comprehension of pointing cues, supporting the results of Udell and colleagues (Udell, Dorey & Wynne, 2008; Udell et al., 2012), Lyn and her colleagues (Lyn, 2010; Lyn, Russell & Hopkins, 2010), and numerous others (e.g., Bard, Bakeman, Boysen, & Leavens, 2014; Hopkins, Russell, McIntyre, & Leavens, 2013; Pedersen, Segerdahl, & Fields, 2009; Scheel, Shaw, & Gardner, 2016) who suggested that exposure to humans and immersion in their environment is a key factor in the development of the comprehension of human communicative cues. As shown in Figure 2, the nonhuman primate subjects in the existing literature come from a wide variety of rearing backgrounds, with only 6% of subjects being enculturated and the majority being nursery-raised or mother-reared in captivity. When this is compared with the dogs' rearing history data, it is clear that the vast majority of dogs are sampled from a pet background. Those who had more impoverished backgrounds in terms of human

experience- for example, stray, shelter and free ranging dogs--demonstrated lower success rates on the OCT (e.g., Udell, Dorey & Wynne, 2008; Udell, Dorey & Wynne, 2010). Thus, to make generalisations about the relative abilities of species without considering their experiential backgrounds (e.g. Herrmann et al., 2007; Kirchhofer et al., 2012) and, furthermore, to base theories of species' evolutionary histories on results from such studies (e.g. Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare & Tomasello, 2005) is not warranted, due to the pervading imbalance in task-relevant pre-experimental experience between dogs and nonhuman primates. Contemporary claims to the effect that dogs have greater social awareness than nonhuman primates are, thus, not supported by compelling experimental evidence.

In addition, there are large differences between the two taxonomic groups in the procedural aspect of whether or not there is an intervening barrier between subject and cue provider in the testing paradigm. For example, excepting three infant subjects, all nonhuman primate subjects experienced testing with a barrier between subject and apparatus compared with less than 1% of dogs. In fact, this 1% consists, entirely, of a sample of 16 dogs in Kirchhofer et al.'s (2012) study, in which they compared performance of dogs tested with and without a barrier and found that those tested with a barrier performed significantly worse than the dogs tested without a barrier. Whilst it is accepted that a barrier is a necessary precaution when working with dangerous animals, the results of this study highlight the impact that this can have on success rate (Kirchhofer et al., 2012). This absence of consistency in testing conditions represents a systemic confound with taxonomic classification in the contemporary scientific literature; dogs and nonhuman primates have not been compared on the same OCT task.

Moreover, there are substantial and systematic differences in the cue types presented to different taxonomic groups, which, again, demonstrates that the

comparisons that are currently being drawn in the literature regarding the relative abilities of different species are not based on like-for-like testing paradigms. Considering that Miklósi and Soproni (2006) and Udell et al. (2013) highlighted the differential abilities involved in the comprehension of the various cue types, with regard to the salience and memory functions necessary, this shows the importance of testing representatives of any given species not just on one pointing cue type, but on several, before drawing conclusions about a species' ability to comprehend human gestural cues, and of making comparisons about ability only when the cue type is matched between samples. This is further demonstrated here, in the findings that there are advantages across different taxonomic groups of particular cue types, specifically those that involve *dynamic* pointing features.

Finally, the differences found in the distances between the containers in the testing paradigm support Mulcahy and colleagues' (Mulcahy & Call, 2009; Mulcahy & Hedge, 2012) assertions that nonhuman primates tend to be tested with a central version of the task, whereas dogs are tested with a more peripheral version. The authors suggest that placing containers close together and within the direct line of vision of the subject can lead to the salience of the containers distracting the subjects' attention from the cue being presented. Alternatively, it may be that placing the containers further apart signifies an additional cost to make a choice, and, as such, there is an increase in the attention afforded the cues by the subject. The positive correlations found in the current review, between inter-object distance and performance on several cue types, provide support for these hypotheses and further evidence that inter-species comparisons without regard for procedural factors such as these is neither appropriate nor scientifically sound.

In conclusion, the current review builds on existing criticisms of the current state

of the OCT literature to further demonstrate that methodological and procedural confounds limit the validity of the results of many studies. In order to gain a more comprehensive understanding of the abilities of different taxonomic groups to understand human communicative cues, it is necessary for further research that controls for the abovementioned factors to be conducted. We recommend that a series of systematic experiments in which these variables are manipulated is required. This should begin by manipulating such variables with human infant participants, such that the effect of, for example, barriers or pointing cues can be established in this ‘enculturated’ sample, and then to broaden the samples of species, maintaining consistency throughout and ensuring that comparisons are only made across truly comparable groups. In addition, our analysis demonstrates that much greater consideration needs to be given to ontogenetic influences on behaviour, rather than the pervasive reliance on phylogenetic explanations that prevails in the literature (Bard & Leavens, 2014; Leavens et al., 2017). Consistent with Udell and colleagues (Udell, Dorey & Wynne, 2008a, 2008b, 2010) we argue that, prior to asserting reductionist interpretations that assume that individuals’ behaviour is solely a function of their evolutionary history, their individual learning experiences must be taken into account.

Acknowledgements

The authors would like to thank the following people for providing additional raw data: Federica Amici, Tanya Behne, Juliane Bräuer, Judith Burkart, William D. Hopkins, Krisztina Kupán, Heidi Lyn, Heidi Marsh, Sadahiko Nakajima, Christian Nawroth, Péter Pongrácz, Christian Schloegl and Monique Udell. We also thank the staff at the University of Sussex Library for exemplary support during all phases of this project.

Supplementary Materials

Table S1. Authors and dates of publication; subject species, with sample size shown in parentheses; and cuing conditions of studies for which data were obtained at the individual level.

Study	Species and Number of Subjects	Pointing Cues
Anderson, Montant & Schmitt (1996) ^b	Rhesus Macaques, <i>Macaca mulatta</i> (3)	<i>Proximal</i>
Anderson, Sallaberry & Barbier (1995) ^b	Capuchins, <i>Cebus apella</i> (3)	<i>Proximal</i>
Barth, Reaux & Povinelli (2005) ^a	Chimpanzees, <i>Pan troglodytes</i> (5)	<i>Dynamic Proximal</i>
Bhattacharjee et al. (2017) ^a	Domestic Dogs, <i>Canis familiaris</i> (209)	<i>Ipsilateral Dynamic Proximal</i>
Bräuer, Kaminski, Riedel, Call & Tomasello (2006) ^b	Domestic Dogs, <i>C. familiaris</i> (21)	<i>Ipsilateral Dynamic</i>
	Bonobos, <i>Pan paniscus</i> (4)	<i>Ipsilateral Momentary</i>
	Chimpanzees, <i>P. troglodytes</i> (12)	
Burkart & Heschl (2006) ^b	Common Marmosets, <i>Callithrix jacchus</i> (10)	<i>Ipsilateral Static Proximal</i>
		<i>Ipsilateral Static Distal</i>
Byrnit (2004) ^a	Orangutans (4), <i>Pongo pygmaeus</i>	<i>Proximal</i>
Byrnit (2009) ^a	Gorillas, <i>Gorilla gorilla</i> (3)	<i>Dynamic Proximal</i>

Call & Tomasello (1994) ^b	Orangutans, <i>P. pygmaeus</i> (2)	No properties known.
Call, Hare & Tomasello (1998) ^b	Chimpanzees, <i>P. troglodytes</i> (6)	No properties known.
Call, Agnetta & Tomasello (2000) ^b	Chimpanzees, <i>P. troglodytes</i> (15)	No pointing cues given.
Carballo, Freidin, Casanave & Bentosela (2016) ^a	Domestic Dogs, <i>C. familiaris</i> (12)	<i>Proximal Dynamic,</i> <i>Distal Dynamic</i>
Dalla Costa, Cannas, Minero & Palestini (2010) ^a	Domestic Dogs, <i>C. familiaris</i> (37)	No properties known
D’Aniello et al. (2017) ^b	Domestic Dogs, <i>C. familiaris</i> (29)	<i>Distal Dynamic,</i> <i>Proximal Dynamic</i>
Dorey, Udell & Wynne (2010) ^b	Domestic Dogs, <i>C. familiaris</i> (33)	<i>Ipsilateral Momentary Proximal</i>
Essler, Schwartz, Rossettie & Judge (2017) ^a	Capuchins, <i>C. apella</i> (10)	<i>Ipsilateral Dynamic Proximal</i>
Gácsi et al. (2009) ^a	Domestic Dogs, <i>C. familiaris</i> (23)	<i>Momentary Distal</i> <i>Momentary Proximal</i>
Gácsi et al. (2009) ^a	Domestic Dogs, <i>C. familiaris</i> (180)	<i>Momentary Distal</i>
Gácsi, McGreevy, Kara & Miklósi (2009) ^b	Domestic Dogs, <i>C. familiaris</i> (140)	<i>Momentary Distal</i>

Hare & Tomasello (1999) ^b	Domestic Dogs, <i>C. familiaris</i> (10)	<i>Contralateral</i>
Hare, Brown, Williamson & Tomasello (2002) ^a	Domestic Dogs, <i>C. familiaris</i> (12)	<i>Contralateral Proximal</i>
Hare & Tomasello (2004) ^b	Chimpanzees, <i>P. troglodytes</i> (12)	<i>Contralateral Proximal</i>
Hare et al. (2005) ^a	Domestic Dogs, <i>C. familiaris</i> (11)	<i>Dynamic</i>
Hattori, Kurashima & Fujita (2007) ^a	Capuchins, <i>C. abella</i> (5)	No pointing cues given.
Hegedüs Bálint, Miklósi & Pongrácz (2013) ^a	Domestic Dogs, <i>C. familiaris</i> (61)	<i>Momentary Distal</i>
Hernádi, Kis, Turcsán & Topál (2012) ^a	Domestic Dogs, <i>C. familiaris</i> (18)	<i>Momentary Proximal</i>
Herrmann, Melis & Tomasello (2005) ^b	Chimpanzees, <i>P. troglodytes</i> (12)	No pointing cues given.
	Orangutans, <i>P. pygmaeus</i> (6)	
	Gorillas, <i>G. gorilla</i> (6)	
	Bonobos, <i>Pan paniscus</i> (4)	
Herrmann et al. (2007) ^a	Chimpanzees, <i>P. troglodytes</i> (106)	<i>Contralateral Dynamic Proximal</i>
	Orangutans, <i>P. pygmaeus</i> (32)	

Hopkins, Russell, McIntyre & Leavens (2013) ^b	Chimpanzees, <i>P. troglodytes</i> (35)	<i>Proximal</i>
Inoue, Inoue & Itakura (2004) ^b	White-Handed Gibbon, <i>Hylobates lar</i> (1)	<i>Proximal</i>
Itakura & Tanaka (1998) ^b	Chimpanzees, <i>P. troglodytes</i> (2)	<i>Proximal</i>
	Orangutan, <i>P. pygmaeus</i> (1)	
Itakura, Agnetta, Hare & Tomasello (1999) ^b	Chimpanzees, <i>P. troglodytes</i> (13)	No properties known.
Kaminski, Schulz & Tomasello (2011) ^a	Domestic Dogs, <i>C. familiaris</i> (50)	<i>Contralateral Momentary Distal</i>
Kirchhofer, Zimmermann, Kaminski & Tomasello (2012) ^b	Domestic Dogs, <i>C. familiaris</i> (32)	<i>Ipsilateral Dynamic Distal</i>
	Chimpanzees, <i>P. troglodytes</i> (20)	
Kraus, van Waveren & Huebner (2014) ^a	Domestic Dogs, <i>C. familiaris</i> (40)	<i>Ipsilateral Momentary Proximal</i>
Lakatos, Dóka, Miklósi (2007) ^a	Domestic Dogs, <i>C. familiaris</i> (14)	<i>Ipsilateral Momentary Distal</i>
		<i>Contralateral Momentary Distal</i>
Lakatos, Soproni, Dóka & Miklósi (2009) ^a	Domestic Dogs, <i>C. familiaris</i> (15)	<i>Ipsilateral Momentary Distal</i>
		<i>Contralateral Momentary Distal</i>

Lyn, Russell & Hopkins (2010) ^b	Chimpanzees, <i>P. troglodytes</i> (10)	<i>Proximal</i>
	Bonobos, <i>P. Paniscus</i> (7)	
Maclean, Krupenye & Hare (2014) ^a	Domestic Dogs, <i>C. familiaris</i> (40)	<i>Ipsilateral Dynamic</i>
Maclean, Herrmann, Suchindran & Hare (2017) ^a	Domestic Dogs, <i>C. familiaris</i> (552)	<i>Ipsilateral Dynamic</i>
Marsh (2012) ^b	Orangutans, <i>P. pygmaeus</i> (5)	No pointing cues given.
McKinley & Sambrook (2012) ^b	Domestic Dogs, <i>C. familiaris</i> (16)	<i>Dynamic</i>
Miklósi et al. (2005) ^a	Domestic Dogs, <i>C. familiaris</i> (21)	<i>Dynamic Proximal</i>
		<i>Dynamic Distal</i>
		<i>Momentary Proximal</i>
		<i>Momentary Distal</i>
Mulcahy & Call (2009) ^b	Chimpanzees, <i>P. troglodytes</i> (12)	<i>Contralateral Dynamic Distal</i>
	Bonobos, <i>P. paniscus</i> (4)	<i>Contralateral Dynamic Proximal</i>
	Orangutans, <i>P. pygmaeus</i> (3)	
Mulcahy & Suddendorf (2011) ^a	Orangutan, <i>P. pygmaeus</i> (1)	<i>Dynamic Proximal</i>
		<i>Dynamic Distal</i>

Nakajima, Fukuoka, Takamatsu & Chin (2009) ^b	Domestic Dogs, <i>C. familiaris</i> (9)	<i>Contralateral Dynamic Distal</i>
Neiworth, Burman, Basile & Lickteig (2002) ^a	Cotton-Top Tamarins, <i>Saguinis oedipus</i> (6)	<i>Proximal</i>
Okamoto-Barth, Tomonaga, Tanaka & Matsuzawa (2008) ^a	Chimpanzees, <i>P. troglodytes</i> (3)	<i>Proximal</i>
Okamoto et al. (2002) ^a	Chimpanzee, <i>P. troglodytes</i> (1)	<i>Proximal</i>
Peignot & Anderson (1999) ^a	Gorillas, <i>G. gorilla</i> (5)	<i>Proximal</i>
Pettersson, Kaminski, Herrmann & Tomasello (2011) ^b	Domestic Dogs, <i>C. familiaris</i> (76)	<i>Contralateral Momentary Distal</i>
Plaude & Fiset (2013) ^b	Domestic Dogs, <i>C. familiaris</i> (10)	<i>Ipsilateral Momentary Proximal</i>
Pongrácz, Gácsi, Hegedüs, Péter & Miklósi (2013) ^b	Domestic Dogs, <i>C. familiaris</i> (115)	<i>Ipsilateral Momentary Distal</i>
		<i>Contralateral Momentary Distal</i>
Povinelli, Nelson & Boysen (1990) ^a	Chimpanzees, <i>P. troglodytes</i> (4)	No properties known.
Povinelli, Parks & Novak (1991) ^a	Rhesus Macaques, <i>M. mulatta</i> (4)	No properties known.
Povinelli, Reaux, Bierschwale, Allain & Simon	Chimpanzees, <i>P. troglodytes</i> (7)	<i>Distal</i>

(1997)^bPovinelli, Bierschwale & Čech (1999)^aChimpanzees, *P. troglodytes* (7)

No properties known.

Riedel, Schumann, Kaminski, Call & Tomasello

Domestic Dogs, *C. familiaris* (64)*Contralateral Dynamic Proximal*(2007)^bSchmidjell, Range, Huber & Virányi (2004)^aDomestic Dogs, *C. familiaris* (102)*Ipsilateral Momentary Distal*Schmitt, Schloegl & Fischer (2014)^bLong-Tailed Macaques, *Macaca fascicularis* (10)*Contralateral Dynamic Proximal*Takaoka, Maeda, Hori & Fujita (2015)^aDomestic Dogs, *C. familiaris* (65)*Momentary Proximal*Tan, Tao & Su (2014)^bGolden Snub-Nosed Monkeys, *Rhinopithecus roxellana**Ipsilateral Dynamic*

(4)

Tomasello, Call & Gluckman (1997)^bChimpanzees, *P. troglodytes* (6)*Dynamic Proximal*Orangutans, *P. pygmaeus* (3)Udell, Dorey & Wynne (2008)^bDomestic Dogs, *C. familiaris* (46)*Momentary Distal*Udell, Dorey & Wynne (2010)^bDomestic Dogs, *C. familiaris* (23)*Ipsilateral Momentary Distal*

		<i>Ipsilateral Dynamic Proximal</i>
Udell, Ewald, Dorey & Wynne (2014) ^a	Domestic Dogs, <i>C. familiaris</i> (36)	<i>Ipsilateral Momentary Distal</i>
Udell et al. (2013) ^a	Domestic Dogs, <i>C. familiaris</i> (58)	<i>Static Proximal</i>
		<i>Dynamic Proximal</i>
		<i>Momentary Proximal</i>
		<i>Static Distal</i>
		<i>Dynamic Distal</i>
		<i>Momentary Distal</i>
Udell, Giglio & Wynne (2008) ^b	Domestic Dogs, <i>C. familiaris</i> (6)	<i>Momentary Distal</i>
Udell, Spencer, Dorey & Wynne (2012) ^b	Domestic Dogs, <i>C. familiaris</i> (7)	<i>Dynamic Proximal</i>
		<i>Contralateral Dynamic Distal</i>
Vick & Anderson (2000) ^a	Capuchins, <i>C. apella</i> (3)	<i>Proximal</i>
Vick & Anderson (2003) ^a	Olive Baboons, <i>Papio anubis</i> (4)	No pointing cue given.
Wobber et al. (2009) ^a	Domestic Dogs, <i>C. familiaris</i> (59)	<i>Contralateral Dynamic Distal</i>

Zaine, Domeniconi & Wynne (2015) ^a	Domestic Dogs, <i>C. familiaris</i> (60)	<i>Ipsilateral Momentary Distal</i> <i>Ipsilateral Momentary Proximal</i> <i>Ipsilateral Dynamic Proximal</i>
Zlatev, Madsen, Lenninger, Persson, Sayehli et al. (2013) ^b	Chimpanzees, <i>P. troglodytes</i> (4)	<i>Dynamic Proximal</i>

Notes: ^a Denotes studies for which only life history data was available. ^b Denotes studies for which life history and individual performance data were available (except Bräuer et al., 2006, where performance data were only available for dogs and Itakura and Tanaka, 1998, where performance data were only available for orangutans).

Table S2: Rearing histories as reported in the original studies categories according to the human experience scale.

Human Experience Scale	Rearing History
Close	Pet Enculturated (incl. language-trained) Human-Reared Hand-Raised Riding School Sea World Working Gun Dog
Occasional	Nursery Mother-Captivity Farm- Enriched Farm- Standard Research Facility Stables Zoo Kennel Free-ranging (dogs)
Seldom	Wild Stray

Paper 2

Changes in the Referential Problem Space of Infants and Toddlers

Hannah Clark¹, Zoe M. Flack ² and David A. Leavens¹

¹ University of Sussex

² University of Brighton

A version of this paper is submitted to the Journal of Comparative Psychology as:

Clark, H., Flack, Z. & Leavens, D. A. (under review) Changes in the referential space of infants and toddlers.

Parts of these data have been accepted for presentation at the 4th Lancaster Conference on Infant and Early Child Development and Protolang 6.

Author Note

HC designed the experiments, collected the data for Study 1, completed all the analyses and wrote the manuscript. ZF collected the data for Study 2. DAL advised on the study design and provided comments on drafts of the manuscript.

Abstract

Recent analyses suggest that procedural and methodological differences in the experimental set-ups of the Object Choice Task (OCT) may better explain apparent species differences in performance than theories that appeal to hypothetical, evolutionarily adaptive histories. We tested human children with versions of the OCT that have been previously used with dogs and nonhuman primates to see if manipulating the set-up would lead to behavioural changes. In Study 1, we compared the responses of 18-month-old and 36-month-old children when tested with a barrier versus without. The presence of a barrier between the child and the reward did not suppress the children's ability to choose the correct container but did lead to a difference in the type of responses elicited from the children, specifically, an increase in communicative behaviour. Moreover, the barrier had a greater facilitating effect on the younger children, who displayed more communicative behaviour in comparison with older children, who more frequently reached through the barrier in acts of direct prehension. In Study 2, we compared the behaviour of 36-month-old children when the reward was within reaching distance (proximal) and when it was out of reach (distal). The children used index-finger points significantly more in the distal condition and grabbed more in the proximal condition, showing that they were making spatial judgements about the accessibility of the reward rather than just grabbing *per se*. We discuss the implications of these within species differences in behavioural responses for cross-species comparisons.

The Object Choice Task (OCT) is used to assess an individual's ability to comprehend human gestural cues and involves an experimenter presenting a deictic cue to indicate to the subject in which of two or three containers a reward has previously been hidden (Anderson, Sallaberry & Barbier, 1995). Results of OCT studies have been used as the bases for theories pertaining to the evolutionary roots of social cognition in a number of species, in particular nonhuman primates and domestic dogs. For example, Hare and Tomasello (2005), on the basis of domestic dogs' consistently adept performance, argued for a theory of convergent evolution between dogs and humans, in which the former have developed specialised socio-cognitive skills to comprehend human gestural cues as a result of centuries of artificial selection during domestication. The Cultural Intelligence hypothesis (Herrmann et al., 2007; van Schaik & Burkart, 2011), based, in part, on nonhuman primates' generally poor performance on the OCT, states that humans have developed species-specific socio-cognitive skills in order to facilitate cultural group living, and, as such, the comprehension of human gestural cues is a human-unique ability within the primates. Bräuer, Kaminski, Riedel, Call and Tomasello (2006) proposed that differences in the performance levels between dogs and nonhuman primates are due to species-specific specialisations where dogs have been selected for specialised social abilities, which thus enable them to follow human gestural cues, whereas apes' foraging behaviour has led to increased physical abilities, which explains their ability to use physical but not social cues on the OCT.

Recent meta-analyses (e.g. Clark, Elsherif & Leavens, 2019; Lyn, 2010; Mulcahy & Hedge, 2012), however, have identified procedural and methodological differences in the testing protocols used with different taxonomic groups on the OCT that may provide more comprehensive explanations of the performance differences found than theories that attribute them to phylogenetic causes. First, human infants'

abilities to comprehend pointing gestures develop over the first year of life (Butterworth, 2001; Butterworth & Morrisette, 1996) in an environment rich in human interaction. It is this developmental process that Bard and Leavens (2014) argued is essential to consider when making cross-species comparisons in socio-cognitive tasks, and they highlight the contemporary prevalence of basing phylogenetic theories on the performance of subjects unmatched for developmental experience. Indeed, Lyn (2010) argued that pre-experimental exposure to humans can differentially affect an individual's performance on the OCT, and evidence is accumulating to support this argument. Lyn, Russell & Hopkins (2010) found enculturated nonhuman primates to be successful at following human gestural cues and a growing body of work shows that domestic dogs with less exposure to humans perform significantly worse than their pet dog counterparts (e.g., D'Aniello et al., 2017; Durantón & Gaunet, 2016; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2008), whereas other canids raised in environments rich in human interaction perform well on the task (Barrera & Bentosela, 2016; Udell, Dorey, Spencer & Wynne, 2012; Udell, Dorey & Wynne, 2008).

Leavens, Bard and Hopkins (2017) argued that a further inconsistency in the testing protocols adopted with different species is that of the presence of a barrier between the subject and the baited container in the testing environment. Due to safety issues surrounding testing nonhuman primates, species from this taxonomic group are tested from within cages, therefore imposing a barrier between the subject and the experimenter and test apparatus. Working with individuals from other taxonomic groups, domestic dogs or human infants, for example, does not necessitate the use of such safety precautions, and, as such, there is an absence of this barrier in the testing environment with these species; this constitutes a confound between experimental protocol and species classification in a significant number of contemporary studies. In a

review of 71 published nonhuman primate and dog OCT studies, Clark et al. (2019) found that 99% of nonhuman primates were tested with a barrier present in the testing environment, compared with less than 1% of dogs. They therefore argued that this inconsistency in the test set-ups used across different species represents an experimental confound that may affect individuals' performance and thus makes comparisons across groups tested with such differences invalid.

In the one study to date in which domestic dogs' abilities to follow a pointing cue with a barrier present was measured, Kirchhofer et al. (2012) found that those tested with a barrier had significantly lower success rates than those tested without. Similarly, Udell et al. (2008) found that domestic dogs tested with a partial visual barrier performed significantly worse than those tested without when required to follow a tapping cue on an OCT. This demonstrates the reduced validity of interpreting group differences as phylogenetic traces without regard to the systematically confounded differences in experimental set-ups being used with different taxonomic groups (Leavens et al., 2017). We are not aware of any study with human children on the OCT to date in which a barrier has been present in the testing environment, although human children, at least in Western populations, are well-habituated to conditions of restraint in car seats, feeding chairs, playpens, cots, and so on (Leavens, Hopkins, & Bard, 2005).

In order to investigate the possible confounding effects of this systematic difference in experimental protocols administered to representatives of different species, we tested children with and without a barrier on an OCT. Children from 14 months of age have been shown to reliably follow pointing cues on the OCT (Behne, Lizkowski, Carpenter & Tomasello, 2012) and so we tested children aged 18 months and 36 months in order to ensure that any differences in performance or behavioural responses between the two conditions were as a result of our experimental manipulation, rather than the

lack of emergence of these skills.

Study 1: Barrier vs. No Barrier

In Study 1, we looked at the effects of the imposition of a barrier, testing human children aged 18 months and 36 months on a within-subjects design, in which the children completed an OCT with and without a barrier present. To recreate as closely as possible the conditions in which nonhuman primates are tested, that is, from within a test cage, in the barrier condition children were tested from within a child's playpen, thus imposing a physical barrier between the child (inside the enclosure) and the experimenter and testing apparatus (outside the enclosure).

Method

Participants

The study was approved by the Science and Technology Cross-Schools Research Ethics Committee (C-REC) at the University of Sussex. Participants were nineteen 18-month-old children ($M = 18$ mos 18 days, $range = 18$ mos 3 days – 18 mos 27 days) and twenty 36-month-old children ($M = 36$ mos 8 days, $range = 33$ mos 10 days– 39 mos 0 days), comprised of 22 males and 17 females (18-month-old children: 11 males, 8 females; 36-month-old children: 11 males, 9 females). Children were recruited from a participant database where parents had registered their interest in participating in developmental studies with their children, and from advertisements on social media sites. Parents gave informed consent for their children to participate. Data were collected between April and November 2016.

Procedure

On arrival at the testing suite, the children and their parents were given time to

become familiar with the surroundings, with children playing freely in the playroom and interacting with the experimenter during this free play time. When parents judged their child to be settled and comfortable, the experimenter, child, and parent moved to the testing room, where the experimenter demonstrated a “ball run” toy to the child, and then encouraged the child to play with the toy. The test room was set up as shown in Figure 1. The experimenter then informed the child that they were going to play a fun hiding game with the balls, and that if the child found the balls they could put them in the ball run. The experimenter then asked the child to sit on the playmat with their parent and explained that she would hide the ball under one of two cups, and then give the child a “clue” to see if they could find it. In the *barrier* condition, a child’s playpen was set up, such that the playmat was inside the pen, and the ball run was outside, but accessible to the child. In the *no barrier* condition, the playmat and ball run were in the same positions, but the playpen was not in place. The experimenter hid the ball under one of two cups, behind a cardboard occluder, then made eye contact with the child, asking “are you ready for your clue?” The experimenter then presented an ipsilateral, dynamic, index-finger pointing cue, whilst alternating her gaze between the container and the child. A dynamic point is one in which the pointing gesture is carried out in front of the child and remains in place until the child makes a choice. The distance between the experimenter’s fingertip and the container was approximately 5cm. The experimenter maintained this position until the child made a choice. If the child was unresponsive, the experimenter encouraged the child to make a choice by giving verbal encouragement such as “can you find that ball?” If the child failed to respond after approximately 2 minutes, or was fussy (for example, trying to get out of the playpen), then the trial was terminated, and the experimenter attempted to increase motivation by again demonstrating the ball in the ball run. If the child made a correct choice, they

were given the ball and encouraged to put it in the ball run. If the child made an incorrect choice, the experimenter lifted the incorrect cup and said, for example, “Oh no! It’s not in that one! Let’s see if it was in the other one!” and then lifted the correct cup, showed the child the ball, and said “Never mind! Let’s hide it again!” Children took part in blocked trials, with 8 trials in the *barrier* condition and 8 trials in the *no barrier* condition. Order of administration was counterbalanced across participants, and in between conditions, children left the test room with their parent and were engaged in another task, such as looking at wall stickers of animals. The baited container was on the right or left an equal number of times in each condition, and the order was counterbalanced, such that the reward was never on the same side for more than two consecutive trials.

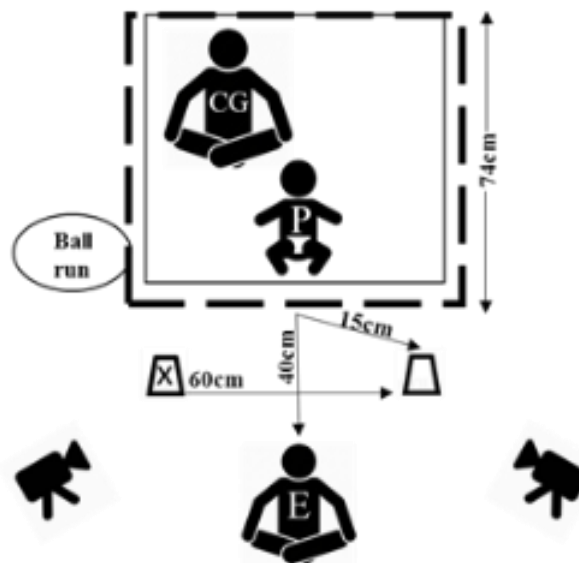


Figure 1: The experimental set-up. P = participant; CG = caregiver; E = experimenter. Barrier represented by dotted line. Drawing not to scale.

Materials

The playpen used in the barrier condition was a Dream Baby Royal Converta 3-

in-1 Playpen Gate, measuring 380 x 4 x 74cm (Rosyth Business Centre, 16 Cromarty Campus, Rosyth, Fife, KY11 2WX). Children and their parents sat on a playmat made up of 16 interlocking JSG Accessories Outdoor/ Indoor Protective Flooring Mats (JSG Accessories, Unit 6 Hughes Business Centre, Wilverley Road, BH23 3RU). The containers used to hide the reward were two white opaque plastic cups measuring 7.8 x 10cm. A John Lewis Junior Ball Run was used as the stimulus, measuring 52 x 56 x 47.5cm (John Lewis Partnership, 71 Victoria Street, London, SW1E 5NN). The occluder was a piece of brown cardboard measuring 65 x 80cm. All testing sessions were recorded on two Sony Handycam HDR-PJ410 video-cameras (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan).

Data Scoring

Test sessions were video-recorded and coded at a later date. For each trial, data were coded for whether or not the choice made was correct, latency of response (from maximum extension of the index-finger to the child choosing a cup), type of response, the direction of the child's gaze whilst giving the response, and whether the response was accompanied by a vocalisation. Response types were categorised according to the following scheme:

Index-finger point: The arm and index-finger are extended towards the referent, with the other fingers curled under the hand (Masataka, 2003).

Whole-hand point: An indicative gesture categorised by outstretched arm and extended fingers, which is not a direct attempt to obtain the container (Leavens & Hopkins, 1999).

Indicative gesture other than index-finger/ whole-hand point: Where a child indicated a choice using a gesture other than an index-finger point or whole-hand point.

An example of this is one child “pointed” to the container with their foot.

Direct Grab: A response was categorised as a grab when the child reached for and contacted the container with their hand.

Reach: An attempt to obtain the container, categorised by hand outstretched and fingers grasping (Leavens & Hopkins, 1999).

Other: Responses other than those described above. An example of an *other* response is a child who used the parent’s arm to indicate the choice.

Analyses

Children were excluded from the analyses if they failed to complete at least four trials in each condition. This led to the exclusion of four 18-month-old children and one 36-month-old child. There was no significant difference in the number of trials completed between the 18-month-old ($Mdn = 15$) and 36-month-old children ($Mdn = 16$), $Z = -2.70$, $p = .458$.

Reliability

An independent coder who was blind to the purpose of the study coded 20% of the videos. For correct choices, there was complete agreement between the two coders, Cohen’s kappa, $\kappa = 1.00$, $p < .001$. There was excellent agreement for both response latency, $r_s = .87$, $p < .001$, and response type, $\kappa = 0.66$, $p < .001$.

Results

Correct Choices

18-month-old children

The 18-month-old children, as a group, performed above chance both with a

barrier (binomial test, $p < .001$) and without a barrier (binomial test, $p < .001$). There was no significant difference in the proportion of correct choices made between the *barrier* ($Mdn = 1.00$) and *no barrier* ($Mdn = 1.00$) conditions, $Z = -0.60$, $p = .552$. This shows that the barrier did not have an effect on the younger groups' ability to effectively use the pointing cue to find the hidden reward.

36-month-old children.

The 36-month-old children also performed above chance as a group in both the *barrier* (binomial test, $p < .001$) and the *no barrier* (binomial test, $p < .001$) conditions. There was no significant difference in the proportion of correct trials between the *barrier* ($Mdn = 1.00$) and the *no barrier* ($Mdn = 1.00$) conditions, $Z = -1.36$, $p = .175$. This shows the older children were also able to effectively use the pointing cue despite the presence of a barrier.

Age comparisons.

There was no significant difference in the proportion of correct trials between the 18-month-old ($Mdn = 1.00$) and the 36-month old children ($Mdn = 1.00$) in the *barrier* condition, Mann-Whitney $U = 112.5$, $p = .302$, nor in the *no barrier* condition (18-month-old children $Mdn = 1.00$, 36-month-old children $Mdn = 1.00$), Mann-Whitney $U = 135.0$, $p = .811$. This shows that the barrier did not have an effect on performance for either age group and that the children of both age groups were equally adept at using the cue to find the hidden reward.

Response Latency

18-month-old children.

There was a significant effect of barrier on mean latency to respond within 18-

month-old children, with increased latencies in the *barrier* condition ($Mdn = 13.00s$) compared with the *no barrier* condition ($Mdn = 4.75s$), $Z = -2.44$, $p = .015$. This shows that the younger children were slower in responding when a barrier was present.

36-month-old children.

There was no significant difference in response latency between the two conditions for the 36-month-old children (*barrier* $Mdn = 2.88s$, *no barrier* $Mdn = 3.13s$), $Z = -0.22$, $p = .825$. This shows that the older children's response times were unaffected by the barrier.

Age comparisons.

In the *barrier* condition, the 18-month-old children ($Mdn = 13.00s$) were significantly slower to respond than the 36-month-old children ($Mdn = 2.88s$), Mann-Whitney $U = 18.0$, $p < .001$. The 18-month-old children ($Mdn = 4.75s$) were also significantly slower to respond than the 36-month-old children ($Mdn = 3.13s$) in the *no barrier* condition, Mann-Whitney $U = 61.5$, $p = .005$. This shows that the 18-month-old children were generally slower to respond than the older children, and these response times were further increased by the presence of a barrier in the testing environment.

Response Type

18-month-old children.

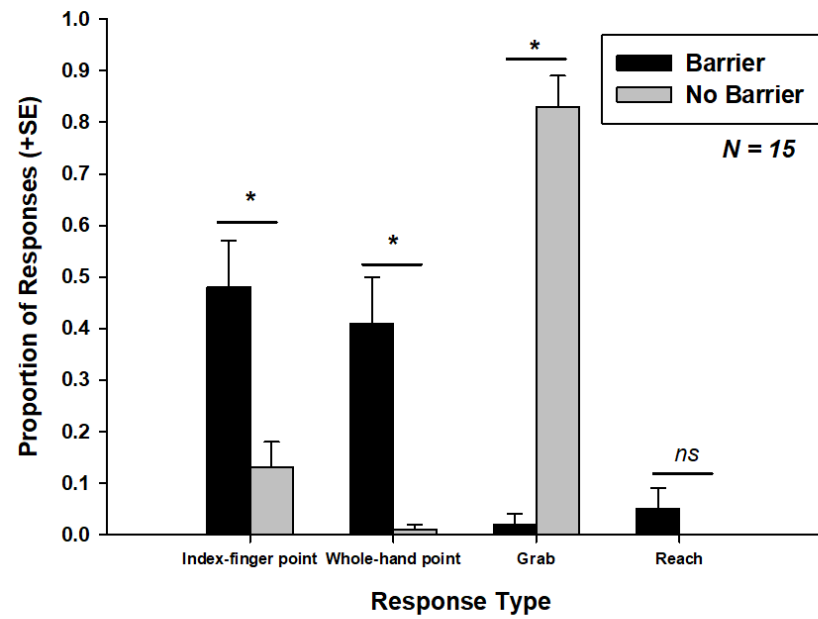
There were a number of differences in the response types elicited from the younger group as a function of the presence of a barrier. 18-month-old children used significantly more *index-finger points* in the *barrier* ($Mdn = .43$) than in the *no barrier* ($Mdn = .13$) condition, $Z = -3.18$, $p = .001$ as well as significantly more *whole-hand points* in the *barrier* ($Mdn = .43$) than in the *no barrier* condition ($Mdn = .00$), $Z = -$

2.94, $p = .003$. They grabbed the container significantly less in the *barrier* condition ($Mdn = .00$) than in the *no barrier* ($Mdn = .88$) condition, $Z = -3.45$, $p = .001$. There was no significant difference in 18-month-old children's tendency to reach for the container between the *barrier* ($Mdn = .00$) and *no barrier* ($Mdn = .00$) conditions, $Z = -1.34$, $p = .180$. This shows that the younger group were more likely to respond using a communicative cue such as an index-finger point or a whole-hand point when there was a barrier present, and more likely to grab the container when there was no barrier present. Analyses were not performed where responses were categorised as *other indicative gesture* or *other*, as these only constituted 0.36% and 1.47% of the total responses (for both age groups combined), respectively. Figure 2a shows the distribution of the response types for the younger children.

36-month-old children.

There was no significant difference in the proportion of responses that were index-finger points for the 36-month-old children between the *barrier* ($Mdn = .00$) and *no barrier* ($Mdn = .00$) conditions, $Z = -1.83$, $p = .066$, but they did use significantly more whole-hand points when the barrier was present ($Mdn = .00$) than when it was not ($Mdn = .00$), $Z = -2.20$, $p = .028$. The 36-month-old children grabbed significantly more when in the *no barrier* condition ($Mdn = 1.00$) than in the *barrier* condition ($Mdn = .38$), $Z = -3.24$, $p = .001$. 36-month-old children were significantly more likely to reach in the *barrier* ($Mdn = .00$) than in the *no barrier* condition ($Mdn = .00$), $Z = -2.69$, $p = .007$. This shows that the older children were also more likely to use some communicative gestures when the barrier was present and again, more likely to grab, or try to grab, the container when the barrier was absent. Figure 2b shows the distribution of response types for the older group.

a) 18-month-olds



b) 36-month-olds

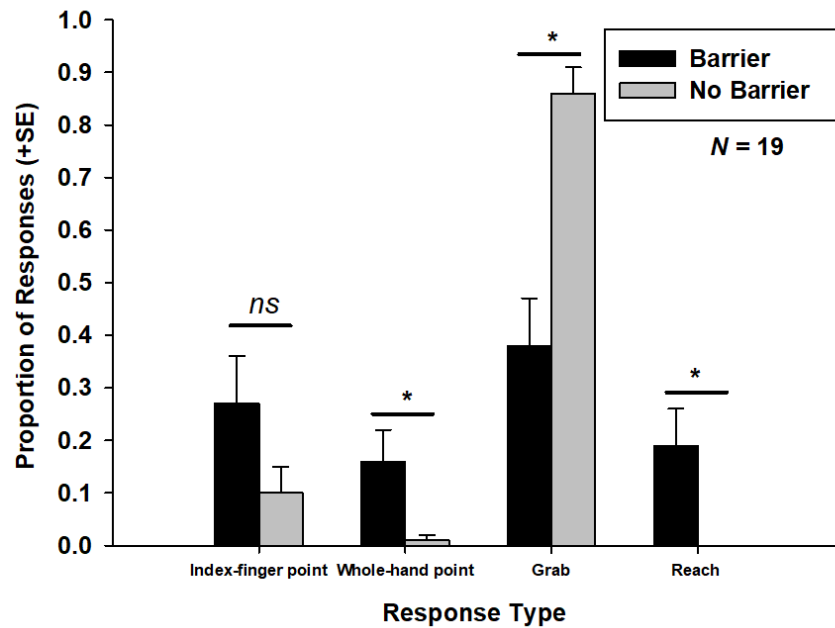


Figure 2: The mean proportion of response types in the *barrier* and *no barrier* conditions by a) 18-month-old and b) 36-month-old children, with standard errors. Means and standard errors are depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. * $p < .05$.

Age comparisons.

The proportion of 18-month-old children's responses that were *index-finger points* was significantly higher than that of 36-month-old children in the *barrier* condition, Mann-Whitney $U = 80.5$, $p = .030$, but there were no significant effects of age on use of *index-finger points* in the *no barrier* condition, Mann-Whitney $U = 125.5$, $p = .560$. This shows that there were effects of both age and barrier on the use of this type of response, with 18-month-old children using *index-finger points* to indicate the container in which they thought the reward was hidden more than the 36-month-old children in the barrier condition.

The proportion of 18-month-old children's responses that were *whole-hand points* was significantly greater than that of 36-month-old children's in the *barrier* condition, Mann-Whitney $U = 81.0$, $p = .022$, but not in the *no barrier* condition, Mann-Whitney $U = 140.5$, $p = .945$. This shows that children from both age groups were more likely to indicate their choice using a *whole hand point* when there was a barrier in the testing environment than when there was not, and that this effect of the barrier was particularly pronounced for the 18-month-old children.

18-month-old children were significantly less likely to grab in the *barrier* condition than the 36-month-old children, Mann-Whitney $U = 64.5$, $p = .006$, but not in the *no barrier* condition, Mann-Whitney $U = 133.0$, $p = .758$. This shows that both age groups tended to grab the container in which they thought the reward was hidden more when there was no barrier present in the testing environment than when there was, but that this effect was less pronounced for the 36-month-old children. There was no significant effect of age on reaching behaviours in either the *barrier*, Mann-Whitney $U = 95.5$, $p = .050$, or *no barrier*, Mann-Whitney $U = 142.5$, $p = 1.00$, conditions. This shows that the 36-month-old children were more likely to reach for the container in which they thought the reward was hidden in the barrier condition, however the 18-

month-old children were not.

Order of Administration and Trial by Trial Analyses

There was no effect of order of administration on correct choices, response latency or type of response in either age group in either condition. There were no significant difference across trials in correct choices or response latency (all $ps > .05$). With regard to response type, for the 18-month-old children, there was a significant difference across trials in the use of index-finger points in the *barrier* condition, Cochran's $Q = 17.50$, $p = .014$. Pairwise comparisons revealed significant differences between trials 1 and 2 ($Z = -2.45$, $p = .014$), trials 1 and 4 ($Z = -2.12$, $p = .034$), trials 1 and 5 ($Z = -2.00$, $p = .046$) and trials 1 and B8 ($Z = -2.24$, $p = .025$). This suggests the children's use of index-finger points increased significantly between the first trial and these later trials. There was also a significant difference in the use of index-finger points across trials in the *no barrier* condition, Cochran's $Q = 14.127$, $p = .049$. Pairwise comparisons showed there was a significant difference between trials 1 and 7, $Z = -2.00$, $p = .046$. There were no other significant comparisons. For the 36-month-old children, there was a significant difference across trials in the tendency to grab in the *barrier* condition, Cochran's $Q = 19.24$, $p = .007$. Pairwise comparisons revealed there were significant differences between trials 1 and 3 ($Z = -2.24$, $p = .025$), trials 1 and 8 ($Z = -2.45$, $p = .014$), trials 2 and 8 ($Z = -2.00$, $p = .046$), trials 3 and 4 ($Z = -2.00$, $p = .046$), trials 3 and 5 ($Z = -2.00$, $p = .046$) and trials 4 and 8 ($Z = -2.24$, $p = .025$).

There was a significant difference across trials in the use of index-finger points in the *barrier* condition, Cochran's $Q = 17.50$, $p = .014$. Pairwise comparisons revealed significant differences between trials 1 and 2 ($Z = -2.45$, $p = .014$), trials 1 and 4 ($Z = -2.12$, $p = .034$), trials 1 and 5 ($Z = -2.00$, $p = .046$) and trials 1 and B8 ($Z = -2.24$, $p =$

.025). This suggests the children's use of index-finger points increased significantly between the first trial and these later trials. There was also a significant difference in the use of index-finger points across trials in the *no barrier* condition, Cochran's $Q = 14.127$, $p = .049$. Pairwise comparisons showed there was a significant difference between trials 1 and 7, $Z = -2.00$, $p = .046$. There were no other significant comparisons. There was a significant difference across trials in the tendency to grab in the *barrier* condition, Cochran's $Q = 19.24$, $p = .007$. Pairwise comparisons revealed there were significant differences between trials 1 and 3 ($Z = -2.24$, $p = .025$), trials 1 and 8 ($Z = -2.45$, $p = .014$), trials 2 and 8 ($Z = -2.00$, $p = .046$), trials 3 and 4 ($Z = -2.00$, $p = .046$), trials 3 and 5 ($Z = -2.00$, $p = .046$) and trials 4 and 8 ($Z = -2.24$, $p = .025$).

Discussion

In Study 1, we investigated the effects of a barrier on 18-month-old and 36-month-old children's behaviour on the OCT. There were no differences in performance as a function of either age or the imposition of a barrier. That both groups of children demonstrated ceiling-level performance when there was no barrier present was expected; however, it is interesting to find that the imposition of a barrier did not have a decreasing effect on success levels for the human children in the way that Kirchhofer et al. (2012) found for domestic dogs. This shows that human children from 18 months are reliably and flexibly able to follow index-finger pointing cues, even with a partial visual barrier, although it must be noted that the two studies differed procedurally in terms of the distances between the containers (1.5m in Kirchhofer et al.) and the locomotor demands on the participants in retrieving the rewards (dogs in Kirchhofer et al.'s study were required to retrieve the object, turn around, and locomote to give it to the experimenter).

The 18-month-old children were slower than the 36-month-old children to choose a container in both the barrier and no barrier conditions, and also showed a marked difference in latency to respond between the conditions, being significantly slower when there was a barrier present than when there was not. Interestingly, however, differences in latency were not associated with performance differences, likely due to the ceiling level performances by both age groups. It may be that these differences in response latencies were due to the unfamiliarity of the situation affecting the younger children more than the older children, or alternatively, due to superior skill in responding to deictic gestures in older children as a function of increased experience with such cues. Leung and Rheingold (1981) found an increase in the ability to comprehend pointing cues associated with age and pointing production in children from 10.5 to 16.5 months, suggesting that comprehension abilities increase with children's own use of these cues. It is worth noting that response latency is not discussed in any of the OCT studies with humans that we reviewed, but according to this explanation, it seems evident that as children become more proficient in both producing and comprehending and gestural cues, they also become quicker to interpret them.

There were differences in the types of response produced by the children as a function of both the imposition of a barrier and age. In the no barrier condition, children of both ages showed a preference for grabbing the container, that is, they overturned the container themselves in order to look inside for the reward. When there was a barrier present, however, both age groups showed an increase in gesturing behaviour, that is, they were more likely in this condition to indicate their choice to the experimenter by gesturing, in the form of an index-finger or whole hand point, rather than reaching through the bars to overturn the container themselves. This bias towards gesturing in the barrier condition was particularly prominent in the 18-month-old children, with 36-

month-old children often choosing to grab the container themselves, despite the presence of the barrier, something which the younger children did significantly less frequently. Interestingly, in previous studies of human children's performance on the OCT (e.g. Behne, Carpenter & Tomasello, 2005; Behne et al., 2012; Herrmann et al., 2007; Pflandler, Lakatos & Miklósi, 2013), descriptions of the children's behaviour when responding to the cue tend to refer to them "searching" or "looking" in the containers. Thus, it can be inferred that typically on the OCT, when no barrier is present, children choose to look inside the container for themselves, as no mention is made of any of the children in these previous studies choosing instead to gesture to the experimenter to indicate their choice. Studies with nonhuman primates differ in the ways in which subjects make their choices, varying from the subject being able to reach through a plexiglass hole to overturn the container themselves (e.g. Barth, Reaux & Povinelli, 2005) to them being required to "indicate" the correct container through wire mesh (e.g. Herrmann et al., 2007). That human children are varying their behavioural responses according to whether or not a barrier is present demonstrates that there is an effect of this experimental manipulation and demonstrates the need for consistency in testing environments when comparing across species. The increased use of gesturing, particularly by the younger children, may be explained in terms of the referential problem space theory proposed by Leavens et al. (2005), that the children see the bars of the playpen as a barrier between themselves and a desirable, but out-of-reach object, and thus use a communicative gesture in order to influence another to retrieve said desirable object. That the object itself was not actually out of reach, but was instead simply partially obstructed, has interesting implications for the way the children perceived the barrier, perhaps as a form of psychological restraint.

Regarding the use of whole-hand points, 18-month-old children used

significantly more of these than did 36-month-old children, consistent with Cochet and Vauclair's (2010) findings, in a sample of French children, that the incidence of whole-hand points tends to decrease with age, with a preference for index-finger points emerging. They found that, when points were analysed separately according to function, this correlation between age and hand shape remained for declarative points, but not for imperative points, and this, they suggest, can be taken as evidence for distinctive origins of these two pointing types. Specifically, similarly to Franco and Butterworth (1996), they hypothesised that declarative gestures have a communicative root, whereas imperative gestures originate in failed grasps. They thus assert that the absence of a correlation of age with the use of whole-hand points in an imperative context can be explained by the children preferentially utilising a hand shape that would permit them to grasp the desired object, rather than an index-finger point, which would not allow them to do so. Alternatively, our finding that the incidence of whole-hand pointing decreased with age could be explained through its being a product of increasing experience with conventionalised human gestures. That is, the 36-month-old children, as a result of their superior level of experience of, and exposure to, human deictic gestures, are responding in a more conventionalised manner than the 18-month-old children, in terms of the shape of the hand when gesturing (Leavens & Hopkins, 1999).

Study 2 Manipulating the distance of the reward

Given our finding that the children's behavioural responses differed as a function of the imposition of a barrier, which they may have perceived as a physical restraint to obtaining the reward themselves, we thought it would be of interest to investigate the effects of the distance of the reward. In Study 2, therefore, we focused on manipulating the distance between the child and the containers, specifically whether the child was able to reach the reward or not, in order to examine the effects of placing the

reward out of reach on children's communicative behaviour. All children were tested from within the playpen, and took part in a *proximal* condition, comparable to the *barrier* condition in Study 1, in which the containers were outside of the playpen but within reach of the child, and a *distal* condition, in which the containers were placed outside of the barrier and out of reach of the child.

Method

Participants

Participants were seventeen 36-month-old children ($M = 36$ mos 4days, $range = 31$ mos 30 days – 39 mos 26 days), comprised of 6 males ($M = 37$ mos 4 days, $range = 31$ mos 30 days – 39 mos 19 days) and 11 females ($M = 36$ mos 11 days, $range = 32$ mos 27 days – 39 mos 26 days). Children were recruited from a participant database, where parents had registered their interest in participating with their children in cognitive studies, and from advertisements on social media sites. Data were collected between December 2017 and January 2018.

Procedure

The procedure was the same as in Study 1, except that all children were tested from within the playpen, and the rewards used were stickers instead of a ball run. The containers were placed in either a proximal or a distal position (see Figure 5). Children completed two proximal and two distal trials and order of administration was counterbalanced across participants. Children were tested with only two trials in each condition because we expected their choices to be correct on almost all trials given our findings in Study 1, and it was their initial response types that we were interested in recording, before these became habituated through the use of many repeated trials. The baited container was on the right or left an equal number of times in each condition, and

the order was counterbalanced.

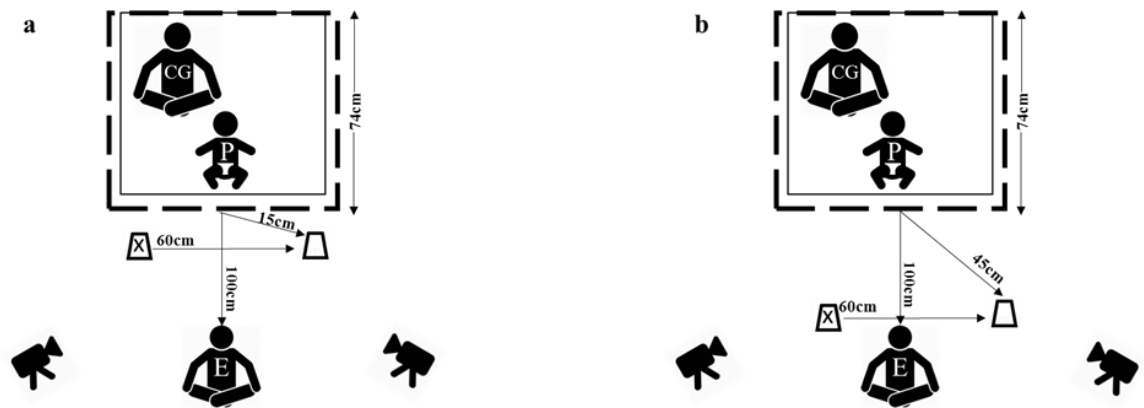


Figure 5: The experimental set-up for a) the proximal condition and b) the distal condition in Study 2. CG = caregiver, P = participant, E = experimenter. Barrier represented by dotted line. Drawing not to scale.

Data Scoring

Data were coded according to the same coding scheme as in Study 1.

Analysis

Five children were excluded from the final analyses due to experimenter error during testing. Three of these cases were due to the experimenter using a momentary, rather than a dynamic, point, and two were because the cameras were placed such that the experimenter was not in shot in the videos, and therefore the moment of pointing could not be ascertained.

An independent coder who was blind to the purpose of the study coded 20% of the videos. For correct choices, there was complete agreement between the two coders, $\kappa = 1.00$, $p = .005$. There was also complete agreement for response type, $\kappa = 1.00$, $p < .001$ and excellent agreement for latency, $r_s = .88$, $p < .001$.

Results

Correct Choices

The data were not normally distributed and so non-parametric tests were used throughout the analyses. There was no significant difference in the proportion of correct choices made in the *proximal* ($Mdn = 1.00$) and *distal* ($Mdn = 1.00$) conditions, $Z = -1.00$, $p = .317$. This shows that the children performed at ceiling level in both conditions, as expected.

Response Latency

There was no significant difference in the mean latency to respond between the *proximal* ($Mdn = 3.00$ secs) and *distal* ($Mdn = 3.75$ secs) conditions, $Z = 0.00$, $p = 1.00$. This shows that the children were equally quick to respond in both conditions.

Type of Response

Only two of the possible response types were used by the children, these were index-finger points and grabs. Figure 6 shows the mean proportion of each type of response used in the two conditions. The proportion of trials in which the children used an index-finger point to indicate their choice of container was significantly lower in the *proximal* condition ($Mdn = 1.00$) than in the *distal* condition ($Mdn = 1.00$), $Z = -2.24$, $p = .025$. This shows that when the containers were out of reach, the children were more likely to respond by using an index-finger point than when the containers were within reach.

The proportion of trials in which the children grabbed the container was significantly higher in the *proximal* ($Mdn = 1.00$) than in the *distal* condition ($Mdn = .00$), $Z = -2.24$, $p = .025$. This shows that the children's response types differed as a

function of whether the reward was or was not within reach- they always pointed when the container was out of reach, but if the container was within reach, then they grabbed it on 20% of trials.

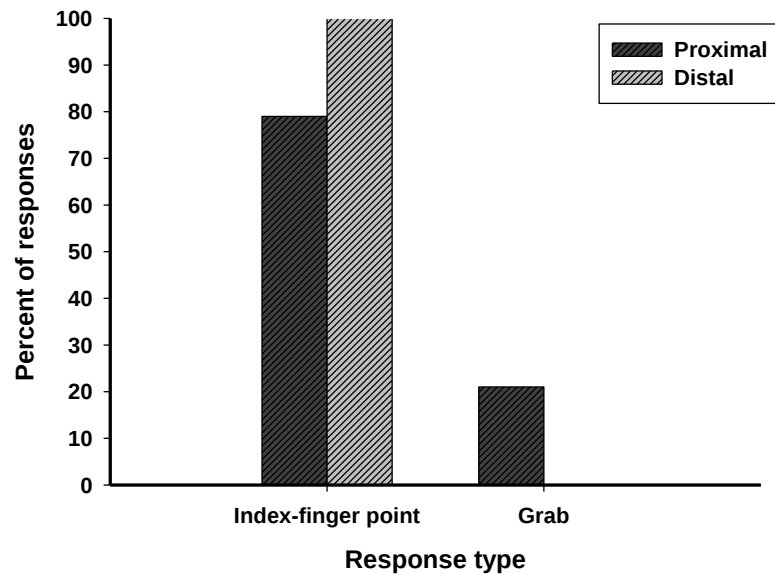


Figure 6: The percentage of trials in which the children responded with index-finger points and grabs in the *proximal* and *distal* conditions.

Order of administration and grabbing behaviour

In order to investigate whether the order of presentation had an effect on the response types used, children were categorised as ‘grabbers’ (grabbed the container on at least one trial in the *proximal* condition or ‘non-grabbers’ (did not grab on either trial in the *proximal* condition). There was no significant effect of order of administration on the likelihood of grabbing (*proximal* first $Mdn = 1.00$; *distal* first $Mdn = .00$), Mann-Whitney $U = 9.00$, $p = .093$. There was a trend, however, that those tested with the *proximal* condition first were more likely to grab in the *proximal* condition, whereas those tested with the *distal* condition first were slightly less likely to grab at all in the

proximal condition (see Figure 7).

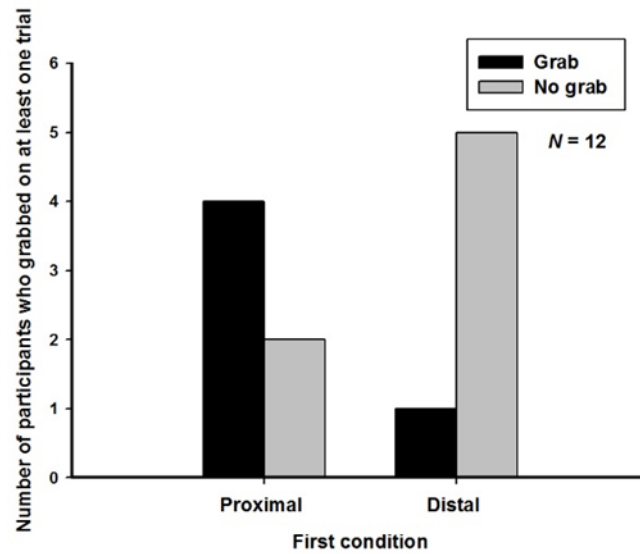


Figure 7: The number of children tested with either the *proximal* or *distal* condition first who grabbed in at least one trial.

Discussion

In Study 2, we investigated the effect of placing a hidden reward either within or out of reach on an OCT with 36-month-old children. As expected, the children performed at ceiling level in both conditions, further demonstrating that 3-year-olds were able to reliably follow a pointing gesture to find a hidden reward. There were no differences in the children's response latencies between the two conditions, showing that this ability is flexible even across increased distances between the child and the object being signalled.

The children in this study used only two response types to indicate the container in which they thought the reward was hidden, index-finger points and direct grabs of the container. These were the responses most often utilised by the 3-year-olds in Study 1, however, it is notable that there was an absence of the use of whole-hand points and

reaches in the current study. The absence of whole-hand points is congruent with Cochet and Vauclair's (2010) findings that the incidence of whole-hand pointing in a declarative context decreases with age and that, here, the children were responding to the experimenter by demonstrating where they believed the reward to be hidden, rather than demanding the cup in an imperative manner. As in Study 1, the children chose to grab the container to look inside themselves in the *proximal* condition on a number of trials. That they did not try to reach for the container on any of the *distal* trials, nor did they exhibit any whole-hand points - which Cochet and Vauclair (2010) argue could be the result of failed grasping attempts in an imperative context- demonstrates that it is not the case that 3-year-olds are categorically 'grabbers', but rather that their grabbing responses are a result of a spatial evaluation. When the container is not in reach, they do not even try to grab it.

Although there was no significant difference, there was a trend towards increased grabbing in the *proximal* condition when this was the first condition administered than when it followed the *distal* condition. Specifically, only one of the six children tested with the *distal* condition first grabbed the container in either of the two trials in the *proximal* condition, compared with four out of six children tested with the *proximal* condition first. This, like Study 1, has interesting implications for the way the children perceive the barrier, with one possible explanation being that those tested with the *distal* condition formed a perception of the barrier as a restraint that prevented them from being able to retrieve the reward themselves, and maintained this perception once the containers were actually moved within reach, such that they continued to use communicative cues to indicate their choice rather than grab for it themselves. An alternative explanation may be that the index-finger pointing became a perseverative response- once this had been effective as tool to retrieve the desired out-of-reach object

in the *distal* condition, they habitually continued to use this response in the *proximal* condition.

General Discussion

Here we present the results of two studies with children in which elements of the configuration of the OCT were manipulated, in order to investigate whether such manipulations affected the children's behavioural responses, especially their decisions to either elicit aid from the experimenter or to act directly on the apparatus. In Study 1, we tested 18-month-old and 36-month-old children on a standard version of the OCT, in which children were tested with and without a barrier in an attempt to mimic the testing conditions used with nonhuman primates. In Study 2, we tested 36-month-old children with a barrier, manipulating the distance of the reward, such that it was placed either within or out of reach of the child.

The children performed at ceiling level in both studies, which was expected, given that the ability to follow pointing gestures has been reliably found in children from around 14 months of age (Behne et al., 2012). We did, however, find that the children's behavioural responses differed as a result of the experimental manipulations. In Study 1, children of both ages used significantly more communicative gestures to indicate their choice of container when there was a barrier present than when there was not. This difference was particularly pronounced for the 18-month-old children, with the 36-month-old children still choosing to grab on a proportion of trials even when the barrier was present. In Study 2, the 36-month-old children varied this grabbing behaviour as a function of the distance of the reward, that is, they chose to grab when the container was within reach, despite the presence of a barrier, but not when it was out of reach, nor did they attempt to reach for it. Although the differences we found here

were not statistically significant, there was a trend towards decreased grabbing when the order of administration was such that the children were tested with the distal condition first, which has implications for their perception of the barrier, an avenue which would be interesting to explore further with an increased sample size, and thus, increased statistical power.

That the children, all of an age at which their comprehension of the pointing cue is at mastery level, displayed differential behavioural responses according to the configurational set up of the experiment, shows that these manipulations do have an effect on communicative behaviour. This has implications for the wealth of OCT literature which compares across species with little-to-no regard for matching experimental conditions (Leavens et al., 2017). Human children, with 18- or 36-month long histories rich in human interaction and exposure to human pointing cues, who are experts in using these cues, themselves, to influence the behaviour of others, react differently when tested with a barrier in the testing environment to when tested without. We argue, therefore, that attributing differences in performance level between species to species differences (i.e., to the effects of evolutionarily adaptive histories, e.g. Hermann et al., 2007; Maclean, Hermann, Suchindran & Hare, 2017) when one sample is tested from within a cage (e.g. chimpanzees) and the other without (e.g. dogs) ignores the confounding effect of unmatched experimental conditions.

In conclusion, here we add to and extend the arguments put forward in recent reviews that detail the procedural and methodological flaws in the OCT literature (Lyn, 2010; Mulcahy & Hedge, 2012), and analyses of ape-human comparisons more generally (Leavens, 2014, 2018; Leavens et al., 2017) and emphasise the necessity of ensuring matched conditions in experimental testing. Furthermore, we demonstrate that these subtle manipulations of the testing environment can lead to differences in the

behavioural responses of members of the taxonomic group most experienced in the use of human gestural cues, and, therefore call into question theories that generalise the ability of one sample to their whole species on studies which fail to control for testing environment.

Acknowledgements

We thank the School of Psychology and the Doctoral School, University of Sussex for funding this research. Dan Hyndman and Martha Casey provided technical support. We thank Dr. Jessica Horst and Professor Nicola Yuill for advice and sharing of laboratory and office space. Finally, we are grateful to the parents and children who volunteered to take part in these studies.

Paper 3

The Performance of 3-Year-Olds on Two Versions of an Object Choice Task

Hannah Clark and David A. Leavens

University of Sussex

This paper is written in the style of an article appropriate for Developmental Psychology.

Author Notes

HC and DAL designed the experiment, HC collected the data, conducted the analyses and wrote the manuscript. DAL provided comments on drafts of the manuscript.

Abstract

Recent reviews of object choice task studies highlight the prevalence of a number of procedural and methodological issues in cross-species comparisons, such as failing to match for pre-experimental history and incommensurate testing environments. There is a tendency to test nonhuman primates with a central version of the task, and dogs with a peripheral version, differentiated by differences in subject-experimenter, subject-container, and inter-object distances. Performance differences associated with the different versions have been found in representatives of both taxonomic groups. In order to investigate the possible effects of confounds in spatial configuration, we tested 3-year-old children with a central and a peripheral version of the task. Half of the children were tested from within a playpen, to match as closely as possible the testing conditions experienced by nonhuman primates, and half without, in order to enable us to delineate the effects of the configurational manipulation. We found no effects of configuration on children's accuracy, latency or behavioural responses, but we did find an effect of the barrier, such that its imposition led to a decrease of acts of direct prehension. We discuss this in terms of previous research which has found differential responses as a function of a barrier and provide further evidence of the need to match experimental conditions when testing across species.

One developmental milestone for human infants, which emerges over the first year of life is the ability to comprehend pointing cues (Butterworth, 2001; Butterworth & Morrisette, 1996), which forms an important part of joint attention (Leavens & Clark, 2017), and is a precursor to speech (Colonnese, Rieffe, Koops, & Perucchini, 2008; Iverson & Goldin-Meadow, 2005). Debate exists as to the extent to which other species possess the ability to comprehend human pointing cues and as to the complexity of socio-cognitive skills required to do so. For example, Behne, Carpenter and Tomasello (2005) argued that one-year-olds' success in using pointing cues in a hiding-finding game demonstrates their understanding of the referential nature of the cue, specifically that the adult's gesture is motivated by a communicative intention to inform the child of the hidden reward's location. Similarly, Tomasello and Carpenter (2007) suggested that successful use of these cues is evidence of shared intentionality, that is, the sharing of psychological states which enables children to interpret the cue as intended to indicate the location of a hidden reward. A number of studies with nonhuman primate subjects report that they fail to successfully comprehend human pointing cues (e.g. Herrmann, Call, Hernandez-Lloreda, Hare & Tomasello, 2007; Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012), leading Herrmann et al. (2007), Moll and Tomasello (2007), and Tomasello and Moll (2007) to argue that this is, among primates, a human-unique skill which emerged in our evolutionary lineage after our separation from the other hominids in order to facilitate living in complex social groups in which collaboration, cooperation and cultural learning drove technological advances and social cognition. Dogs' repeatedly high performance (e.g. Hare & Tomasello, 1999) has been attributed to the emergence of specialised socio-cognitive skills for understanding humans, evolved through years of domestication, known as the domestication hypothesis (Hare, Brown, Williamson & Tomasello, 2002; Hare & Tomasello, 2005).

In order to assess the ability to comprehend pointing cues, the object choice task (OCT) is frequently employed. This experimental paradigm involves an experimenter baiting one of two or three opaque containers with a reward, and then indicating the location of the reward to the subject using a pointing or gaze cue. A number of reviews of the published OCT literature have demonstrated the prevalence of methodological and procedural flaws in these studies, which draw into question the validity of theories based on their results. For example, Lyn (2010) and Russell, Lyn, Schaeffer and Hopkins (2011) highlighted the consistent failure of studies to match subjects on their pre-experimental exposure to humans, an important consideration given the developmental nature of the ability to comprehend pointing cues in human infants (Bard & Leavens, 2014). Subsequent studies have demonstrated increased success rates in “enculturated” nonhuman primates, that is, those who have been raised in environments rich in human interaction or have undergone language training programmes, (Lyn, Russell & Hopkins, 2010) and lower success rates in samples of dogs who have little pre-experimental exposure to humans (D’Aniello et al., 2017; Duranton & Gaunet, 2016; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2008). Leavens, Bard and Hopkins (2017) discussed the use of test cages with nonhuman primates, a necessary safety precaution, which, however, imposes a barrier between the subject and the testing apparatus and experimenter. Clark, Elsherif and Leavens (2019), in a meta-analysis of published nonhuman primate and dog OCT studies found that almost 100% of nonhuman primates were tested with this barrier present, compared with fewer than 1% of dogs. Kirchhofer et al. (2012) compared the performance of dogs tested with and without a barrier, and reported a decrease in performance associated with the barrier, as did Udell et al. (2008) when testing dogs’ ability to use a tapping cue on the OCT. Clark, Flack and Leavens (under review) looked at 18-month- and 36-month-old human

children's responses on the OCT when tested with and without a barrier and found the presence of a barrier led to an increase in communicative responses (i.e., pointing to indicate their choice of container) and a suppression of acts of direct prehension, demonstrating a clear effect on behaviour as a result of this experimental manipulation.

Mulcahy and Hedge (2012) argued that incommensurate testing environments used when administering the OCT to nonhuman primate subjects and dogs could disadvantage the former. They distinguished between central versions of the task, in which the distance between the containers is less than 1m, and peripheral versions, in which the distance is greater. Central versions, they argued, involve the containers being placed in the direct line of vision of the subject and between the subject and the experimenter, such that the subject must look past the containers to attend to the experimenter's cue. Therefore, the salience of the container (and the food held within) may distract the subject's attention from the cue being presented and may also require a greater level of behavioural inhibition to attend to the cue and utilise it to inform the choice of cup. In contrast, in peripheral versions, the containers, in addition to being out of the subject's direct line of vision, are placed further away, and thus retrieving the reward implies greater locomotor effort, possibly increasing the subject's motivation to attend to the cue. In a review of published OCT studies, Mulcahy and Hedge (2012) found that 21 out of 22 dog studies used a peripheral version of the task, in contrast with 6 out of 20 studies with apes. Clark et al. (2019), in a review of 71 dog and nonhuman primate OCT studies, similarly found a bias towards greater inter-object distances when testing dogs, with a median of 58cm for nonhuman primates and 155cm for dogs. Mulcahy and Hedge (2012) reported a 67% pass rate in the peripheral experiments with nonhuman primates, compared with 17% for the central versions, congruent with the increased success rates associated with peripheral versions found by Mulcahy and Call

(2009) and Mulcahy and Suddendorf (2011). Clark et al. (2019) also found increased inter-object distances were associated with greater success rates on a number of pointing cues. We are aware of only one study to date which has directly compared the performance of dogs on a central vs. peripheral version of the OCT. Kraus, van Waveren and Huebner (2014) found dogs' performance was impaired by 15% when tested with the central version, comparable to the 17% difference that Mulcahy and Call (2009) reported in their ape subjects, which Kraus et al. (2014) attributed to the distraction hypothesis.

In order to investigate further the potential effects of this difference in configuration on behaviour, to see whether these configurational effects would influence representatives of our own species, we tested 36-month-old children on a central and a peripheral version of the OCT, using a within-subjects manipulation of configuration. Human children reliably follow pointing cues from around 14 months of age (Behne, Lizkowski, Carpenter & Tomasello, 2012), so this age group were chosen in order to ensure that any behavioural differences observed were as a result of the experimental manipulations rather than a lack of emergence of pointing comprehension. In an attempt to replicate as closely as possible Mulcahy and Call's (2009) study, half of the children were tested within a child's playpen, designed to mimic the testing cages from within which nonhuman primate subjects are tested, which impose a barrier between subject and the testing apparatus and experimenter. The other half of the sample were tested without this barrier to ascertain whether the barrier itself had an effect on behaviour, such that a between-subjects design was used in testing the effect of a barrier.

Method

Participants

Participants were thirty-seven 36-month-old children ($M = 3$ yrs 5 days, $range = 32$ mos, 6 days – 40 mos, 0 days) comprised of 13 males and 24 females. Children were recruited from advertisements on local parenting social media sites. Parents gave informed consent for their children to participate. Data were collected between November 2018 and April 2019.

Procedure

On arrival at the community centre, the experimenter interacted with the child and showed them a box of stickers from which, they were informed, they could choose one they would like to keep. The experimenter then demonstrated a warm-up finding game, where she hid two stickers under two cups, mixed them up, and then asked the child to guess where one of the stickers was. This was repeated three times, in order to familiarise the child with both the experimenter and the testing apparatus. The experimenter then informed the child that they were going to play a fun finding game and the child was asked to sit on some coloured playmats with their caregiver. For children tested with a barrier, these playmats were inside the playpen and children were informed that this was the ‘finder’s house’. In order to centre the child, the experimenter labelled one of the coloured playmats ‘the finder’s seat’ and the child was asked to sit on this particular mat. The experimenter then informed the child that she was going to hide one of the stickers and then give the child a clue to see if they could find it. She then used another playmat to bait one of the cups (and sham-bait the other), then made eye contact with the child and asked, “are you ready for your clue?”. The experimenter then presented a contralateral dynamic pointing cue, as in Mulcahy and Call (2009),

which involves the pointing gesture being carried out in front of the participant and remaining in place until a choice is made. The experimenter also alternated her gaze between the child and the container. In the central condition, the distance between the experimenter's finger and the container was approximately 40cm, in the peripheral condition, it was approximately 1m. The experimenter maintained this position until the child made a choice. If the child was unresponsive, the experimenter gave verbal encouragement such as "where do you think the sticker is?". If the child failed to make a choice after 2 minutes, the trial was terminated, and a new trial began. If the child chose the correct cup, they were given the sticker to keep. If they chose the incorrect cup, the experimenter showed them the correct location of the sticker saying, "Never mind, let's hide it again!" and a new trial began. Children received 4 trials in the central condition and 4 trials in the peripheral condition, with the order of first condition being counterbalanced across participants. The baited container was on the right or left an equal number of times in each condition, and the order was counterbalanced, such that the reward was never on the same side for more than two consecutive trials.

Materials and set-up

The configurations were kept as close as possible to those described in Mulcahy and Call's (2009) study. Figure 1 shows the experimental set-up in the central and peripheral conditions. The playpens used as barrier were two Dream Baby Royal Converta 3-in-1 Playpen Gates, measuring 380 x 4 x 74cm (Rosyth Business Centre, 16 Cromarty Campus, Rosyth, Fife, KY11 2WX). Children and their parents sat on a playmat made up of 16 interlocking JSG Accessories Outdoor/ Indoor Protective Flooring Mats (JSG Accessories, Unit 6 Hughes Business Centre, Wilverley Road, BH23 3RU), and one of these was also used as the occluder. The containers used to hide the reward were two white opaque plastic cups measuring 7.8 x 10cm. All testing

sessions were recorded on a Sony Handycam HDR-PJ410 video-camera (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan).

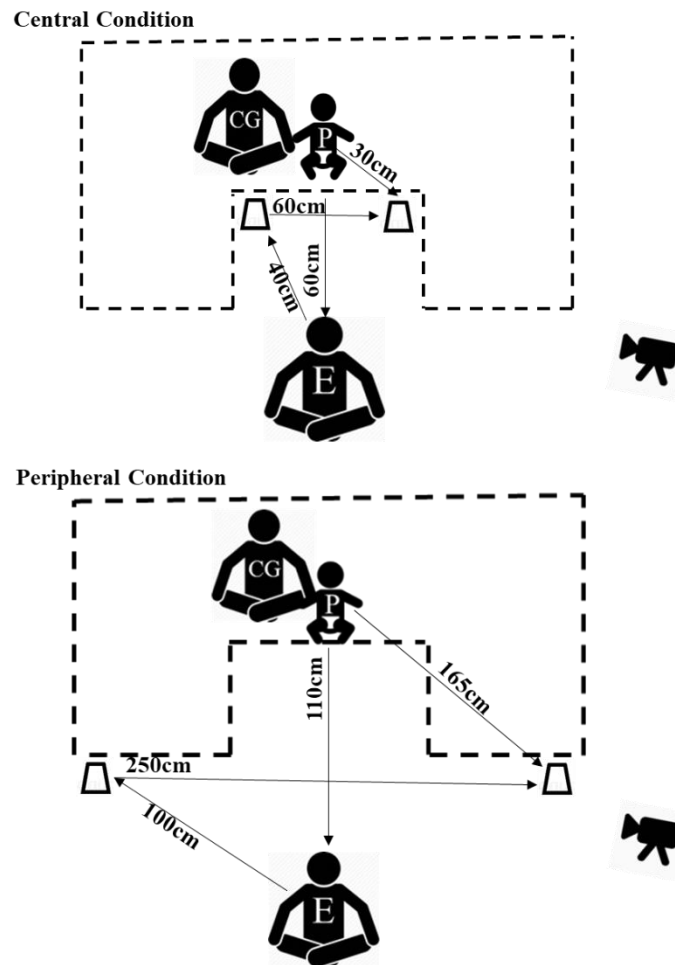


Figure 1: The experimental set-up in the central and peripheral conditions. P = participant, E = experimenter, CG = caregiver. Barrier represented by dashed line.

Drawing not to scale.

Data Scoring

Test sessions were video-recorded and coded at a later date. For each trial, data were coded for whether or not a correct choice was made, response latency and type of response. Response types were categorised according to following coding scheme:

Index-finger point: The arm and index-finger are extended towards the referent,

with the other fingers curled under the hand (Masataka, 2003).

Whole-hand point: An indicative gesture categorised by outstretched arm and extended fingers, which is not a direct attempt to obtain the container (Leavens & Hopkins, 1999).

Direct grab: A response was categorised as a grab when the child reached for and contacted the container with their hand.

Other: Responses other than those described above. An example of an *other* response is a child who pointed to the container using her foot.

No response: When a child failed to choose a cup within 2 minutes, this was coded as a *no response*.

Reliability

An independent coder who was blind to the purpose of the study coded 20% of the videos. There was excellent agreement between coders for correct choices, $\kappa = .84$, $p < .001$, response latency, $r_s = .84$, $p < .001$, and response type, $\kappa = .74$, $p < .001$.

Results

Due to a lack of normal distribution in the data, all analyses were performed using nonparametric tests.

Children Tested with a Barrier

Correct choices.

Children performed above chance in both the central (binomial calculator, $p < .001$) and peripheral (binomial calculator, $p < .001$) versions of the task. There was no difference in the proportion of trials in which children chose the correct container

between the central ($Mdn = 1.00$) and peripheral ($Mdn = 1.00$) versions, $Z = 0.00$, $p = 1.00$. This shows that children tested with a barrier performed at ceiling level in both conditions.

Response latency.

The children's latency of response was significantly lower in the central version ($Mdn = 2.75s$) than in the peripheral version ($Mdn = 4.75s$), $Z = -3.62$, $p < .001$. This shows that, when a barrier was present, the children were slower to respond when the containers were placed further away

Response type.

There was no significant difference in the proportion of trials in which children responded using index-finger points (central $Mdn = .25$; peripheral $Mdn = .50$, $Z = -0.85$, $p = .398$), whole-hand points (central $N = 0$; peripheral $Mdn = .00$, $Z = -1.00$, $p = .317$), grabs (central $Mdn = .75$; peripheral $Mdn = .25$, $Z = -1.29$, $p = .196$), or other responses (central $Mdn = .00$; peripheral $N = 0$, $Z = 1.00$, $p = .317$) between the two versions of the task. This shows that, for the children tested with a barrier, configuration had no effect on the types of responses elicited.

Children Tested without a Barrier

Correct choices.

The children performed above chance level in both the central (binomial test, $p < .001$) and peripheral (binomial test, $p < .001$) versions of the task. There was no significant difference in the proportion of trials in which children chose the correct container between the central version ($Mdn = 1.00$) and the peripheral version ($Mdn = 1.00$), $Z = -0.97$, $p = .334$. This shows that the children tested without a barrier also

performed at ceiling level in both conditions.

Response latency.

The children's response times were significantly lower in the central version ($Mdn = 2.25s$) than in the peripheral version ($Mdn = 4.00s$), $Z = -2.59$, $p = .010$. This shows that the children again were slower to respond when the containers were placed further away.

Response type.

On trials where children responded, only two response types were elicited, index-finger points and grabs. There was no significant difference in the proportion of trials in which children responded using an index-finger point (central $Mdn = .00$, peripheral $Mdn = .00$, $Z = -1.02$, $p = .306$), grab (central $Mdn = 1.00$, peripheral $Mdn = 1.00$, $Z = -1.14$, $p = .256$), or no response (central $N = 0$, peripheral $Mdn = .00$, $Z = -1.00$, $p = .317$) between the two versions of the task. This shows that, as with the children tested with a barrier, there was no effect of configuration on the types of response elicited.

Barrier vs. No Barrier Comparisons

There was no significant difference between children tested with a barrier and those tested without in the proportion of trials in which a correct choice was made in either the central (barrier $Mdn = 1.00$, no barrier $Mdn = 1.00$), Mann-Whitney $U = 144.00$, $p = .975$ or peripheral (barrier $Mdn = 1.00$, no barrier $Mdn = 1.00$), Mann-Whitney $U = 123.00$, $p = .318$, versions of the task. Similarly, there was no difference in response times between children tested with a barrier and children tested without a barrier in either the central (barrier $Mdn = 2.75s$, no barrier $Mdn = 2.25s$), Mann-Whitney $U = 143.00$, $p = .955$, or the peripheral (barrier $Mdn = 4.75s$, no barrier $Mdn =$

4.00s), Mann-Whitney $U = 113.5$, $p = .285$, versions of the task.

Figure 2 shows the mean proportion of trials in which the different response types were elicited in a) the central and b) the peripheral condition by children tested with a barrier and without a barrier. There were no significant differences in the central version of the task between children tested with a barrier and children tested without a barrier in the proportion of trials in the types of responses elicited (all $ps > .05$). In the peripheral version of the task, children tested with a barrier grabbed the container on a significantly lower proportion of trials ($Mdn = .25$) than children tested without a barrier ($Mdn = 1.00$), Mann-Whitney $U = 90.5$, $p = .048$. There were no other significant differences. This shows that in the peripheral version of the task, the presence of a barrier suppressed the tendency to grab the container but had no other effects on the types of responses elicited.

Order of Administration and Trial by Trial Analyses

There were no significant effects of order of administration on the proportion of correct choices, response latency or type of response in either of the conditions (all $ps > .05$). There were no differences across consecutive trials except with regard to response latency. For children tested without a barrier, there was a significant difference across trials in response latency in the central version, Friedman $\chi^2 = 10.27$, $p = .017$. Pairwise comparisons revealed that children were significantly slower to respond on the first trial than the second ($Z = -2.24$, $p = .025$), third ($Z = -2.27$, $p = .023$) and fourth ($Z = -2.42$, $p = .015$) trials. There were no other significant differences. There was also a significant difference across trials in response latency in the peripheral version, Friedman $\chi^2 = 15.17$, $p = .002$. Pairwise comparisons revealed that children were significantly slower to respond on the first trial than the second ($Z = 2.20$, $p = .028$) and the third ($Z = -3.09$,

$p = .002$) trials. There were no other significant differences. This shows that in both conditions, the children's speed of response increased with successive administrations of the test.

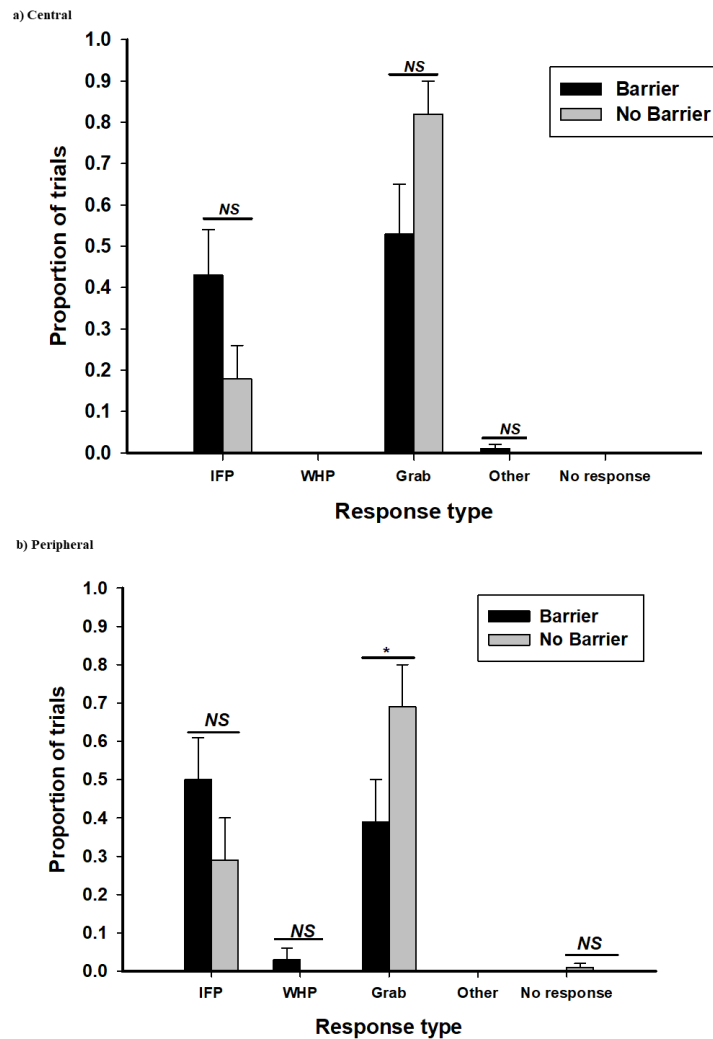


Figure 2: The mean proportion of response types, with standard errors, in a) the central and b) the peripheral condition. Means and standard errors are depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. IFP = index finger point, WHP = whole-hand point. * $p < .05$.

Discussion

In the current study, we investigated children's behaviour on a central and a peripheral version of the OCT. There was no effect of configuration on children's behavioural responses in terms of accuracy or the types of responses elicited. This contrasts with the findings of Mulcahy and Call (2009), Mulcahy and Hedge (2011) and Mulcahy and Suddendorf (2011) with ape subjects, and Kraus et al. (2014) with dog subjects who all found detrimental effects on performance associated with the central version. One explanation is for this is that by the age of 36 months, children are beginning to develop self-regulation skills that better enable them to comply with requests and control impulsivity (Chang, Shaw & Cheung, 2015; Kopp, 1982). This may make them less susceptible to distraction from attending to the cue, as appears to be the case with nonhuman primate and dogs. They may also be more motivated to participate in the 'game,' feeling pride at having won the reward through their own achievement and thus be more attentive to cues that will aid them with this. A simpler explanation may be that children are more practiced in following pointing cues than dogs or apes owing to the frequency of joint attentional interactions with caregivers in the first years of life in which children may learn that following a caregiver's focus of attention often leads to events or objects of interest (see Triesch, Teuscher, Deak & Carlson, 2006, for a discussion of how this may explain the emergence of gaze-following).

The difference in response latency we found - that is, greater response times associated with the peripheral version - can be explained through the increased distance between the child and the containers. Given that children tested both with and without a barrier grabbed the cup on 42% and 72% of trials, respectively, the increased time taken for children on these trials to locomote to the container in order to do so would result in

an increased mean latency in the peripheral condition.

We also found no effects of the presence of a barrier on either accuracy or latency of response. We tested half of the children with a barrier in order to more closely match the testing conditions experienced by the nonhuman primate subjects in Mulcahy and Call's (2009) study, and also those experienced by almost 100% of nonhuman primate subjects in OCT studies in general (Clark et al., 2019). We did, however, find a difference in the types of response elicited from the children as a function of the presence of a barrier, in that the barrier had a suppressing effect on the children's tendency to grab the container in acts of direct prehension in the peripheral version, and, although not significant, there was also a similar trend in the central version. This echoes the findings of Clark, Flack and Leavens (under review), who tested 18-month-old and 36-month-old children on an OCT with and without a barrier and found the presence of a barrier was associated with an increase in communicative responses (i.e. pointing gestures) and a suppression of grabbing in both age groups. The observation of these differences in the behavioural responses of young children who, at the age of 36 months, are experts in effectively utilising human gestural cues (Behne et al., 2012), and resulting from the imposition of a permeable barrier, further highlights the necessity of ensuring matched experimental conditions when comparing across species. 3-year-old children who have pre-experimental histories rich in human interaction and intense exposure to human gestural cues demonstrate differential responses when a barrier is present in the testing environment. It follows then, as Leavens et al. (2017) argued, that we cannot assume that failure to match for the presence of this barrier when comparing across representatives of different taxonomic groups has no confounding effect on results.

In conclusion, here, we demonstrate that differences in the spatial configurations

of the OCT have no effect on 3-year-old children's responses to the task, in contrast to previous findings with nonhuman primates and dogs. It would, however, be interesting to investigate whether such differences would occur in younger children with less developed skills in behavioural control. We do show that the presence of a barrier in the testing environment has an effect on the behavioural responses elicited from children. The children exhibited ceiling-level performance showing that this particular confound did not affect their ability to use the cue effectively, and we predict that this may also be the case with representatives of other species who have similar pre-experimental exposure to human directional cues. However, the children did show differential response types as a function of the presence of a barrier, showing that this confound, even in experienced point followers can affect behaviour. We therefore provide further support for Clark et al. (2019) and Leavens et al.'s (2017) arguments for the necessity of matching experimental conditions before appealing to phylogenetic theories to explain cross-species differences in performance; imposition of a barrier, simulating cage mesh, does have systematic effects on responses to the OCT in human 3-year-olds.

Acknowledgements

We thank the School of Psychology and the Doctoral School, University of Sussex for funding this research. Dan Hyndman and Martha Casey provided technical support. We thank Milton Village Hall for being so accommodating and special thanks to all the parents and children who took part. Finally, we thank Zoe Flack for inter-rater reliability coding.

Paper 4

Children's Responses to Different Reward Types on an Object Choice Task

Hannah Clark and David A. Leavens

This paper is written in the style of an article appropriate for *Animal Behaviour and Cognition*.

Author Note

HC and DAL designed the experiment, HC collected the data, conducted the analyses and wrote the manuscript. DAL provided advice on the manuscript and comments on drafts.

Abstract

Comparative studies investigating the differential socio-cognitive skills possessed by humans and our nearest relatives, the great apes, have been subject to criticism as a result of the prevailing tendency to compare across species in experimental tests that fail to match pre-experimental history, sample selection methods, test administration procedures and spatial configurations of testing paradigms. Here, we describe a number of studies that have directly compared the performance of human children and apes on tasks designed to measure a number of cognitive skills, all of which involved a reward being given when a correct response was elicited. We note that the studies differed in their treatment of the representatives of different species, in that nonhuman primates were rewarded with a piece of food whereas children were rewarded with toys or stickers (most commonly the latter). In order to investigate whether there are confounding effects of presenting different reward types, we tested 38 36-month-old children on an object choice task in which they received either a piece of food or a sticker on trials where they correctly followed a pointing cue to find the hidden reward. We found no effects of reward type on either correct choices, response latency or the type of response elicited suggesting that failure to match species on this particular aspect of testing may not have the same confounding effects as the other aforementioned factors have been demonstrated to have.

In searching for the traits that set humans apart from our nearest relatives, the great apes, many studies compare the performance of preschool children and representatives of the nonhuman hominid species. Children of this age are selected due to their lack of exposure to formal education, leading researchers to believe that this allows an insight into phylogenetically rooted cognitive skills, which, if evidenced in our species but not other nonhuman primates, may represent abilities which emerged after our separation from our common ancestors in the hominid lineage. Such comparative studies have been conducted across a number of domains and have led some researchers to argue that the fundamental difference between humans and the other ape species is not one related to increased general intelligence, but rather specialised socio-cognitive skills evolved to facilitate living in cooperative cultural groups, such as shared intentionality (Tomasello & Carpenter, 2007), and theory of mind (Povinelli & Eddy, 1996a). For example, Call and Tomasello (1999) compared the performance of 4- and 5-year-old children with that of chimpanzees and orangutans on a false-belief task and concluded that the latter do not possess a theory-of-mind based on their poor performance on the task. Povinelli, Bierschwale, and Čech (1999) compared 3-year-old children and chimpanzees on a gaze-following task in which the experimenter's eye orientation indicated their being focused on or distracted from a target and concluded that the children showed a high level understanding of attention as an unobservable mental state whereas the apes did not. Liszkowski, Schäfer, Carpenter and Tomasello (2009) claimed that the possession of cognitive skills for displaced reference is a human-unique skill based on their comparison of 12-month-olds and chimpanzees on a task designed to measure the ability to communicate about absent entities (but see Lyn et al., 2014, for discussion of the validity of these findings).

One area in which there is substantial debate about the relative abilities of humans and great apes is that of the production and comprehension of deictic cues, specifically declarative points. The function of declarative points is to direct another's attention to an item of interest or share information with another, differentiated from imperative points whose aim is to manipulate another to perform a certain behaviour, for example, give the pointer the indicated, desired object (Bates, Camioni & Volterra, 1975). Although it is now widely accepted that apes, in certain ecological conditions produce points (see Krause, Udell, Leavens, & Skopos, 2018; Leavens & Bard, 2011), different interpretations exist regarding the cognitive processes underlying pointing production and comprehension in apes and humans. For example, Van der Goot, Tomasello, and Liszkowski (2014) compared 12-month-olds with chimpanzees and bonobos on a task designed to elicit imperative points, and argued that the behavioural differences they observed could be explained through the apes' pointing resulting from ontogenetic ritualization of reaching, in contrast to the infants who demonstrated awareness of the communicative nature of the gesture (but see Leavens et al., 2015, for a critique).

Some authors argue that apes fail to comprehend human pointing cues and a wealth of research (e.g. Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012) demonstrates seemingly poor performance by apes in comprehending such cues, leading to a variety of arguments as to the complexity of the socio-cognitive skills required to successfully pass the task (for a lean interpretation see Lyn, Russell & Hopkins, 2011; Russell, Lyn, Schaeffer & Hopkins, 2011; for rich interpretations see Behne, Carpenter & Tomasello, 2005; Tomasello & Carpenter, 2007; Tomasello & Moll, 2010, and for discussions of the validity of attributing such abilities based on these studies see Leavens, 2018, and

Leavens, Bard & Hopkins, 2017). One experimental paradigm frequently employed to measure this ability is the Object Choice Task (OCT), which involves an experimenter hiding a reward in one of two or three opaque containers out of view of the subject and then using a deictic cue to indicate to the subject the container in which the reward is hidden. Povinelli, Reaux, Bierschwale, Allain, and Simon (1997) compared chimpanzees' performance on an OCT with that of 2-year-old children and concluded that the children's superior performance in using distal cues to find rewards indicated their understanding of the referential nature of these gestures, whereas the apes' failure to generalise their use of proximal points to more distal pointing cues indicated that they failed to understand the cues as referential. Similarly, Tomasello, Call, and Gluckman (1997) attributed the successful use of pointing, marker, and replica cues to find a hidden reward to two- and three-year old children's understanding of others' communicative intentions and the failure of chimpanzees and orangutans to do so as evidence that they lack this socio-cognitive ability. In a large-scale comparison of apes and children, Herrmann et al. (2007) found 2-year-old children performed better than apes on a battery of tasks designed to measure socio-cognitive skills, including the OCT, which they argued evinced support for the cultural intelligence hypothesis.

Recently, the validity of such studies has been called into question by a series of reviews which suggest that much of the comparative literature fails to control for confounds in sampling, methods and procedures when comparing across species (e.g. Byrmit, 2015; Clark, Elsherif & Leavens, 2019; Lyn, 2010; Mulcahy & Hedge, 2012). Leavens et al. (2017) noted that studies fail to match human and nonhuman subjects in terms of pre-experimental history (see Call & Tomasello, 1994; Lyn et al., 2010; for discussions of the effects that pre-experimental exposure to humans can have on performance) and age (see Bard, Bakeman, Boysen & Leavens, 2014 for a discussion

of, and evidence for, developmental similarities in apes and humans), use selective sampling when recruiting from one species and not the other, and fail to administer the same test procedure to the representatives of different species. Here, we address a further procedural difference, specifically the use of different reward types across species. The experimental paradigms used to test the cognitive skills listed above often involve, as previously mentioned with respect to the OCT, producing a certain behavioural response in anticipation of receiving a reward. In all of the abovementioned studies which directly compared the performance of nonhuman primates with human children (Call & Tomasello, 1999; Herrmann et al., 2007; Liszkowski et al., 2009; Povinelli & Eddy, 1996a; Povinelli et al., 1997; Povinelli et al., 1999; Tomasello et al., 1997; Van der Goot et al., 2014), the ape subjects received a food reward, whereas the human children were rewarded with a toy or sticker. In order then, to investigate whether behavioural responses may differ as a result of qualitatively different incentives being offered, we conducted an OCT with human children in which participants were offered a food reward in half of the trials and a sticker in the other half. Children were tested from within a child's playpen to imitate the conditions in which nonhuman primate testing takes place, that is, from within a cage, thus introducing a partial barrier between subject and experimenter and testing apparatus (Clark et al., 2019; Leavens et al., 2017). This was with the exception of six children who were tested without a barrier in order to ascertain that any effects found resulted from the reward type manipulation. The spatial configuration of the test set-up was designed to match that of Mulcahy and Call's (2009) "central" version of the OCT, representative of the design most commonly used with nonhuman primate subjects. The age of the children (36 months) was chosen because, from 14 months of age, children reliably follow pointing cues on the OCT (Behne, Liszkowski, Carpenter & Tomasello, 2014) and so we could ensure that any

differences in behavioural responses we observed were due to the experimental manipulation rather than the lack of emergence of this skill.

Method

Participants

Participants were 38 3-year-old children ($M = 36\text{m}, 4\text{d}$, $\text{range} = 32\text{m } 27\text{d} - 39\text{m } 7\text{d}$), comprising 17 males ($M = 36\text{m } 18\text{d}$, $\text{range} = 33\text{m } 5\text{d} - 39\text{m } 7\text{d}$) and 15 females ($M = 35\text{m } 27\text{d}$, $\text{range} = 32\text{m } 27\text{d} - 38\text{m } 26\text{d}$). Children were recruited from a participant database, where parents had registered their interest in participating with their children in cognitive studies, and from advertisements on social media sites. Data were collected between February and April 2017.

Materials

The playpen used was a Dream Baby Royal Converta 3-in-1 Playpen Gate, measuring 380 x 4 x 74cm (Rosyth Business Centre, 16 Cromarty Campus, Rosyth, Fife, KY11 2WX). Children and their parents sat on a playmat made up of 16 interlocking JSG Accessories Outdoor/ Indoor Protective Flooring Mats (JSG Accessories, Unit 6 Hughes Business Centre, Wilverley Road, BH23 3RU). The containers used to hide the reward were two white opaque plastic cups measuring 7.8 x 10cm. The occluder was a piece of brown cardboard measuring 65 x 80cm. All testing sessions were recorded on two Sony Handycam HDR-PJ410 video-cameras (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan).

Procedure

Children were familiarised with the laboratory setting and the experimenter during a “settling in” period, in which they played freely in the lab’s playroom. When

parents indicated that their child was settled, experimenter, parent and child moved to the test room, where the child was introduced to the test area, and was asked to sit on a child's playmat, accompanied by their parent. The children were asked to choose one sticker from a bowl, which they were told that they could keep, in order for the child to become familiar with the rewards that could be won, and for the experimenter to ascertain which of the stickers were most rewarding for the individual child. In order to centre the child, they were informed that they would only receive the "clue" to find the reward if they were sat on the "finder's seat" (two raised segments of the playmat positioned to be equidistant from the containers). The experimenter showed the child which reward she was going to hide (i.e. a piece of food or a sticker), baited one of the two containers behind a cardboard occluder and then placed the containers in position, such that the experimenter, child and containers were positioned as shown in Figure 1. The child was reminded to sit on the "finder's seat" if they were not already doing so, and the experimenter established eye contact with the child and told them they were going to see their "clue". The experimenter then turned her head towards the baited container, whilst simultaneously pointing, with index-finger, across the body (as in Mulcahy & Call, 2009). The experimenter maintained this position, alternating her gaze between the child and the baited container, until the child made a choice. If the child chose correctly, the experimenter retrieved the baited container, opened it and gave the child the reward, also giving verbal praise such as "well done" or "good finding!". If the child made an incorrect choice, the experimenter opened the chosen container in front of the child, made a comment such as "oh no, it's empty! Let's see if it's in the other one" and proceeded to open the baited container, retrieve the reward, and return it to the reward bowl. Each child received 8 trials in one testing session. The reward used was either a sticker or a small piece of food, provided by the children's parents, such that

each child received 4 trials with a sticker reward and 4 trials with a food reward. The baited container was on the right or the left an equal number of times, with the order counterbalanced such that the reward was never on the same side for more than two consecutive trials.

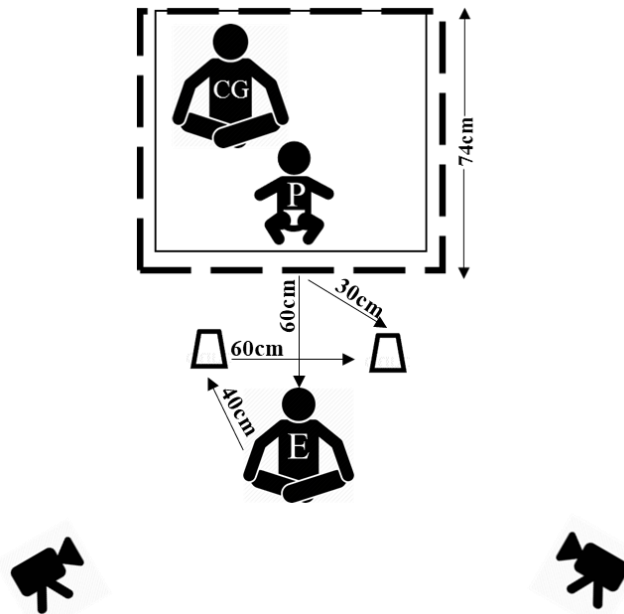


Figure 1: The experimental set-up. Dashed lines indicate playpen; CG = caregiver; P = participant; E= experimenter. Picture not to scale.

Data Scoring

Each test session was recorded on video-camera and later coded. For each trial, data were coded for whether or not the choice made was correct, latency of response (from maximum extension of the index-finger to the child choosing a cup) and type of response, with the following six response types identified (as in Clark, Flack & Leavens, under review):

Index-finger point: The arm and index-finger are extended towards the referent, with the other fingers curled under the hand (Masataka, 2003).

Whole-hand point: An indicative gesture categorised by outstretched arm and extended fingers, which is not a direct attempt to obtain the container (Leavens & Hopkins, 1999).

Indicative gesture other than index-finger/ whole-hand point: Where a child indicated a choice using a gesture other than an index-finger point or whole-hand point. An example of this is one child “pointed” to the container with their foot.

Direct Grab: A response was categorised as a grab when the child reached for and contacted the container with their hand.

Reach: An attempt to obtain the container, categorised by hand outstretched and fingers grasping (Leavens & Hopkins, 1999).

Other: Responses other than those described above. An example of an *other* response is a child who used the parent’s arm to indicate the choice.

Inter-Rater Reliability

An independent coder who was blind to the purpose of the study coded 20% of the videos. For correct choices, there was complete agreement between the two coders, $\kappa = 1.00, p < .001$. There was also excellent agreement between coders for latency, $r_s = .76, p < .001$. Response type agreement was initially poor, $\kappa = .28, p = .001$, however inspection of the data showed that this was due to the coders differentially coding the response of one child, who, in all trials tapped the container with her hand, which she then withdrew. The first coder categorised this as an “other indicative gesture”, whereas the second coder categorised this as a “grab”. This difference was due to a lack of clarity in the coding scheme which resulted in an easily identifiable coding error in response to one participant’s idiosyncratic response. Following Leavens, Hopkins and Bard (1996), in which a similar instance of two coders applied the coding system in

systematic and different ways, the second coder's categorisations of this behaviour were therefore recoded as "other indicative gesture" and the analyses rerun, with the final statistic demonstrating good agreement, $\kappa = .708$, $p < .001$.

Results

Due to a lack of normal distribution in the data, all analyses were performed using nonparametric tests.

Children Tested with a Barrier

Correct choices.

As a group, children performed above chance in both the food reward trials, binomial test, $p < .001$, and the sticker reward trials, binomial test, $p < .001$. There was no significant difference in the proportion of trials where children chose the correct cup between food reward trials ($Mdn = 1.00$) and sticker reward trials ($Mdn = 1.00$), $Z = -0.35$, $p = .725$. This shows that the children performed at ceiling level, as expected, both when the hidden reward was a piece of food and when it was a sticker.

Response Latency.

There was no significant difference in the children's mean response latency between the food reward trials ($Mdn = 2.25s$) and the sticker reward trials ($Mdn = 2.25s$), $Z = -0.81$, $p = .416$. This shows the children were equally quick to respond when both piece of food and a sticker were offered as a reward.

Response Type.

There were no significant differences between the food reward trials and the sticker reward trials in the proportion of responses that were index-finger points (food $Mdn = .75$; sticker $Mdn = .75$, $Z = -1.04$, $p = .299$), whole hand points (food $Mdn = .00$;

sticker $Mdn = .00$, $Z = -1.30$, $p = .193$), other indicative gestures (food $Mdn = .00$; sticker $Mdn = .00$, $Z = -0.82$, $p = .414$), grabs (food $Mdn = .00$; sticker $Mdn = .00$, $Z = -0.88$, $p = .380$), reaches (food $Mdn = .00$; sticker $Mdn = .00$, $Z = -1.73$, $p = .083$) or other responses (food $Mdn = .00$; sticker $Mdn = .00$, $Z = -1.41$, $p = .157$). This shows that the type of reward offered had no effect on the type of response the children used to indicate their choice of container.

Children Tested without a Barrier

Correct Choices.

As a group, children tested without a barrier performed above chance both in the food reward trials, binomial test, $p < .001$, and in the sticker reward trials, binomial test, $p < .001$. There was no significant difference in the proportion of correct choices between the food reward trials ($Mdn = 1.00$) and the sticker reward trials ($Mdn = 1.00$), $Z = -0.45$, $p = .655$. This shows that children tested without a barrier also performed at ceiling level both when tested with a food reward and a sticker reward.

Response Latency.

There was no significant difference in the mean response latency between the food reward trials ($Mdn = 2.63s$) and sticker reward trials ($Mdn = 2.25s$), $Z = -0.53$, $p = .599$. This shows that the children tested without a barrier were equally quick to respond with both types of reward.

Response Type.

Only three response types were elicited from the children tested without a barrier: index-finger points, grabs and responses categorised as “other”. There was no significant difference between the food reward and sticker reward trials in the

proportion of responses that were index-finger points (food $Mdn = .63$; sticker $Mdn = .63$, $Z = 0.00$, $p = 1.00$), grabs (food $Mdn = .13$; sticker $Mdn = .13$, $Z = 0.00$, $p = 1.00$) or categorised as “other”(food $Mdn = .00$; sticker $Mdn = .00$, $Z = .00$; $p = 1.00$). This shows that there was no effect of the type of reward on the types of responses elicited from the children tested without a barrier.

Barrier vs. No Barrier Comparisons

No effects of reward type were found for any of the measures for either children tested with a barrier or children tested without therefore results for the food and sticker reward trials were combined and analyses were conducted to assess whether the barrier itself had any effect on the children’s behavioural responses. There were no differences between the children tested with or without a barrier in the proportion of trials in which children chose the correct container, Mann-Whitney $U = 89.00$, $p = .717$, response latency, Mann-Whitney $U = 78.50$, $p = .483$, or response type (all $ps > .05$). Figure 2 shows the distribution of response types by children tested with and without a barrier.

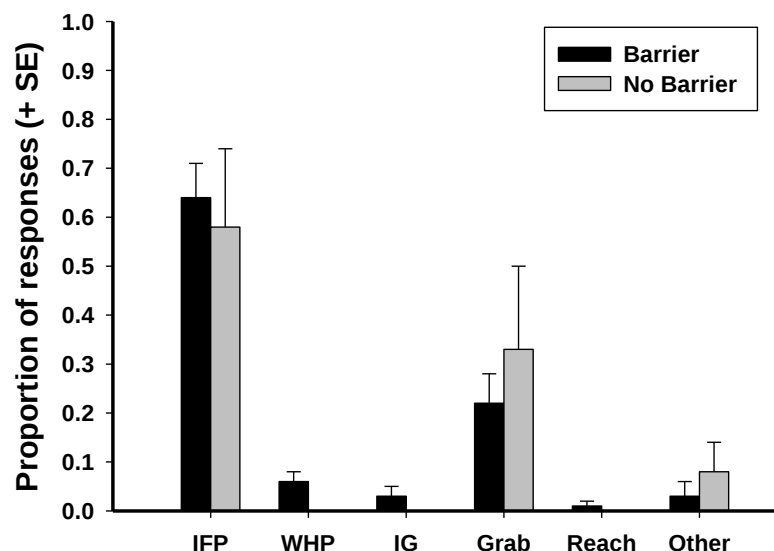


Figure 2: The mean proportion of response types, with standard errors, by children tested with a barrier and children tested without. Means and standard errors are

depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. IFP = index finger point; WHP = whole hand point; IG = other indicative gesture.

Discussion

In the current study, we investigated whether different types of reward would affect children's behaviour on an OCT. We found no differences in the behavioural responses elicited from children in terms of the accuracy, latency or response type on an OCT according to the type of reward that was hidden, specifically a food reward or a sticker. A number of recent papers have highlighted incommensurate testing protocols in comparative studies (Leavens et al., 2015; Leavens et al., 2017), such as pre-experimental exposure to humans (Lyn, 2010), failure to administer the same test procedure (Leavens et al., 2015) and differences in the spatial configurations of test set-ups (Mulcahy & Hedge, 2012) and studies investigating their effects have demonstrated these inconsistencies to represent confounds which can have detrimental effects on individuals' performance (see Lyn et al., 2010, for the effects of pre-experimental history; and see Mulcahy & Call, 2009 and Mulcahy & Suddendorf, 2011, for effects of spatial configuration). Our results suggest that failing to match reward types may not impact the validity of the results in the same way as these other factors. One possible explanation for this, however, could be the way in which we presented the two different types of reward, specifically, our use of no more than two consecutive trials of the same reward in each eight-trial test. A possibility for a future direction would be to present eight consecutive trials of the same reward type in a between subjects design in which half of the participants receive food rewards and the other half sticker rewards, in order to investigate whether children's motivation to work for the reward continues for the full eight trials. Finding, for example, that children continue to participate in exchange

for the stickers for the full eight trials, yet reach satiation with food rewards after fewer trials, would have implications for testing with nonhuman animals who are often tested with food rewards for large amounts of consecutive trials.

We also found that there were no differences in the behavioural responses elicited from children as a function of the presence of a barrier in the testing environment. Thirty-two of the 38 children sampled here were tested with this barrier present, in an attempt to match the testing environments experienced by nonhuman primates based on Clark et al.'s (2019) findings that almost 100% of nonhuman primates tested with the OCT are tested from within a cage. A further 6 children were tested without the barrier, here, in order to ensure that any differences we found were due to the reward type manipulation rather than due to the presence of the barrier. A greater range of response types were elicited from the children tested with a barrier than those tested without, which may be an indication of the barrier having an effect on the children's behavioural responses or may be due to the much greater number of children tested in this condition leading to greater variation in response. That we found no significant effect of the barrier on children's responses contrasts with Clark, Flack and Leavens' (under review) findings that the presence of a barrier elicited more communicative responses (index-finger and whole hand points) from 18- and 36-month-old children whilst the absence of a barrier increased the frequency of grabbing the container in acts of direct prehension. This effect was particularly pronounced in the younger children in their sample and the authors explained this in terms of the referential problem space (Leavens, Hopkins & Bard, 2005), that is, the children perceived the bars of the playpen as a barrier between themselves and the reward, despite the reward actually being within reach, and thus chose to use a communicative gesture to influence the experimenter to obtain the desired object on their behalf. This

finding, that the bars could be perceived as a barrier to the reward, has important implications for comparative studies that fail to match this aspect of the testing environment across species. A possible explanation for the absence of an effect of a barrier in the current study is the small proportion of our sample who were tested with a barrier, meaning that analyses of barrier/ no barrier differences are low in statistical power, and further tests with larger sample sizes are needed to effectively test the reproducibility of Clark et al.'s findings.

In conclusion, although recent reviews highlight the necessity of ensuring that testing conditions are matched across species with reference to a number of procedural and methodological aspects, our results suggest that failure to match reward types may not be one such of these confounds.

Acknowledgements

We thank the School of Psychology and the Doctoral School at the University of Sussex for funding this research. Thanks to Dan Hyndeman and Martha Casey for technical support, to Dr. Jessica Horst and Professor Nicola Yuill for providing testing space, and to Zoe Flack for inter-rater reliability coding. Special thanks to the parents and children who took part.

Paper 5

Testing Dogs in Ape-Like Conditions: The Effect of a Barrier on Dogs' performance on
the Object Choice Task

Hannah Clark and David A. Leavens

University of Sussex

A version of this paper is published in *Animal Cognition* as:

Clark, H. & Leavens, D. A. (2019) Testing dogs in ape-like conditions: The effect of a
barrier on dogs' performance on the Object Choice Task.

Author Note

HC and DAL designed the study, HC collected the data, conducted the analyses and
wrote the manuscript. DAL provided comments on drafts of the manuscript.

Abstract

Recent reviews have found marked procedural and methodological differences in the testing of different taxonomic groups on the Object Choice Task. One such difference is the imposition of a barrier in the testing environment of nonhuman primates in the form of a cage necessitated to ensure the experimenter's safety. Here we conducted two studies with domestic dogs (*Canis familiaris*) in which we compared the performance of dogs tested from within a child's playpen and dogs tested without this barrier present. In Study 1, in a within-subjects design, we found no effect of the barrier on dogs' ability to use a pointing cue, but there was an increase in instances in which dogs failed to choose a cup. In Study 2, in a between-subjects design, dogs tested with a barrier failed to perform above chance and were also more likely to fail to make a choice. When dogs tested without a barrier made an incorrect response, these were more likely to be incorrect choices than no choice errors. We discuss the implications of these differences in behavioural responses in function of the presence of a barrier and the necessity of ensuring matched conditions when comparing across species.

The differential performances of domestic dogs and nonhuman primates on the Object Choice Task (OCT), which measures an individual's ability to follow human gestural cues, have led to phylogenetic theories regarding their respective socio-cognitive abilities. Numerous studies (e.g., Riedel, Schumann, Kaminski, Call, & Tomasello, 2007; Viranyi et al., 2008) report that domestic dogs possess specialised skills in comprehending human communicative cues, evidenced by their high levels of performance on the Object Choice Task (OCT). Whether emerging as a by-product of domestication (Hare & Tomasello, 2005) or as a result of humans' active selection for dogs with specific traits (Miklosi et al., 2003), there is a consensus among some researchers that dogs have an evolved ability to follow human gestural cues. In contrast, apes' poor performances have been attributed to a separation in the primate lineage resulting in this being (among primates) a human-unique ability (e.g. Herrmann et al., 2007; Tomasello & Carpenter, 2007; Moll & Tomasello, 2011). Several studies directly compared the performance of nonhuman primates and dogs on the OCT, and concluded that apes, for example, are less similar to human infants than dogs in terms of their socio-cognitive abilities due to convergent evolution (Maclean, Herrmann, Suchindran, & Hare, 2017) and lack the ability to understand communicative intentions (Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012).

Leavens, Bard, and Hopkins (2017) discussed the prevalence of systematic confounds with species classification in the comparative cognition literature in studies which compare across species. Such confounds are by no means absent from OCT research, and, in fact, some have already been addressed. For example, Leavens et al. (2017) discussed the tendency to compare institutionalised apes with non-institutionalised human infants with no regard for the differences that experiential history with humans may confer. Lyn (2010) highlighted this factor of enculturation

with specific reference to the OCT, with several studies demonstrating that apes from backgrounds rich in human interaction are able to follow human communicative cues on the OCT (Call & Tomasello, 1994; Lyn, 2010; Lyn, Russell, & Hopkins, 2010; Russell, Lyn, Schaeffer & Hopkins, 2011). Similarly, dogs from backgrounds lacking in interactive exposure to humans perform much more poorly on the OCT (D’Aniello et al., 2017; Lazarowski & Dorman, 2015; Udell, Dorey, & Wynne, 2010), leading Udell, Dorey, and Wynne (2010b) and Wynne, Udell, and Lord (2008) among others, to emphasise the role of ontogeny in the development of socio-cognitive skills in dogs rather than their being innate (e.g. Kaminski & Nitzchner, 2013).

Leavens et al. (2017) also noted the differences in testing environments when comparing human infants and apes, in that apes are tested from within cages for safety measures, whereas the infants are not. Such a criticism can be extended to research comparing apes with several other species, notably domestic dogs. In fact, Clark, Elsherif, and Leavens (2019), in a review of 71 published nonhuman primate and dog OCT studies, found that fully 99% of the nonhuman primates tested were tested with a barrier in the testing environment, compared with less than 1% of domestic dogs. These 16 dogs were all subjects in Kirchhofer et al.’s (2012) comparison of nonhuman primate and dog performance, and the portion of their dog sample which were tested with a barrier performed significantly worse than those tested without. Clark, Flack, and Leavens (under review) compared the performance of 18-month- and 36-month-old human children tested with and without a barrier on the OCT. These age groups were chosen because from 14-months of age, human children reliably follow pointing cues (Behne, Lizkowski, Carpenter, & Tomasello, 2012) and so any behavioural differences could reliably be attributed to the manipulation of the presence of a barrier rather than lack of emergence of these skills. As predicted, all of the children performed at ceiling

level, however, there were marked differences in the behavioural responses in both age groups, with increased communicative responses compared with acts of direct prehension when a barrier was present. That the children frequently chose to communicate their choice rather than directly lifting the cup as they chose to do in the “no barrier” condition, suggests they perceived the barrier as an impediment to their ability to obtain the reward themselves. Thus, as Leavens et al. (2017) argued, it cannot be assumed that such differences in testing environment have no effect on performance.

In the current studies, we aimed to investigate whether testing pet dogs in “ape-like” conditions would have an effect on their behavioural responses. We compared the performance of dogs tested with and without a barrier on an OCT using an *ipsilateral proximal dynamic* pointing cue in Study 1, and a *contralateral proximal dynamic* pointing cue in Study 2. Reviewing the OCT literature, Clark et al. (2019) found that nonhuman primates tend to be tested with *contralateral* rather than *ipsilateral*, *dynamic* rather than *momentary* and *proximal* rather than *distal*, pointing cues. Dogs, in contrast, tend to be tested with *ipsilateral*, *distal* cues, with a nearly equal proportion being *momentary* and *dynamic*. The experimental configuration in both studies was such that it would be categorised as a “central” version of the task, according to Mulcahy and Hedge’s (2012) distinctions. Mulcahy and Hedge (2012) and Clark et al. (2019) found that apes are more frequently tested with this version of the OCT, in which the placement of the containers is such that they are close together, and within the direct line of vision of the subject. This is compared to the “peripheral” version, in which the containers are separated by a greater distance and not within the subjects’ direct line of vision. They found that this latter version is more frequently used when testing dogs, and this, they argue, may affect performance, because having the containers within the direct line of vision may distract the subjects’ attention away from the cue owing to the

salience of the food reward held within. Hence, in the present study, we also explore the effects of administration of task features more typical of presentations used with nonhuman primates.

Study 1: Barrier vs. no barrier within-subjects

Method

Subjects

Thirty-two pet dogs (15 male, 17 female) took part in the study. Dogs ranged in age from 4 months to 13 years old ($M = 4.97$, $SD = 3.50$) and comprised a variety of breeds (see Table 1 for individual subject data). Dogs were recruited through advertisements on social media, word of mouth, and flyers distributed. Although some dogs had taken part in other cognitive tests before, none had previously been tested on an OCT. All subjects were tested individually and by an unfamiliar experimenter. Testing took place inside in a community hall, and dogs were randomly assigned to the first condition prior to testing. One dog was excluded from the final analyses because she failed to complete more than two trials in one condition due to being apparently nervous and unable to settle.

Procedure

On arrival at the hall where testing took place, subjects were given time to freely explore the test room, off-lead, in order to become familiar with the environment. Whilst owners read the information and completed the consent forms, the experimenter interacted with the dog and offered them a treat. When the owner stated that the dog was comfortable and ready to begin, they were asked to put the dog on a 1m-long lead and to stand in a marked position in the test room (for the *barrier* condition, this was within a child's playpen, for the *no barrier* condition, this was in the same place but without the playpen present). The experimenter then kneeled in position (see Figure 1

for experimental set-up) and placed the two cups upside down on the floor. She then proceeded to bait both cups with a piece of dry dog food, in sight of the subject, and prior to doing so she called the dog's name and said "Look", ensuring that the subject was watching while baiting took place. The experimenter then called the dog's name, made eye contact and, using an ipsilateral, dynamic point, indicated one of the cups. This cue was held until the subject made a choice. If the subject chose the correct cup, the experimenter gave the subject the piece of food (if they had not already retrieved it themselves), and if the subject chose the incorrect cup, both pieces of food were removed and placed back in the food container. If the subject did not make a choice, the trial continued for one minute before the experimenter stopped giving the cue and the next trial began. Following recommendations from Udell et al. (2010b) the beginning of the trial was counted from when the cue was presented and the subject released to make a choice. Owners were asked to hold the lead whilst baiting took place, and to drop the lead so that the dog was free to move independently as soon as the experimenter pointed. Subjects took part in blocked trials with 4 trials per condition, with this number of trials chosen in order to avoid detrimental effects on performance associated with loss of motivation or satiation to the food reward from repeated testing as trials in the two conditions took place consecutively. The order of conditions was counterbalanced prior to testing. The baited container was on the right or left an equal number of times and the order was counterbalanced, such that the container was never on the same side for more than two consecutive trials.

Materials and set-up

The playpen used in the barrier condition was a Dreambaby Royal Converta 3-in-1 Playpen Gate, measuring 380 x 4 x 74cm (Dreambaby, Unit 53, Rosyth Business Centre, 16 Cromarty Campus, Rosyth, KY11 2WX, Scotland). The containers used to

hide the bait were two opaque plastic cups. A premium commercial dry dog food was used for baiting the cups. All dogs were tested on a 1m long lead. All testing sessions were recorded on two Sony Handycam HDR-PJ410 video-cameras (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan). The owner stood at a distance of 184cm from the experimenter holding the subject on a 1m lead, such that the nearest distance between subject and experimenter (depending on the size/ position of the subject) was 60cm. The experimenter was positioned 60cm from the edge of the barrier in the barrier condition. The distance between the two containers was 60cm, and the distance between the experimenter's pointing finger and the container was approximately 10cm.



Figure 1: The experimental set-up in the no barrier and barrier conditions.

Data Scoring

All test sessions were video-recorded and coded at a later date. For each trial, data were recorded for whether a correct choice was made according to Udell et al.'s (2010b) recommendations for the standardisation of OCT tests. Thus, a “correct choice” was defined as the subject first touching or coming within 10cm of the correct container with their snout. Any other response, including the trial timing out, was marked as an “incorrect response”. Incorrect responses were further categorised into “incorrect choice”, where the subject first touched or went within 10cm of the incorrect container, and “no choice”, where the subject failed to come into contact with either of the containers before the end of the 1 min trial.

Reliability

All trials were coded by the first author, and 20% of the dogs' trials were coded by a second coder, who was blind to the hypotheses under test (six dogs, with eight trials each, for a total of 48 trials). Inter-observer reliability as to whether each dog was correct on each trial was high: Cohen's $\kappa = .76$.

Results

Correct Choices

Due to a lack of normal distribution in the data, all analyses used nonparametric tests. Dogs performed above chance when tested both without a barrier (binomial test, $p < .001$) and when tested with (binomial test, $p < .001$). There was no significant difference in percentage of correct choices between the barrier ($Mdn = 88\%$) and the no barrier ($Mdn = 100\%$) conditions, $Z = -0.72$, $p = .470$. This shows that the barrier did not have a suppressing effect on the dogs' ability to use an *ipsilateral proximal dynamic* pointing cue on the OCT.

Incorrect Choice vs. No Choice

There was no significant difference in the percentage of incorrect responses that were incorrect choices between the barrier and no barrier conditions, $Z = -1.41$, $p = .157$. There was a significant difference in the percentage of incorrect responses that were "no choice" responses, with dogs failing to choose one of the cups on significantly more trials in the barrier condition ($Mdn = 0\%$) than in the no barrier condition (no "no choice" responses were recorded in this condition), $Z = -2.24$, $p = .025$. This shows that, although there were no significant differences in performance between the two conditions, there were differences in the behavioural responses, with dogs failing to make a choice more when a barrier was present. Figure 2 shows the percentage of correct choices, incorrect choices and no choice responses in the two conditions.

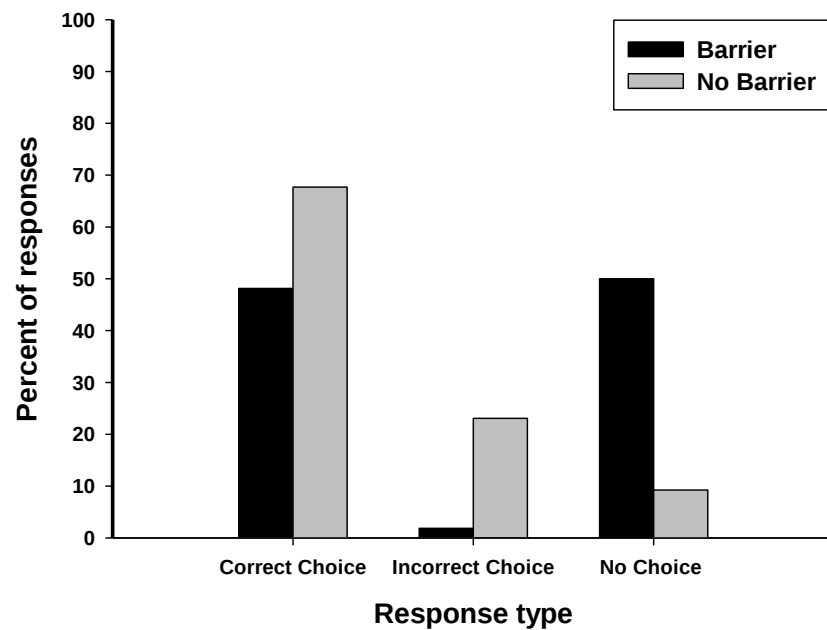


Figure 2: The percentage of trials in which dogs made a correct choice, an incorrect choice and no choice in Study 1. Total number of trials in no barrier condition = 124; total number of trials in barrier condition = 120. * denotes significant at $p < .05$.

Order of Administration

There was no significant effect of order of administration on the proportion of trials correct in the barrier condition (barrier first $Mdn = 1.00$; no barrier first $Mdn = .75$), Mann-Whitney $U = 109.00$, $p = .875$, or in the no barrier condition (barrier first $Mdn = 1.00$; no barrier first $Mdn = 1.00$), Mann-Whitney $U = 119.00$, $p = .965$. This shows that the dogs' performance was not affected by the order of administration of the barrier and no barrier conditions.

Trial by Trial Analyses

There was no difference in correct choices across trials in the barrier condition, Cochran's $Q = 2.33$, $p = .506$, or in the no barrier condition, Cochran's $Q = 2.61$, $p =$

.456. This shows that the dogs' performance was not affected by successive administrations of the task.

Discussion

Here, the imposition of a barrier in the testing environment did not affect dogs' ability to use an *ipsilateral dynamic proximal* pointing cue on an OCT. There was, however, a subtle difference in the dogs' behaviour, with significantly more of the incorrect responses comprising "no choice" responses, suggesting that the barrier had a suppressing effect on the frequency of a choice being made.

Study 2: Barrier vs. No barrier Between-subjects

Method

Subjects

Thirty-seven (15 male, 19 female) pet dogs took part in the study. Dogs ranged in age from 5 months to 11 years old ($M = 4.23$; $SD = 2.94$) and comprised a variety of breeds (see Table 2 for individual subject data). The dogs were recruited through advertising on social media, word-of-mouth, and flyers distributed. None of the dogs had previously taken part in an OCT study. All testing was completed inside in a community hall by an unfamiliar experimenter. Three dogs were excluded from the final analyses because they failed to complete at least two trials, due to being inattentive or unable to settle.

Procedure

The procedure was the same as in Study 1, except that the pointing cue used *contralateral proximal dynamic* point and half of the subjects were tested with a barrier and half without; thus, each dog received four trials in one of the two conditions, Barrier, or No Barrier.

Materials

The materials used were the same as in Study 1.

Data Scoring

The data were scored in the same way as in Study 1.

Reliability

All trials were coded by the first author, and 20% of the dogs' trials were coded by a second coder, who was blind to the hypotheses under test (seven dogs, with four trials each, for a total of 28 trials). Inter-observer reliability as to whether each dog was correct on each trial was high: Cohen's $\kappa = .73$.

Results

Correct Choices

Dogs tested without a barrier chose the correct container significantly above chance (binomial test, two-tailed, $p = .003$). Dogs tested with a barrier did not perform significantly above chance (binomial test, two-tailed, $p = .089$). There was no significance difference in the percentage of trials in which the dogs chose the correct container between subjects tested with a barrier ($Mdn = 50\%$) and those tested without a barrier ($Mdn = 75\%$), Mann-Whitney $U = 88.5$, $p = .215$. This shows that the barrier had a suppressing effect on the dogs' ability to use a *contralateral proximal dynamic* pointing cue.

Incorrect Choice vs. No Choice

Dogs tested with a barrier responded by making an incorrect choice on a significantly greater percentage of trials ($Mdn = 25\%$) than those tested without a barrier ($Mdn = 0\%$), Mann-Whitney $U = 48.00$, $p = .001$. Dogs tested with a barrier failed to make a choice on a significantly greater percentage of trials ($Mdn = 50\%$) than those tested without ($Mdn = 0\%$), Mann-Whitney $U = 60.00$, $p = .007$. These findings, as in

Study 1, show that, although the imposition of a barrier did not lead to a difference in performance, overall, it did elicit different behavioural responses, and increased the likelihood of dogs failing to choose one of the containers. Figure 3 shows the percentage of responses that were correct choices, incorrect choices and no choice made in the two groups.

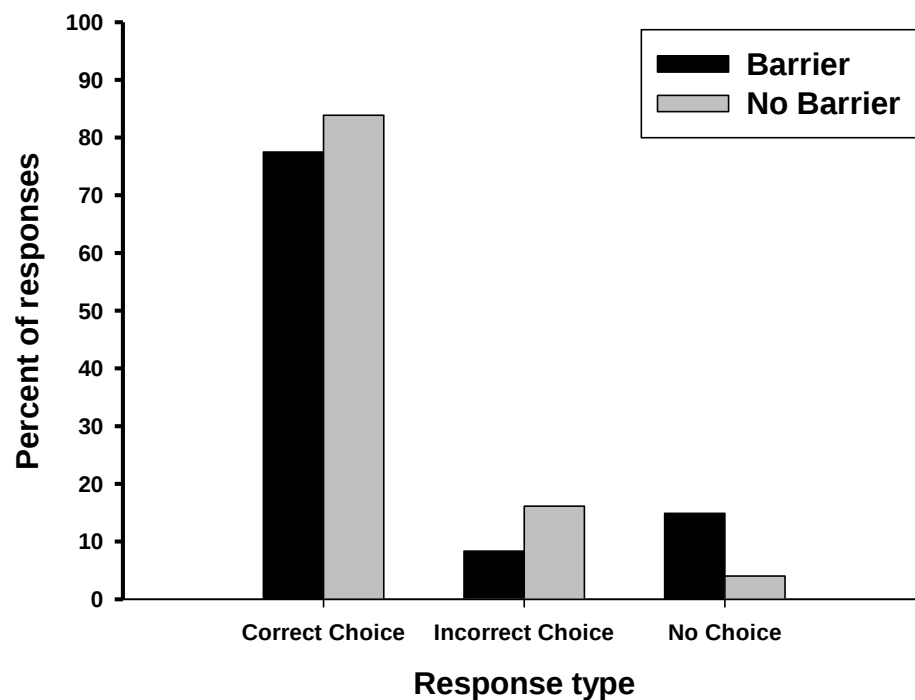


Figure 3: The percentage of trials in which dogs tested with and without a barrier made correct choices, incorrect choices and no choices in Study 2. Total number of trials in no barrier condition = 65; total number of trials in barrier condition = 54. * denotes significant at $p < .05$.

Trial by Trial Analyses

There was no significant difference in correct choices made across trials in dogs tested with a barrier, Cochran's $Q = 3.33$, $p = .343$, or dogs tested without a barrier, Cochran's $Q = 1.50$, $p = .682$. This shows that the dogs' tendency to choose the correct container was not affected by successive administrations.

Discussion

Here, the presence of a barrier had a suppressing effect on dogs' ability to use a *contralateral proximal dynamic* pointing cue on an OCT. This differs from our findings in Study 1, in which dogs performed above chance when tested both with and without a barrier,

As in Study 1, we found a difference in the behavioural responses elicited from the dogs between those tested with and without a barrier. Where incorrect responses were recorded, in the no barrier condition these tended to be due to subjects choosing the incorrect container, whereas in the barrier condition, they were due to the subjects failing to make a choice. This shows that the presence of a barrier does affect subjects' responses on the OCT and thus, comparing across groups when one is tested with and the other without a barrier, represents an experimental confound.

General Discussion

Here, we present two studies in which we examine the effect of the imposition of a barrier on dogs' behaviour on an OCT. In Study 1, we found no effect of the barrier on dogs' ability to use an *ipsilateral proximal dynamic* pointing cue on the OCT, whereas in Study 2, we did find that the dogs' performance was significantly worse on a *contralateral proximal dynamic* pointing cue. The latter finding supports the results of Kirchhofer et al. (2012), who also found a reduction in success rate in dogs tested with a barrier compared to those without. That we did not find this reduction in performance in the first study may be due to the different types of pointing cue used. Miklósi and Soproni (2005) and Udell et al. (2013) reported that the *ipsilateral proximal dynamic*

pointing cue is one of the simpler cues to follow, and even dogs with minimal prior exposure to humans can succeed at using this cue on the OCT (Udell, Dorey & Wynne, 2010). Clark et al. (2019) in their review, found dogs' performance was significantly lower with *contralateral* than *ipsilateral* pointing cues and *momentary distal* points, as used in Kirchhofer et al. (2012) are reportedly more difficult to follow (Miklósi & Soproni, 2006; Udell et al., 2013), and therefore it may be that the increased difficulty of the cue led to an increase in the suppressive effect of the barrier.

In both studies, we found different behavioural responses were elicited from the dogs according to whether the barrier was present, specifically in the frequency in which incorrect responses were constituted by incorrect choices or failure to choose a cup. In both studies, there were significantly more instances in which the dogs failed to make a choice. Clark, Flack and Leavens (under review), when testing human children with and without a barrier on the OCT, found an increase in communicative behaviour when the barrier was present, as opposed to acts of direct prehension to obtain the reward, a finding which noted could be explained with reference to Leavens, Hopkins and Bard's (2005) Referential Problem Space. Specifically, when the barrier was present, the children perceived the containers as out of reach (although they were not) and so chose to communicate with the experimenter to influence her behaviour to receive the reward. In the current study, the dogs, when the barrier was present, may have perceived the containers (and thus the reward) as unobtainable and therefore, lacking the gestural communicative skills of human children, failed to try to obtain the reward.

In Study 2, we also found that a significant difference in the dogs' behavioural responses in the two conditions related to incorrect responses. Here, although there was no significant difference in the number of correct choices, the dogs made incorrect

choices significantly more when the barrier was not present than when it was. This could be explained with reference to Mulcahy and Call's (2009) distraction hypothesis. That is, when the containers are within the subjects' direct line of vision, and they therefore have to look past the containers to attend to the cue, the salience of the container and the reward contained within may distract attention away from the cue being given. Kraus, van Waveren and Huebner (2014) compared dogs' performance on a central and a peripheral version of the OCT and found that, when tested with a *momentary proximal* pointing cue, the dogs' performance, although still above chance, was significantly lower in the central version, providing support for the distraction hypothesis. In Study 2, we matched the testing conditions in Mulcahy and Call (2009) as closely as possible, by also using the same cue type, and found a 67% success rate when no barrier was present and a 48% success rate when the barrier was present (and thus, the dogs were tested in the same conditions as the apes); the apes in their study had a success rate of 58% on the central version of the task. This suggests that the spatial configuration in the current study also affected the dogs' responses, supporting the distraction hypothesis of Mulcahy and Call (2009), and further showing that cross-species comparisons that do not control for these factors have reduced validity.

The results reported here provide further evidence of the necessity of matching testing conditions when comparing performance across different groups. That we found differences in the behavioural responses of the dogs according to whether or not a barrier was present in the testing environment, together with Clark, Flack and Leavens' (under review) similar findings with human children, provides support to Leavens et al.'s (2017) arguments that we cannot assume that performance is not affected by differences in the testing environment (also see Hopkins, Russell, McIntyre, & Leavens, 2013). This is further highlighted by the dogs' below-chance performance in the barrier

condition in Study 2. In sum, here we report an empirical test of predictions from recent reviews of the OCT, confirming that methodological and procedural differences in testing environments in this experimental paradigm may more parsimoniously explain apparent species differences in performances than evolution- or selection-based accounts. Thus, appeals to the effects of domestication on cognitive performance (e.g., Hare & Tomasello, 2005) are predicated, in part, on between-species comparisons that have confounded procedural factors, such as the presence of a barrier, with species classification. So-called species differences between apes and dogs in the OCT have not adequately controlled for these systematic confounds; these group differences may be simple artefacts of the radically different protocols administered to dogs, compared with the protocols administered to non-dog species (e.g., Hopkins et al., 2013).

Ethical Note

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the Animal Welfare and Ethical Review Board at the University of Sussex, United Kingdom.

Acknowledgements

We thank the School of Psychology and the Doctoral School, University of Sussex for funding these studies. Dan Hyndman and Martha Casey provided technical support. We also thank the many pets and owners who volunteered their time for this study. We thank Kassandra Giragosian for assistance with reliability coding. Finally, we thank the

Milton Village Community Association for generous and helpful access to their facilities.

Table 1

Study 1 individual subject and performance data

Name	Breed	Sex	Age (years)	First condition	Trials Complete Barrier condition	Correct trials Barrier condition	Trials complete No barrier condition	Correct trials No barrier condition
Roxy	Labrador Retriever x Poodle	F	4	No barrier	4	0	4	4
Luka	Parson Russell Terrier	F	4	Barrier	4	2	4	3
Trixie	Yorkshire Terrier	F	2	No barrier	4	3	4	1
Mali	Sussex Spaniel	F	2	Barrier	4	4	4	4
Jack	Yorkshire Terrier	M	5	No barrier	4	0	4	4
Charlie	Shih Tzu	M	4	Barrier	4	2	4	2
Stan	Jack Russell Terrier	M	6	No barrier	4	3	4	4
Freddie	Shih Tzu	M		Barrier	4	4	4	2
Topsy	Cocker Spaniel x Poodle	F	2	No barrier	4	0	4	3
Missy	Whippet	F	2	No barrier	4	4	4	2
Toby	Springer Spaniel	M	8	Barrier	4	3	4	3
Lionel	Lhasa Apso	M	4	No barrier	4	1	4	3
Ruby	Lhasa Apso	F	6	Barrier	4	4	4	4
Badger	Border Collie	M	3	Barrier	4	4	4	3
Spock	Jack Russell Terrier	M	2	No barrier	4	3	4	0
Muffins	Lurcher x Spaniel	F	7	Barrier	4	2	4	0
Mabel	Labrador Retriever x Pointer	F	2	No barrier	4	1	4	2
Amber	Mongrel	F	2	No barrier	4	4	4	4
Marley	Labrador Retriever x Pointer	M	13	No barrier	4	4	4	4

Table
continues...

Beth	Labrador Retriever x Pointer	F	3	No barrier	4	3	4	3
Solo	Papillon	M	4	No barrier	4	4	4	4
Ruby	Cocker Spaniel x Poodle	F	0	No barrier	4	4	4	4
Bailey	Staffordshire Bull Terrier	M	0	Barrier	4	4	4	4
Bella	Yorkshire Terrier	F	7	Barrier	0		1	0
Choco	Chihuahua x Podenco	F	4	Barrier	4	4	4	4
Bear	Staffordshire Bull Terrier	M	9	No barrier	4	4	4	3
Spike	English Bulldog x Staffordshire Bull Terrier	M	13	Barrier	4	4	4	4
Eric	Staffordshire Bull Terrier x	M	10	Barrier	4	4	4	4
Inca	Labrador Retriever	F	3	No barrier	4	4	4	2
Maisie	Jack Russell Terrier	F	5	Barrier	4	2	4	2
Arlo	Springer Spaniel	M	3	Barrier	4	2	4	4
Maisie	Labrador Retriever x Airdale Terrier	F	8	Barrier	4	3	4	4

Table 2
Study 2 subject and performance data

Name	Breed	Sex	Age (Years)	Condition	Trials completed	Trials correct
Kiko	Yorkshire Terrier	M	6	Barrier	4	2
Charlie	King Charles Cavalier Spaniel	M	1	Barrier	4	1
Poppy	King Charles Cavalier Spaniel	F	7	Barrier	2	0
Elliot	King Charles Cavalier Spaniel	M	10	Barrier	4	4
Daisy	King Charles Cavalier Spaniel	F	4	Barrier	0	-
Amber	Miniature Dachshund	F	2	Barrier	4	4
Nacho	Chihuahua	M	3	Barrier	4	4
Hoover	Labrador Retriever x Springer Spaniel	M	0	Barrier	4	3
Bo	Setter	M	5	Barrier	4	1
Rupert	Setter	M	6	Barrier	4	4
Chilli	Chilli	F	0	Barrier	4	2
Angel	Siberian Husky	F	4	Barrier	4	2
Missy	Siberian Husky x Staffordshire Bull Terrier	F	4	Barrier	4	1
Kano	Staffordshire Bull Terrier	M	4	Barrier	4	3
Marj	Irish Water Spaniel	F	7	Barrier	4	4
Kobe	Siberian Husky	M	1	Barrier	4	2
Digby	Cocker Spaniel x Poodle	M	2	No Barrier	4	1
Phantom	Siberian Husky	M	9	No Barrier	4	0
Jet	Siberian Husky x Malamute	M	8	No Barrier	4	3
Lucy	Labrador Retriever	F	7	No Barrier	4	4
Saffron	Mini Pinscher x	F	1	No Barrier	4	2
Lucy	Cavalier King Charles Spaniel x Mini Poodle	F	8	No Barrier	4	3
Lily	Jack Russell x Shih Tzu	F	5	No Barrier	4	2
Lady	Mongrel	F	3	No Barrier	4	2
Reggie	Whippet x Collie	M	0	No Barrier	4	3
Jax	Greyhound Rhodesian Ridgeback	M	3	No Barrier	4	2

Table continues..

Sandy	Labrador Retriever x Poodle	F	11	No Barrier	4	2
Ronnie	French Bulldog	M	2	No Barrier	4	3
Margot	Miniature Dachshund	F	2	No Barrier	4	4
Bella	Chihuahua x Jack Russell	F	3	No Barrier	4	1
Tommy	Chihuahua	M	5	No Barrier	2	0
Baggins	Labrador Retriever	M	5	No Barrier	4	4
Alfie	Chihuahua	M	5	No Barrier	4	3
Blossom	Chihuahua	F	1	No Barrier	3	1

Paper 6

The Performance of Domestic Dogs (*Canis familiaris*) on Two Versions of the Object
Choice Task

Hannah Clark and David A. Leavens

University of Sussex

This paper is written in the style of article appropriate for *Animal Cognition*.

Author Notes

HC and DAL designed the experiment, HC collected the data, conducted the analyses and wrote the manuscript. DAL provided advice on analyses and comments on drafts.

Abstract

Object Choice Task (OCT) studies are used as the evidence base for theories concerning the evolutionary roots of certain socio-cognitive skills, such as the comprehension of communicative intention in a number of species. Recent reviews have found systematic procedural and methodological differences in studies which compare performances across species on the OCT. One such difference concerns the spatial configuration of the test set-up, specifically the distances between the two containers (inter-object distance) and the subject-experimenter distance. Here, we tested dogs on a central and peripheral version of the task, with half of the subjects tested with a barrier in the testing environment (as nonhuman primates are tested) and the other half without. We found that dogs tested with a barrier performed significantly better in the central version and were more likely to fail to make a choice in the peripheral version. Dogs tested without a barrier showed comparable performance on the two versions. We discuss possible explanations for these differences in terms of the Referential Problem Space.

Object Choice Task (OCT) studies are frequently used as evidence in support of theories which speculate about the evolutionary roots of socio-cognitive skills in comprehending deictic, referential gestures and intention in a range of vertebrate species. The OCT is designed to assess an individual's ability to comprehend deictic cues, such as pointing, and involves an experimenter baiting one of two or three containers and then indicating the location of the hidden reward to the subject using a directive cue, such as pointing or direct gaze. Over multiple trials, above-chance performance in selecting the baited container constitutes evidence of subjects' comprehension of the communicative cue. Interest in pointing comprehension abilities stems from gestural origins theories of language, which assert that spoken language evolved from complex gestural communication developed as an adaptive response to life on the savannah, after humans (*Homo sapiens*) separated from the rest of the hominid lineage (Arbib, Liebel & Pika, 2008; Corballis, 1999). Specific focus is given to declarative points, which are defined as points intended to direct attention (Bates, Camaioni & Volterra, 1978), distinguished from imperative points, which function to obtain a desired object, and which have been linked to the emergence of joint attentional skills and language (Dawson, Toth, Abbott, Osterling & Munson, 2004). The relative abilities of nonhuman primates, therefore, in comprehending pointing cues have been widely studied, and their apparently poor performance (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Itakura, Agnetta, Hare & Tomasello, 1999; Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012; Tomasello, Call & Gluckman, 1997) used to suggest that the understanding of declarative points is, among primates, a human-unique ability and, further, that it entails a suite of socio-cognitive skills such as theory of mind (Baron-Cohen, 1995), shared intentionality (Tomasello &

Carpenter, 2007) and cooperation (Moll & Tomasello, 2007) that other primates do not possess.

In contrast, domestic dogs' apparently consistently good performance on the OCT (e.g. Riedel, Schumann, Kaminski, Call, & Tomasello, 2007; Viranyi et al., 2008) has been used to promote domestication theories, built on the premise that, through years of domestication, dogs have evolved specialised socio-cognitive skills which endow them with the ability to comprehend human gestural cues (Hare & Tomasello, 2005). Both the domestication hypothesis and the claims for human uniqueness in declarative cue comprehension have been drawn into question recently by reviews that argue that a number of procedural confounds that are prevalent in the comparative literature prohibit group-to-species generalisations, on which evidence claims of species differences are grounded. For example, Lyn (2010) highlighted the importance of rearing history in apes to their understanding of human communicative conventions. Moreover, Lyn, Russell & Hopkins (2010) demonstrated the importance of pre-experimental exposure to humans in the development of pointing comprehension, finding that enculturated apes, who have backgrounds rich in exposure to humans, display significantly higher success rates in point-following than institutionalised apes with less pre-experimental history with humans. Russell, Lyn, Schaeffer and Hopkins (2011), whose ape samples were matched for age, sex and species, reported similar results, with the enculturated sample performing at similar levels to those reported for 2.5-year-old human children in previous studies. Similar effects have been found for dogs, with those with less experience of human interaction than the typically tested pet dogs (e.g., kennel-raised or shelter dogs) demonstrating lower success rates on the OCT (D'Aniello et al., 2017; Lazarowski & Dorman, 2015; Udell, Dorey, & Wynne, 2010).

Mulcahy and Hedge (2012) suggested that it may be methodological differences which better account for the ape-dog disparities in performance compared with phylogenetic explanations. They defined two versions of the OCT, the central version and the peripheral version. The central version involves the subject, experimenter and containers being positioned such that they are in close proximity, usually with subject and experimenter facing each other at a small table on which the containers are placed. In the peripheral version, in contrast, the subject and experimenter face each other at a distance of around 2m, both equidistant to the containers which are placed on the floor, around 2m apart. They outlined two ways in which the differential positioning of the testing apparatus may impact an individual's performance. First, in the central version of the task, the subject has both the experimenter and the containers in their direct line of vision. Thus, the salience of the containers, one of which contains food (as the subject is aware), may distract the subject's attention away from the cue being given. Alternatively, the subject may lack the behavioural inhibition to prevent it from grabbing one of the in-reach containers. Second, in the peripheral version, the containers are placed at such a distance that the subject must locomote to the container in order to retrieve the hidden reward. Such extra effort required to obtain the hidden food may result in the subject paying increased attention to the cue being given by the experimenter. In a review of OCT studies, Mulcahy and Hedge (2012) argued that there is a tendency to test apes with the central version of the task, and dogs with the peripheral, and this was supported by Clark, Elsherif, and Leavens (2019), who, in a review of 71 dog and ape OCT studies, found that dogs were tested with significantly greater inter-container distances than nonhuman primates. Mulcahy and Call (2009) and Mulcahy and Suddendorf (2011) both found increased performance by great apes when tested with a peripheral rather than central version of the task, providing support for

Mulcahy and Call's (2009) assertion that this methodological difference can affect individuals' performance on the OCT. Similarly, Clark et al. (2019) found inter-object distance to be positively correlated with performance for both dogs and apes on a number of pointing cues on the OCT.

In the one study to date that has directly compared dogs' performances on a peripheral versus a central version of the task, Kraus, van Waveren and Huebner (2014) found that dogs performed above chance in both versions, but that there was a significant decrease in performance associated with the central version. Specifically, they found a success rate difference of 15% between the two conditions and they noted the similarity with Mulcahy and Call's (2009) ape subjects whose performance in the two conditions differed by 17%. Kraus et al. (2014) therefore argued that their results provided support for the distraction hypothesis in dogs. Whilst Kraus et al. (2014) aimed to match their study design as closely as possible to that of Mulcahy and Call (2009), it differed in a number of ways. First, dogs were not tested within test cages, a cross-species difference that Leavens, Bard & Hopkins (2017) and Clark et al. (2019) note is prevalent in much of the comparative OCT literature and which Clark and Leavens (2019) and Kirchhofer et al. (2012) found can have a detrimental effect on dogs' performance. Second, the configuration of the test set-up differed such that the peripheral inter-object distance was 140cm and the central inter-object distance was 45cm, whereas Mulcahy and Call used distances of 250cm and 60cm, respectively. Finally, Kraus et al. (2014) presented subjects with an ipsilateral momentary proximal point in both conditions, which involves presenting the pointing cue for 3 seconds before retracting the hand (as per Miklósi & Soproni, 2006), with the distance between the experimenter's finger and the container being around 20cm in the central condition and 30cm in the peripheral condition. In comparison, Mulcahy and Call (2009) used a

cross-lateral dynamic pointing cue, which involves pointing across the body with the hand contralateral to the correct container and maintaining the position until the subject makes a choice. The distance between the experimenter's finger was approximately 100cm in the peripheral condition (a distal point according to Miklósi & Soproni, 2006) and 40cm in the central condition (a proximal point). Given that the use of barriers in the testing environment in the form of cages, inter-object distance, and point type (Miklósi & Soproni, 2006; Udell, Hall, Morrison, Dorey & Wynne, 2013) have all been shown to differentially affect an individual's performance, direct comparison between Kraus et al.'s (2014) dogs and Mulcahy and Call's (2009) ape subject is subject to the effects of a number of confounds in the test setups. In the current study, therefore, we aimed to replicate as closely as possible the testing conditions used by Mulcahy and Call (2009) with a sample of pet dogs, in order to investigate further whether differences in spatial configuration affect the behavioural responses of dogs on the OCT. We matched the spatial configuration, the point cue presented, and half of the dogs were tested within a child's playpen, designed to emulate a cage. For control purposes, we also tested half of the dogs without the playpen, in order to ascertain effects which were due to the configurational manipulation and effects, if any, which were due to the presence of a barrier between the dogs and the experimenter.

Method

Subjects

Twenty-six pet dogs (14 male, 12 female) took part in the study. Dogs ranged in age from 4 months to twelve years ($M = 4.0$ years, $SD = 3.1$) and comprised a variety of breeds (see Table 1 for individual subject data). Dogs were recruited through advertisements placed on social media and face-to-face recruitment in a local park.

None of the dogs had ever taken part in any studies before. Subjects were tested individually by an unfamiliar experimenter, and testing took place in a local community hall. Subjects were randomly assigned to conditions prior to testing. Two dogs were excluded from the final analysis; one because she failed to complete testing due to being apparently highly anxious within the playpen, and the other because the video-recording of the test session was lost due to a technical error.

Procedure

On arrival at the hall, dogs were given time to explore the room, off lead, and the experimenter interacted with the dog and offered them treats, in order for them to become familiar with both the testing environment and the experimenter. The experimenter then showed the dogs the cups, allowing them to sniff them, and then, in sight of the dog, hid a treat under one of the cups, then showed it to the dog and gave it to them. This was repeated twice to show the dog that the experimenter had treats, and that these could be found under the test cups. The owner was then asked to put the dog on a 1m lead and to stand with the subject on a marked point in the test room (this was inside a playpen for dogs tested with a barrier; for those tested without, it was the same location in the room but without the barrier present). The experimenter then baited both cups in view of the dog, saying “*name... look*” as she hid a treat under each cup. She then returned to a point equidistant between the two cups, called the dog’s name and established eye contact, then pointed to one of the cups using a cross-lateral, dynamic pointing cue. In the central version, the distance between the experimenter’s finger and the cup was around 40cm; in the peripheral version it was around 100cm (see Figure 1 for the configurational set-up of the two conditions). The owner held the subject on the lead while the baiting took place and released the lead when the experimenter pointed. The experimenter maintained the pointing cue until either the subject made a choice

(described below). If the subject made a correct choice, the experimenter gave the subject the piece of food (if they had not already retrieved it). If the subject made an incorrect choice, both pieces of food were removed from the containers and the trial ended. If the subject failed to make a choice, the trial ended after 1 minute. The beginning of the trial was counted from when the pointing cue was presented and the subject was released to make a choice (Udell, Dorey & Wynne, 2010b). Subjects were given 4 trials per condition, with the order of conditions counterbalanced prior to testing. The cued (correct) container was on the left or the right an equal number of times and the order was counterbalanced such that it never appeared on the same side for more than two consecutive trials.

Materials and Set-up

The playpens used in the barrier condition were two Dreambaby Royal Converta 3-in-1 Playpen Gates, measuring 380 x 4 x 74cm (Dreambaby, Unit 53, Rosyth Business Centre, 16 Cromarty Campus, Rosyth, KY11 2WX, Scotland). The containers used to hide the bait were two opaque plastic cups. A premium commercial dry dog food was used for baiting the cups. All dogs were tested on a 1m long lead. All testing sessions were recorded on a Sony Handycam HDR-PJ410 video-camera (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan). In the central version of the task, the experimenter kneeled at a distance of 60cm from the subject and the distance between the two containers was 60cm; in the peripheral version, the subject-experimenter distance was 110cm and the inter-object distance was 250cm (see Figure 1).

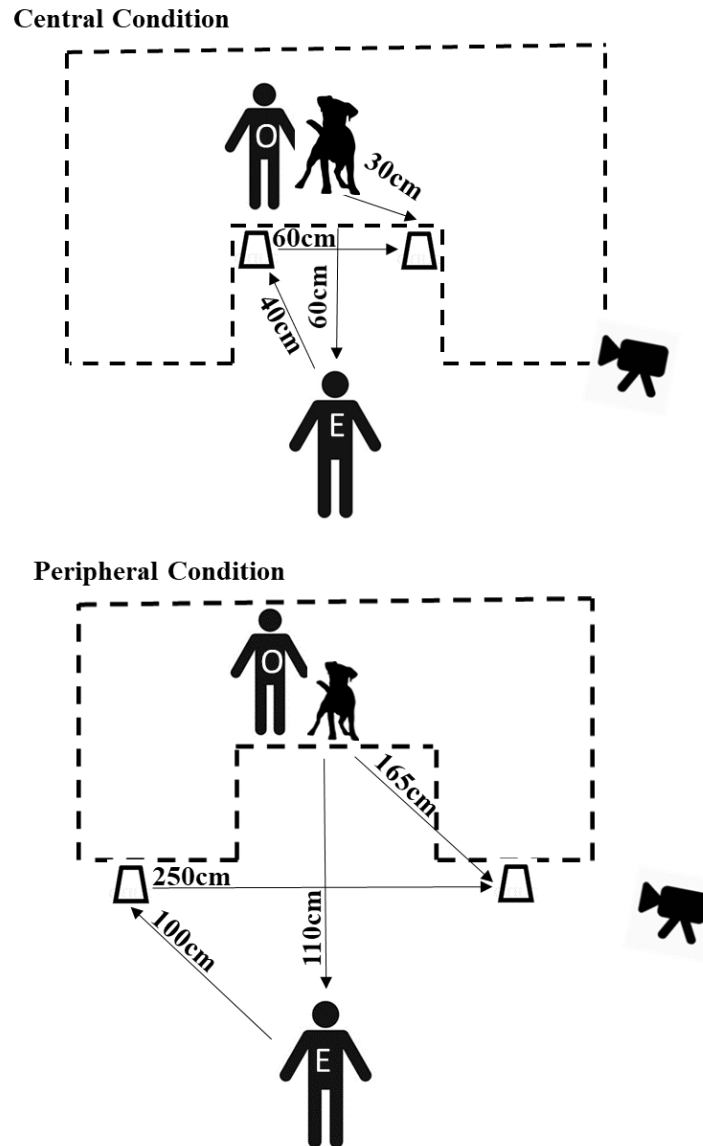


Figure 1: The configurational set-up of the central and the peripheral conditions. O = Owner; E = experimenter. Dashed line indicates playpen. Figure not to scale.

Data Scoring

All testing sessions were video recorded and coded at a later date. Trials were coded for correct/ incorrect response and response latency. Following Udell et al's (2010b) recommendations, a correct response was categorised as one in which the subject first touched or came within 10cm with their snout of the correct container. Any other response was categorised as incorrect. Incorrect responses were further

categorised into incorrect choices, in which the subject first touched or came with 10cm with their snout of the incorrect cup, and no choice, in which the subject failed to choose one of the cups before the end of 1 minute.

Reliability

All trials were coded by the first author, and a second coder, who was blind to the purposes of the study, coded a random sample of 20% of subjects' testing sessions, that is, five dogs' testing sessions constituting 40 trials. For correct choices, there was 100% agreement between the two coders, $\kappa = 1.00$, $p < .001$. There was excellent agreement for whether or not the subject made a choice, $\kappa = .81$, $p < .001$, and response latency, $r_s = .99$, $p < .001$.

Results

Dogs Tested with a Barrier

Correct choices.

Due to a lack of normal distribution in the data, all analyses used nonparametric tests. As a group, dogs performed above chance in the central version of the task, (binomial test, $p < .001$) but not in the peripheral version (binomial test, $p = .665$). Dogs chose the correct container on a significantly higher proportion of trials in the central version ($Mdn = .88$) than in the peripheral version ($Mdn = .50$) of the task, $Z = -2.46$, $p = .014$. This shows that dogs tested with a barrier were more accurate in their responses in the central version than in the peripheral version. Figure 2 shows the percentage of trials on which the dogs tested with a barrier chose the correct cup, the incorrect cup and failed to make a choice.

Incorrect choice vs. no choice.

There was no significant difference in the proportion of responses that were incorrect choices between the central version ($Mdn = 0.00$) and the peripheral version ($Mdn = 0.00$) of the task, $Z = -1.00$, $p = .317$. The proportion of trials on which the dogs failed to make a choice was significantly lower in the central version ($Mdn = .00$) than in the peripheral version ($Mdn = .50$), $Z = -2.27$, $p = .023$. This shows that the dogs were equally likely to make an incorrect choice in the two versions of the task, but that they were more likely to fail to respond when the cups were placed further apart.

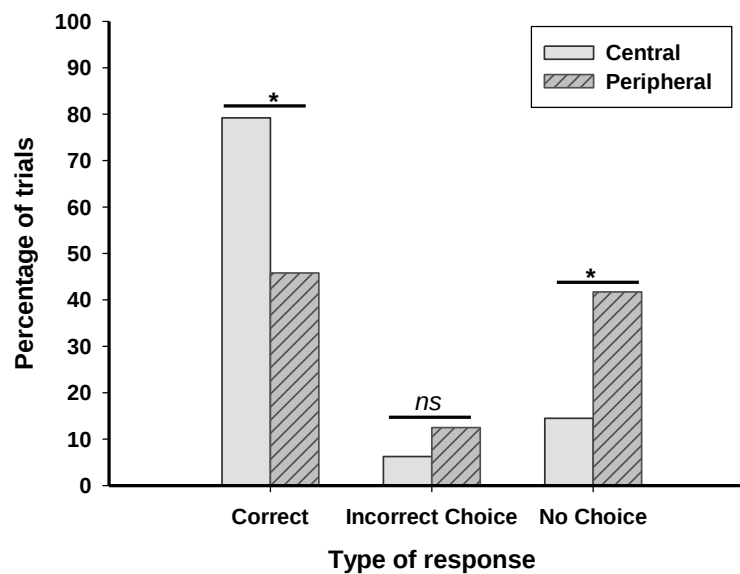


Figure 2: The percentage of trials in which dogs tested with a barrier made a correct choice, an incorrect choice and no choice in the central and peripheral versions of the task. Total number of trials per condition = 48. * denotes significant at $p < .05$.

Dogs Tested without a Barrier

Correct choices.

As a group, the dogs performed above chance in both the central version (binomial test, $p = .002$) and the peripheral version (binomial test, $p = .013$) of the task. There was no significant difference in the proportion of trials on which the dogs chose the correct cup between the central version ($Mdn = .75$) and the peripheral version (Mdn

= .75) of the task, $Z = -0.29$, $p = .774$. This shows the dogs tested without a barrier performed equally well on both versions of the task. Figure 3 shows the percentage of trials on which dogs tested without a barrier chose the correct cup, the incorrect cup and failed to make a choice.

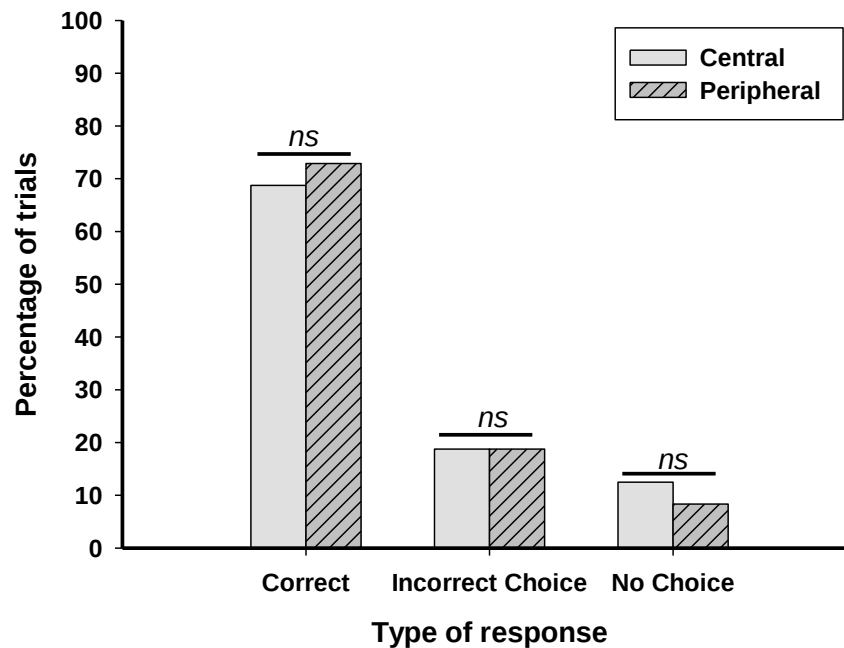


Figure 3: The percentage of trials in which dogs tested without a barrier made a correct choice, an incorrect choice and no choice in the central and peripheral versions of the task. Total number of trials per condition = 48. * denotes significant at $p < .05$.

Incorrect choice vs. no choice.

There was no significant difference between the proportion of trials on which the dogs made an incorrect choice between the central ($Mdn = .25$) and the peripheral version ($Mdn = .13$), $Z = -0.14$, $p = .890$. There was also no significant difference between the proportion of trials in which the dogs failed to make a choice between the central ($Mdn = .00$) and the peripheral version ($Mdn = .00$), $Z = -378$, $p = .705$. This shows that there was no difference between the type of incorrect responses the dogs exhibited between the two versions of the task.

Barrier vs. No Barrier Comparisons

Correct Choices.

On the central version of the task, there was no significant difference in the proportion of trials in which the dogs chose the correct cup between those tested with a barrier ($Mdn = .75$) and those tested without ($Mdn = .88$), Mann-Whitney $U = 58.5$, $p = .411$. Nor was there a significant difference between those tested with a barrier ($Mdn = .50$) and those tested without ($Mdn = .75$) on the peripheral version, Mann-Whitney $U = 42.00$, $p = .073$. This shows that the dogs tested with and without a barrier chose the correct cups on a comparable number of trials in the two versions of the task. Figure 4 shows the percentage of trials in which the dogs tested with and without a barrier chose the correct cup, the incorrect cup and failed to make a response in a) the central version and b) the peripheral version of the task.

Incorrect choice vs. no choice.

There was no significant difference in the proportion of trials in which dogs made incorrect choices in the central version of the task between dogs tested with a barrier ($Mdn = .00$) and dogs tested without a barrier ($Mdn = .25$), Mann-Whitney $U = 45.00$, $p = .075$. There was also no difference in the proportion of incorrect choices on the peripheral version between dogs tested with a barrier ($Mdn = .00$) and dogs tested without ($Mdn = .13$), Mann-Whitney $U = 60.00$, $p = .434$. On the central version of the task, there was no difference in the proportion of trials on which the dogs failed to make a choice between dogs tested with a barrier ($Mdn = .00$) and dogs tested without a barrier ($Mdn = .00$), Mann-Whitney $U = 67.00$, $p = .719$. On the peripheral version of the task, dogs tested with a barrier failed to make a choice on a significantly higher proportion of trials ($Mdn = .50$) than dogs tested without a barrier ($Mdn = .00$), Mann-

Whitney $U = -2.41$, $p = .016$. This shows that the dogs tested with and without a barrier exhibited comparable types of incorrect responses on the central version of the task, but that dogs tested with a barrier showed an increased tendency to fail to make a choice on the peripheral version.

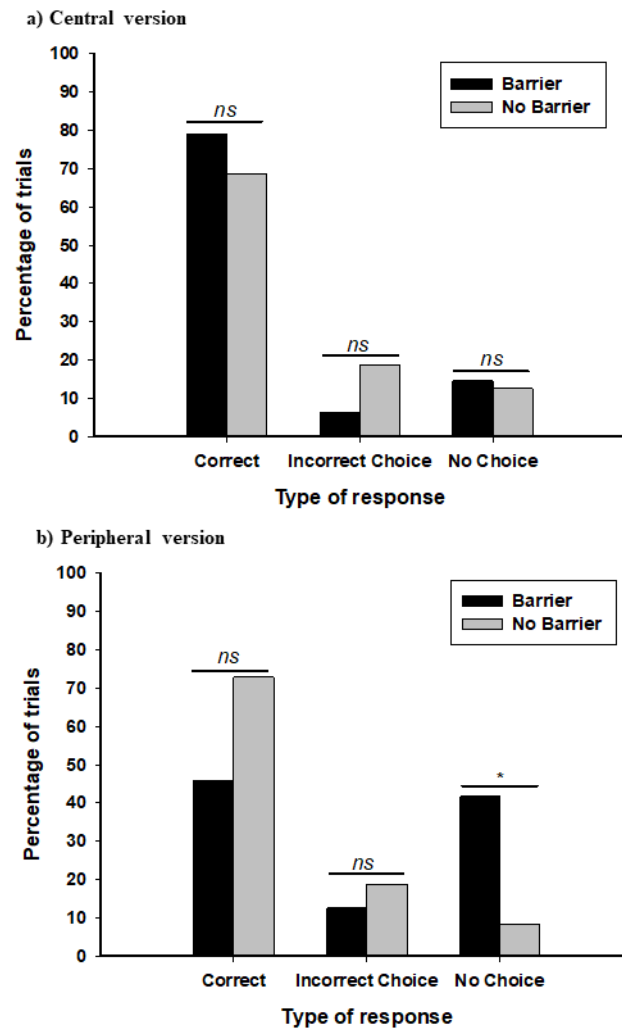


Figure 4: The percentage of trials in which dogs tested with and without a barrier made a correct choice, an incorrect choice and no choice in a) the central and b) the peripheral version of the task. Total number of trials per condition = 48. * denotes significant at $p < .05$.

Order of Administration

For dogs tested with a barrier, there were no effects of order of administration on the proportion of correct choices in the central version (central first $Mdn = 1.00$; peripheral first $Mdn = .75$), Mann-Whitney $U = 17.00$, $p = .930$, nor in the peripheral version (central first $Mdn = .50$; peripheral first $Mdn = .50$), Mann-Whitney $U = 17.50$, $p = 1.00$.

For dogs tested without a barrier, there were no effects of order of administration on the proportion of correct choices in the central version (central first $Mdn = .88$, peripheral first $Mdn = .75$), Mann-Whitney $U = 14.50$, $p = .560$, nor in the peripheral version (central first $Mdn = .63$; peripheral first $Mdn = .88$), Mann-Whitney $U = 12.00$, $p = .312$. This shows that both for dogs tested with and dogs tested without a barrier, there was no effect of order of administration of conditions on performance.

Trial by Trial Analyses

For dogs tested with a barrier, there was no difference in correct choices across trials in the central version, Cochran's $Q = 3.00$, $p = .392$, nor in the peripheral version, Cochran's $Q = 0.55$, $p = .909$.

For dogs tested without a barrier, there was no difference in correct choices across trials in the central version, Cochran's $Q = 1.44$, $p = .697$, or in the peripheral version, Cochran's $Q = 6.14$, $p = .105$. This shows that both for dogs tested with and dogs tested without a barrier, there was no effect of successive administrations on their tendency to choose the correct container.

Discussion

We attempted to replicate, as closely as possible, Mulcahy and Call's (2009) ape study, testing dogs on a central and a peripheral version of the OCT. Dogs tested from within a playpen, thus those tested in the most closely matched conditions to the apes in Mulcahy and Call's (2009) study, performed significantly better in the central version than in the peripheral version; indeed in the latter condition, they failed, as a group, to perform above chance. This contrasts with Mulcahy and Call (2009) and Kraus et al.'s (2014) findings in which apes and dogs, respectively, showed increased performance in the peripheral version. Furthermore, the dogs failed to make a choice significantly more often in the peripheral version of the task, in which the reward was placed further away yet was still retrievable, than in the central version, and thus a possible explanation for this difference in performance is that the presence of a barrier had a suppressive effect on the dogs' tendency to attempt to retrieve the reward when it was further away.

We found no performance differences for dogs tested without a barrier between the two configuration set-ups, again contrasting with Mulcahy and Call (2009) and Kraus et al.'s (2014) findings. Dogs chose the correct cup on 69% of central version trials compared with 73% on peripheral version trials (non-significant difference), whereas Kraus et al. (2014) reported success rates of 63% and 78% (significant difference), respectively. A possible explanation for this difference in findings is the procedural differences between our study and Kraus et al.'s (2014). Kraus et al. (2014) presented subjects with an ipsilateral, momentary proximal pointing cue, meaning subjects were presented with the cue for just 3 seconds. In the current study, we used a dynamic cue, involving presentation of the pointing cue until the subject made a choice. As Miklósi and Soproni (2006) highlight momentary and dynamic cues differ in terms of the memory demands on subjects- with a momentary cue, the subject must attend to

and remember the direction of the cue, whereas with a dynamic cue, no such memory demands exist. It may be, then, that dogs presented with a momentary cue are more likely to be subject to distraction if the containers are in their direct line of vision, because not only must they attend to the cue over the highly salient containers, they must also remember the direction of the cue and use this to inform their response.

When comparing dogs tested with and without a barrier, we found no significant differences in performance in either version of the task. This was surprising, given that Kirchhofer et al. (2012) and Clark and Leavens (2019) found suppressive effects on performance associated with the imposition of a barrier. On the central version of the task, dogs tested with a barrier actually had a higher success rate, choosing the correct cup on 75% of trials, than those tested without, who had a success rate of 69%. One explanation for this is that the barrier perhaps reduces dogs' susceptibility to distraction—having the barrier between the subject and the containers may decrease the salience of the cups and promote attention to the cue. In contrast, in the peripheral version of the task, although there was no significant difference, there was a trend towards poorer performance associated with the presence of a barrier, and dogs failed to make a choice on a significantly greater number of trials. This echoes the findings of Clark and Leavens (2019) who found a similar increase in no choice responses when dogs were tested with a barrier. Clark and Leavens (2019) suggested that this may be explained by the Referential Problem Space (Leavens, Hopkins & Bard, 2005), that is, the barrier may increase perceptions of the reward as being unobtainable (even though, it is, in fact obtainable). This perception coupled with the extra effort associated with obtaining a reward which is placed far away from the subject in the peripheral condition, may explain, in the current study, why the dogs tested with a barrier failed to make a choice on 42% of peripheral trials.

Another possible factor to consider is the visual acuity of the individual and the extent to which they may be able to see either the structure of the barrier (be that bars of wire mesh) and the objects behind it, with this being influenced by the distance between the subject and the apparatus as well as variables such as age and species (see Bard, Street, McCrary & Booth, 1995 for a discussion of the development of visual acuity in chimpanzees and Miller & Murphy, 1995, for a discussion of visual acuity in dogs).

In conclusion, the current study shows a complex picture of the way in which configuration and the presence of a barrier can affect dogs' responses on the OCT. We did not find support for the distraction hypothesis in dogs tested either with or without a barrier, for the former, in fact, the central version facilitated performance, and for the latter, that performance was comparable. This contrasts with Kraus et al.'s findings and further highlights the effects that cue types may have on performance, as noted by Miklósi and Soproni (2006) and Udell et al. (2013). Indeed, an interesting future direction would be to further investigate how configuration and cue type may interact to affect subjects' performance. This could be done by testing dogs across a range of cue types differing in their distance and temporal properties, similarly to Udell et al. (2013), using both a central and a peripheral configuration.

We found the presence of a barrier to affect behavioural responses in the peripheral version, and a statistically non-significant trend towards this finding in the central version which echoes Clark & Leavens' (2019) findings in which greater sample sizes were used.

Here we found a complex interaction between two environmental influences on performance in dogs, with responses to central and peripheral versions of the OCT differing as a function of the presence of a barrier. We found no support for the

distraction hypothesis when no barrier was present, with dogs performing comparably well on both versions of the task, and we found that dogs' performance in the peripheral version was negatively affected by the imposition of a barrier. This study adds to the growing OCT literature which emphasises the necessity of addressing systematic confounds prior to speculating about the evolutionary roots of socio-cognitive skills based on apparent species difference in performance. Without consideration of the effects of these confounding variables, it is illogical to attribute subjects' responses to their selective histories.

Acknowledgements

We thank the School of Psychology and the Doctoral School, University of Sussex for funding these studies. Dan Hyndman and Martha Casey provided technical support. We thank Kassandra Giragosian for assistance with reliability coding and Milton Village Community Association for generous and helpful access to their facilities. Finally, we thank the pets and owners who took part in the study.

Table 1
Subject and performance data

Name	Breed	Sex	Age (years)	Barrier	First condition	Central trials correct	Peripheral trials correct
Hendrix	Shih Tzu x Chihuahua	M	7	Barrier	Central	1	0
Evie	Lurcher	F	3	Barrier	Central	4	2
Abbie	Cavalier King Charles Spaniel	F	0.75	Barrier	Central	2	2
Theo	Cocker Spaniel x Poodle	M	4	Barrier	Peripheral	4	4
Marnie	Yorkshire Terrier x Jack Russell	F	4	Barrier	Peripheral	4	1
Bruce	Manchester Terrier x Jack Russell	M	6	Barrier	Central	4	4
Dudley	Springer Spaniel	M	12	Barrier	Peripheral	4	2
Stanley	Cocker Spaniel x Poodle	M	0.33	Barrier	Central	4	1
Olly	Cocker Spaniel	M	0.83	Barrier	Peripheral	3	0
Eva	Red Fox Labrador	F	6	Barrier	Peripheral	3	4
Axie	Mongrel	F	4	Barrier	Peripheral	3	2
Bob	Jack Russell	M	4	Barrier	Peripheral	2	0
Jake	Springer Spaniel x Poodle	M	1	No Barrier	Peripheral	4	4
Gary	English Mastiff	M	2	No Barrier	Peripheral	3	4
Floki	Border Collie	F	3	No Barrier	Peripheral	3	3
Leyla	Cavalier King Charles Spaniel x Poodle	F	4	No Barrier	Central	0	3
Watson	Border Collie	M	5	No Barrier	Central	4	4
Tallulah	Jack Russell x Poodle	F	3	No Barrier	Peripheral	3	4
Bruce	Border Collie	M	12	No Barrier	Peripheral	2	3
Tilly	Border Collie	F	1	No Barrier	Central	2	0
Ruby	Jack Russell	M	4	No Barrier	Peripheral	1	2
Elvis	Cocker Spaniel	M	7	No Barrier	Central	4	4
Cookie	French Bulldog x Pug	F	1	No Barrier	Central	3	2
Penny	Boston Terrier	F	2	No Barrier	Central	4	2

Thesis Discussion

Hannah Clark

University of Sussex

Summary of Papers

This thesis adds to current understanding about how systematic differences in procedure and method can affect individuals' performances on the Object Choice Task (OCT). In Paper 1, the most comprehensive review of dog and nonhuman primate studies to date was conducted of studies published up until 2017 and highlighted systematic differences in the sampling and testing protocols used in the OCT literature. In examining subjects' prior exposure to humans, we found that 91% of the dog subjects studied had pre-experimental histories rich in human exposure compared with just 6% of nonhuman primate subjects. Furthermore, on a number of pointing cues, there were significant differences in performance associated with pre-experimental history across representatives of both taxonomic groups. In addition, I found that level of prior exposure to humans, spatial configurations, and point types had significant effects on performance on the OCT.

Paper 1 also presented the first quantitative examination to date of the prevalence of barriers in nonhuman primate and dogs subjects, finding that 99% of nonhuman primate subjects are tested with a barrier compared with less than 1% of dogs. In addition, I found markedly greater inter-object distances were used when testing dogs than nonhuman primates and that higher inter-object distances were found to be positively correlated with performance across both dogs and nonhuman primates on a range of pointing cues. Finally, I found that the pointing cues presented to nonhuman primates and dogs systematically differed in their distance, temporal and lateral features, and that different pointing cues were associated with differences in performance.

In Paper 2, I tested 18-month-old and 36-month-old children on an OCT with and without a barrier. These age groups were chosen because children from 14 months reliably follow pointing cues and so this would allow an examination of the effects on behavioural responses attributable to the experimental manipulation rather than the lack of emergence of pointing comprehension skills. Here, clear effects of a barrier were shown, with an increase of communicative behaviour associated with its presence. Thus, when a barrier was present, children's tendency to communicate their choice of container by pointing increased, whereas in the absence of a barrier, children tended to grab the container in acts of direct prehension.

In Paper 3, I investigated whether the use of different spatial configurations, specifically central vs. peripheral set-ups would affect children's responses on the OCT, in the first study to date in which children have been tested with a peripheral version of the task. I tested 36-month-old children on the two versions of the task replicating Mulcahy and Call's (2009) study with apes, finding no effect of configuration on their responses. I did, however, find a decrease in grabbing behaviour associated with the presence of a barrier, echoing the findings of Paper 2.

In Paper 4, I described the differential reward types offered as an incentive in the OCT between nonhuman primates and human children and examined whether presenting food or sticker rewards would affect the 36-month-old children's behavioural responses, finding no effects of reward type on children's responses.

In Paper 5, I investigated the effect of a barrier on dogs' performance on the OCT. I used an ipsilateral dynamic proximal point in Study 1 and a cross-lateral dynamic proximal point in Study 2 and found decreased performance associated with the barrier when using the latter, but not the former. There were also differences in the

types of responses elicited; specifically, the presence of a barrier was associated with increased instances of failing to choose a container and the absence of a barrier with increased instances of an incorrect choice.

Finally, in Paper 6, I examined the effects of differences in spatial configuration on dogs' behavioural responses on the OCT, replicating again Mulcahy and Call's (2009) ape study. I found no effect of configuration when dogs were tested without a barrier, but worse performance in the peripheral version in dogs tested with a barrier. There was also an increase in the number of trials in which dogs failed to choose a container associated with the presence of a barrier, as in Paper 5.

In the Introduction, a number of criticisms of the existing OCT literature were described, with reference to previous reviews. Here, I summarise how this thesis extends current knowledge with respect to these.

Pre-Experimental History

Lyn (2010) highlighted the importance of pre-experimental exposure to humans to apes' performance on the OCT in her review of 9 nonhuman primate studies. Subsequent studies with both apes (e.g. Lyn, Russell and Hopkins, 2010; Russell, Lyn, Schaeffer and Hopkins, 2011) and dogs (e.g. Udell, Dorey & Wynne, 2010; D'Aniello et al., 2017; Lazarowski & Dorman, 2015) showed that individuals with life histories rich in human exposure, for example, enculturated apes and pet dogs, demonstrate superior performance to individuals with backgrounds with less human exposure, for example, institutionalised apes and kennel or shelter dogs. This thesis builds on existing knowledge, first, by providing quantitative evidence of the prevalence of failing to match nonhuman primate and dog subjects for pre-experimental history. Second, it provides support to the numerous empirical studies which report effects of pre-

experimental history on performance in representatives of just one taxonomic group by demonstrating such effects across a much larger sample. Finally, the first incarnation of the meta-analysis reported in the meta-analysis presented in Paper 1 included all published OCT studies, collating data for 43 vertebrate species. As a result of this, the Human Experience Scale developed in this paper featured a wide range of rearing histories, specific to representatives of different taxonomic groups, and is therefore not limited to use with nonhuman primates and dogs. This means that in Paper 1, a Human Experience Scale is provided which can be used by future researchers working in a wide range of comparative domains to facilitate matching subjects for pre-experimental history ensuring that comparisons are made across comparable groups.

Barriers

Leavens et al. (2017) discussed the testing of nonhuman primate subjects from within cages in terms of the barrier this imposed between the subject and the experimenter and testing apparatus, noting that we cannot assume that this does not affect performance. Only one study to date has looked at the effect of barriers on dogs' performance with a barrier on a pointing cue on the OCT (Kirchhofer et al., 2012), and they reported a diminished success rate, relative to a condition with no barrier. The meta-analysis presented in Paper 1 showed a marked difference in the prevalence of barriers between nonhuman primates and dogs and a subsequent objective of the empirical studies in this thesis was to investigate their effect on behavioural responses on the OCT using a systematic series of experiments. This thesis thus builds on existing knowledge by informing us of the marked difference on behavioural response that the presence of a barrier can have. Children of 18 and 36 months of age are already experts in producing and comprehending human pointing cues, therefore finding that their reactions on the OCT can be differentially affected by the presence or absence of a

barrier has implications for studies which fail to match this aspect of the testing environment when comparing across species.

The studies presented here which showed an effect of a barrier on OCTs with dogs build on existing knowledge by providing partial support for Kirchhofer et al.'s (2012) findings with respect to performance, but also by showing that, irrespective of performance, the behavioural responses elicited from dogs systematically differ as a function of the presence of a barrier. This again highlights the necessity of matching for this aspect of the testing environment when comparing across samples.

Central vs. Peripheral Versions of the OCT

Mulcahy and Hedge (2012), in a review of 63 published studies, suggested that there existed disparity in the testing environments used with dogs and nonhuman primates in terms of their spatial configurations, distinguishing between central and peripheral versions of the task. Mulcahy and Call (2009) and Mulcahy and Suddendorf (2011) found performance differences in ape samples of 12 and 1, respectively, with higher success rates exhibited on the peripheral than the central version of the task. The meta-analysis presented in this thesis confirmed and extended this across a larger database showing systematic differences in the inter-object distances used with nonhuman primates and dogs and correlations between inter-object distance and performance.

Mulcahy and colleagues (Mulcahy & Call, 2009; Mulcahy & Hedge, 2012; Mulcahy & Suddendorf, 2011) attributed the differences they found to the distraction hypothesis, suggesting that the salience of the containers, when placed in the direct line of vision of the subject and between the subject and the experimenter in the central version, may distract subjects' attention away from the cue being given. That I found no

effect with 36-month-old children may be due to children of this age group having more developed behavioural control, being accustomed to having to attend to an adult's directions, or being eager to be correct leading to increased attention to the cue being given.

With regard to dogs, one paper (Kraus et al., 2014) reported support for the distraction hypothesis in dogs in a study which compared their performance on a central and a peripheral version of the OCT, however the results of their study were not directly comparable to those of Mulcahy and Call (2009) due to a number of procedural differences. The results reported in Paper 6 in which I investigated the effects of configuration on performance in dogs, interestingly, contrasted with Mulcahy and Call (2009) and Kraus et al.'s (2014) findings as well as the findings the meta-analysis presented in Paper 1, failing to find support for the distraction hypothesis in dogs. This was the first paper to date which tested dogs on two versions of the OCT in conditions comparable to those experienced by apes (i.e. with a barrier) and adds to existing knowledge by showing that an interaction of different aspects of the testing environment can lead to markedly different behavioural responses, that is, the presence of a barrier may influence the configurational effects of central and peripheral versions.

Point Types

Miklósi and Soproni (2006) differentiated between a number of pointing cues used on the OCT in terms of the temporal, distance, and lateral properties and reported that representatives of different taxonomic groups demonstrate differential success rates according to the point type presented. Udell et al. (2013) looked at the effects of point topography on dogs' performance and again found different success rates. In the review of OCT literature presented in Paper 1, systematic differences were found in the types

of cues presented to dogs and nonhuman primates, respectively, and the patterns found were consistent with Miklósi and Soproni (2006) and Udell et al. (2013): performance differences were associated with different cue types across nonhuman primates and dogs. I found a decrease in performance associated with the barrier when using a cross-lateral dynamic proximal point but not an ipsilateral dynamic point. Udell et al. (2013) report that the ipsilateral dynamic proximal point is a simpler point to follow, and this may be due to the experimenter maintaining the cue until the subject makes a choice (thus no demands are made on memory) and because it has local enhancement properties, due to being in close proximity to the container. The studies presented in this thesis extend current knowledge by providing further evidence for the effect of cue types on performance. It also emphasises the importance of both matching for cue types when testing across species, and for testing individuals with a range of cue types before drawing conclusions about their point-following abilities.

Other findings

Reward types.

Leavens et al. (2017) discussed a number of inconsistencies in testing when comparing across representatives of different taxonomic groups in a number of domains, not limited to the OCT. One such difference is the use of different rewards with human children and representatives of other taxonomic groups: toys or stickers for the former, food for the latter. Whilst this issue was not quantitatively addressed in the meta-analysis presented in Paper 1, in Paper 4, the effect of reward type on 36-month-old children's behavioural responses on the OCT was examined. To my knowledge, this is the first study to date which has investigated this, and this paper builds on current

knowledge by finding that this particular difference in procedure in comparative testing may not have the same confounding effects as other differences reported in this thesis.

Age.

A further issue not specifically examined in the review of OCT literature in Paper 1, but that has been raised by Leavens et al. (2017) is the tendency to fail to match cross-species samples for age. Bard, Dunbar, Maguire-Herring, Veira, Hayes and McDonald (2014), Bard and Leavens (2014) and Bard et al. (2014) have all described the developmental trajectory of socio-cognitive skills in nonhuman primates and highlighted its similarity to that followed by human infants, specifically that socio-cognitive skills emerge as a function of intense emotional engagement with a social partner. Leavens et al. (2017) describe the distinct lack of overlap in the ages of subjects in studies that compare humans and nonhuman primates on measures of socio-cognitive skills, noting that differences in performance may be attributable to differences in the developmental stage at which subjects were tested. In Paper 2, there was a marked difference in the responses exhibited by the 18-month-old children compared with those of the 36-month-old children, with the younger children using significantly more communicative behaviour than the older children when a barrier was present in the testing environment. This has implications for the way in which the children of the two age groups perceived the barrier. Whilst children from both age groups increased their communicative behaviour as a function of the barrier, suggesting they may have perceived it as an obstacle to obtaining a desired object, this effect was particularly enhanced in the 18-month-old children. The 3-year-olds, in comparison, showed more ability to use problem-solving tactics to overcome this obstacle: reaching through the bars to grab the container in acts of direct prehension; the 18-month-old children, although able to reach the container on the other side of the playpen, were less likely to

do so. That this difference in human children's responses to the OCT as a function of their age was found shows a developmental trajectory in the cognitive skills required to effectively respond to the problem presented by the OCT and provides support for Leavens et al.'s (2017) argument that matching for age (and developmental stage) is, as well as matching for pre-experimental history, important for cross-species comparisons.

Limitations

Sampling.

Discussed above were a number of issues in the comparative literature with failure to match samples, however, the studies presented in this thesis are not without their limitations in this regard. First, in both the studies with human children and with dogs, I relied on volunteers taking part. Henrich et al. (2010) noted the tendency of researchers to take results from Western, industrialised societies and generalise to humans as a whole. In the studies with children reported here, the samples consisted of children whose parents had responded to social media advertisements to volunteer their time to come to take part at either, in the case of Papers 2 and 4, to a child development laboratory at the University of Sussex, or, in the case of Paper 3, to a local community hall. People who have both the time and the inclination to volunteer to take part are unlikely to be representative of the population as a whole, and as such, this limits the generalisability of the results reported here. In addition, whilst index-finger pointing is, in Western societies, the predominant form of deictic signalling, it is not universal (Cooperrider, Slotta, & Núñez, 2018), and this again limits the extent to which the results here can be generalised.

Similarly, the dogs that took part in Paper 4 were all recruited from advertisements on social media, and those in Paper 5 were largely recruited from

opportunity sampling in local park. Whilst no pre-selection for dogs took place, that is, dogs were not required to have shown prior retrieving behaviours (as in Kirchhofer et al., 2012), this form of recruitment, which relies on owners to volunteer their time may lead to sampling effects (e.g., owners of particularly attentive dogs might differentially respond to recruiting advertisements).

Best Practices for Future Research

In addition to the general limitations of the studies presented in this thesis, reflection on the methodology and design of the studies has enabled me to identify areas of best practice for future research. First, in the studies with human children, children who did not respond or complete testing in a given number of trials were excluded from the final analyses. In contrast, all of the dogs' non-responses were coded as such and included in the analyses, leading to the observation of a reduction in responsiveness in the presence of a barrier. Therefore, in future studies, children's non-responses should be coded as such, in order to avoid an artificially inflated percentage of correct responses resulting from the exclusion of non-responses.

Second, in the studies with human children, the task was presented as a game, with children encouraged to attend and respond to the cue being given, being instructed that doing so correctly would result in their winning a reward. Such a level of instruction is not given when testing dogs, and indeed other nonhuman animals, and so, in future studies, in order to maximise consistency in cross-species testing, the OCT should be presented to children with the verbal directions given matched to those used with representatives of other species.

Finally, the methodology used with the dogs in the studies presented here, specifically the baiting of both containers in order to avoid asymmetrical olfactory cues,

led to some dogs being inadvertently rewarded for the incorrect response when they chose the incorrect container, knocking it over and retrieving the reward before the experimenter was able to. In future studies, in order to avoid this, food should be stored in the two containers prior to testing so that both containers smell equally of the reward, and then neither container should be baited, with the reward for following the cue to the direct container being provided directly by the experimenter after the choice has been made.

Future Directions

This thesis identified and investigated the effects of some of the existing confounds in the OCT literature, and further extension of these could provide us with more understanding of how these issues may influence an individual's performance in comprehending pointing cues. First, as with any psychological research, it would be interesting to investigate the reproducibility of the results reported here. Some of the findings, for example, the effects of pre-experimental history on performance reported in Paper 1, based on a comprehensive review of existing studies, provide further support for arguments put forward by Lyn (2010) and Udell et al. (2010). However, many of the findings here are novel, for example, the effect of a barrier on children's responses on the OCT and the interaction between spatial configuration and the presence of a barrier in inhibiting dogs' responses and are based on fairly small sample sizes. It would be of interest to see if these results are replicated in future studies using the same methods.

Second, here, I found no effect of spatial configuration on 36-month-old children's performance on the OCT, and one possible explanation for this is that, by this age, children have developed sufficient behavioural inhibition to focus on attending to the cue given. It would be of interest to investigate the effects of this manipulation with

younger children, in order to determine whether the distraction hypothesis may be applicable to human children with less developed skills in behavioural inhibition.

Finally, and perhaps, most importantly, it is hoped that the findings reported here will contribute to a culture of more rigorous scientific testing in comparative psychology, where researchers who aim to compare across-species endeavour to match representatives of those species as closely as possible, in pre-experimental history, testing environment, and cues presented, and the proliferation of theories based on studies which fail to do so is called into question.

Theoretical Implications

As discussed in the Introduction to this thesis, there exist several theories regarding the evolutionary roots of human social cognition that use as their evidence base the performance of humans and nonhuman primates on the OCT. At the root of these is the fundamental idea that humans are alone amongst hominids in their ability to comprehend declarative pointing cues, and that the underlying processes that facilitate this skill comprise human-unique abilities in the understanding of others and their attentional and intentional states. The findings presented in this thesis support a role of ontogeny not fully accounted for by such theories. Tomasello and Carpenter (2007) suggested that children comprehend the referential nature of the cue presented on the OCT, and this ability is purported to emerge around the end of the first year of life as humans begin to comprehend others as social agents with intentions that may differ from their own (Tomasello, 1995), representing an evolutionary adaptation to the cooperative and cultural demands of life in social groups. That the findings reported here show differential response profiles both between age groups and as a function of environmental influences, suggests a contribution of life experience and development in

responding to deictic cues on the OCT (see Triesch, Teuscher, Deák, & Carlson, 2006 for a related argument pertaining to gaze-following in human children).

Furthermore, that these environmental influences were found, calls into question the argument for human-uniqueness for this ability and the mental processes that some researchers claim are its basis: nonhuman primates, given a comparable level of exposure to, and interaction with, humans demonstrate comparable abilities in following pointing cues on the OCT. If then, we accept that, given certain ecological circumstances, nonhuman primates too have the capacity to comprehend declarative cues on the OCT, this has implications for the theories that claim that specialised socio-cognitive skills underlie this capacity. If evidence of human children's success on passing the OCT is used to evince theories that this demonstrates understanding of others as intentional agents (Tomasello & Carpenter, 2007) and is the beginning of the development of theory of mind (Tomasello, 1995), then apes' success on the task must too be taken as evidence of these abilities in our nearest relatives. As Leavens (2012) pointed out, such mental state attributions are in fact, immeasurable and unquantifiable: the evidence base, as it stands, could be used to support hypothetical claims as to these underlying cognitive processes, but also, to support leaner interpretations that attribute deictic cue comprehension to learning. What is clear from this thesis, and from the research that it builds on, is that greater consideration of ontogenetic influences on sociocognitive development is necessary before commitment to theories that appeal to phylogenetic explanations is warranted.

References

- Albiach-Serrano, A., Bräuer, J., Cacchione, T., Zickert, N., & Amici, F. (2012). The effect of domestication and ontogeny in swine cognition. *Applied Animal Behaviour Science*, 141, 25-35. <https://doi.org/10.1016/j.applanim.2012.07.005>.
- Anderson, J. R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioural Processes*, 37, 47-55. [https://doi.org/10.1016/0376-6357\(95\)00074-7](https://doi.org/10.1016/0376-6357(95)00074-7)
- Anderson, J. R., Sallaberry, P., & Barbier, H. (1995). Use of experimenter-given cues during object-choice tasks by capuchin monkeys. *Animal Behaviour*, 49, 201-208. [https://doi.org/10.1016/0003-3472\(95\)80168-5](https://doi.org/10.1016/0003-3472(95)80168-5)
- Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–167. <https://doi.org/10.1017/S0140525X05000038>
- Arbib, M. A., Liebel, K. & Pika, S. (2008) Primate vocalisation, gesture, and the evolution of human language. *Current Anthropology*, 49(6), 1053-1076. <https://doi.org/10.1086/593015>
- Bard, K. A., Bakeman, R., Boysen, S. T., & Leavens, D. A. (2014). Emotional engagements predict and enhance social cognition in young chimpanzees. *Developmental Science*, 17(5), 682-696. <https://doi.org/10.1111/desc.12145>
- Bard, K. A., Dunbar, S., Maguire-Herring, V., Veira, Y., Hayes, K. G. & McDonald, K. (2014) Gestures and social-emotional communicative development in chimpanzee infants. *American Journal of Primatology*, 76, 14- 29. <http://dx.doi.org/10.1002/ajp.22189>

- Bard, K. A., & Leavens, D. A. (2014). The importance of development for comparative primatology. *Annual Review of Anthropology*, 43, 183-200.
<https://doi.org/10.1146/annurev-anthro-102313-030223>
- Bard, K. A., Street, E. A., McCrary, C. & Boothe, R. G. (1995) Development of visual acuity in infant chimpanzees. *Infant Behaviour and Development*, 18(2), 225-232.
- Baron-Cohen, S. (1995). *Learning, development, and conceptual change. Mindblindness: An essay on autism and theory of mind*. Cambridge, MA, US: The MIT Press.
- Baron-Cohen, S. (1999) The evolution of a theory of mind. In M. Corballis & S. E. G. Lea (eds.) *The descent of mind: Psychological perspectives on hominid evolution*. (p.261- 277). Oxford: Oxford University Press.
- Barrera, G. & Bentosela, M. (2016) Interspecific social skills in a highly socialized Aguará Guazú (maned wolf, *Chrysocyon brachyurus*). A pilot subject study. *Current Ethology*, 15(1), 55-63.
- Bates, E., Camaioni, L. & Volterra, V. (1975) The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, 21(3), 205-226.
- Barth, J., Reaux, J. E., & Povinelli, D. J. (2005). Chimpanzees' (*Pan troglodytes*) use of gaze cues in object-choice tasks: different methods yield different results. *Animal Cognition*, 8, 84-92. <https://doi.org/10.1007/s10071-004-0235-x>
- Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8(6), 492-499. <https://doi.org/10.1111/j.1467-7687.2005.00440.x>

- Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, 30, 359-375. <https://doi.org/10.1111/j.2044-835X.2011.02043.x>
- Bernabei, G., Camaign, L., & Levi, G. (1998). An evaluation of early development in children with Autism and Pervasive Developmental Disorders from home movies: Preliminary Findings. *Autism*, 2(3), 243-258. <https://doi.org/10.1177/1362361398023003>
- Bhattacharjee, D., Nikhil, D. N., Gupta, S., Sau, S., Sarkar, R., Biswas, A., Banarjee, A., Babu, D., Mehta, D. & Bhadra, A. (2017). Free-ranging dogs show age-related plasticity in their ability to follow human pointing. *PLoS ONE*, 12(7), 1-17. <https://doi.org/10.1371/journal.pone.0180643>
- Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making inferences about the location of hidden food: Social dog, causal ape. *Journal of Comparative Psychology*, 120(1), 38-47. <https://doi.org/10.1037/0735-7036.120.1.38>
- Burkart, J., & Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, 120(2), 120-130. <http://dx.doi.org/10.1037/0735-7036.120.2.120>
- Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of mind: evolution, development and simulation of everyday mindreading* (pp. 223–232). Cambridge, MA: Blackwell.
- Butterworth, G. (2001). Joint visual attention in infancy. In J. G. Bremner & A. Fogel (Eds.), *Blackwell Handbook of Infant Development* (pp. 213-240). Hove: Blackwell.

- Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Mahwah: Erlbaum
- Butterworth, G., & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (Ed.) *Thought without Language* (pp. 5- 24) New York: Clarendon Press.
- Butterworth, G. & Morissette, P. (1996) Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology*, 14(3), 219-231. <https://doi.org/10.1080/02646839608404519>
- Byrnit, J. T. (2004). Nonenculturated orangutans' (*Pongo Pygmaeus*) use of experimenter-given manual and facial cues in an object-choice task. *Journal of Comparative Psychology*, 118(3), 309-315. <http://dx.doi.org/10.1037/0735-7036.118.3.309>
- Byrnit, J. T. (2009). Gorillas' (*Gorilla gorilla*) use of experimenter-given manual and facial cues in an object-choice task. *Animal Cognition*, 12, 401-404. <https://doi.org/10.1007/s10071-008-0200-1>
- Byrnit, J. T. (2015). Primates' socio-cognitive abilities: What kinds of comparisons make sense? *Integrative Psychological and Behavioural Science*, 49, 485-511. <https://doi.org/10.1007/s12124-015-9312-8>
- Call, J., Agnetta, B., & Tomasello, M. (2000). Cues that chimpanzees do and do not use to find hidden objects. *Animal Cognition*, 3, 23-24. <https://doi.org/10.1007/s100710050047>

- Call, J., Hare, B. A., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition*, 1, 89-99. <https://doi.org/10.1007/s100710050013>
- Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, 108(4), 307-317. <http://dx.doi.org/10.1037/0735-7036.108.4.307>
- Call, J. & Tomasello, M. (1999) A nonverbal false belief task: The performance of human children and great apes. *Child Development*, 70(2), 381- 395. <https://doi.org/10.1111/1467-8624.00028>
- Carballo, F., Freidin, E., Casanave, E. & Bentosela, M. (2016). Following human pointing: Where do dogs (*Canis familiaris*) look at to find food? *Behavioural Processes*, 128, 37-40. <https://doi.org/10.1016/j.beproc.2016.04.004>
- Chang, H., Shaw, D. S. & Cheung, J. (2015) The development of emotional and behavioural control in early childhood: Heterotypic continuity and relations to early school adjustment. *Journal of Child and Adolescent Behaviour* 3(3), 204. <https://doi.org/10.4172/2375-4494.1000204>
- Clark, H., Elsherif, M. M. & Leavens, D. A. (2019) Ontogeny vs. phylogeny in primate/ canid comparisons: A meta-analysis of the object-choice task. *Neuroscience and Biobehavioural Reviews*. Advance online publication. <https://dx.doi.org/10.1016/j.neubiorev.2019.06.001>
- Cochet, H. & Vauclair, J. (2010) Pointing gestures produced by toddlers from 15 to 30 months: Different functions, hand shapes and laterality patterns. *Infant Behaviour and Development*, 33, 431-441. <https://doi.org/10.1016/j.infbeh.2010.04.009>

- Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology*, 26, 561-577. <https://doi.org/10.1348/026151008X285660>
- Corballis, M. C. (1999) The gestural origins of language: Human language may have evolved from manual gestures, which survive today as a “behavioural fossil” coupled to speech. *American Scientist*, 87(2), 138-145.
<https://doi.org/10.1511/1999.20.810>
- Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Cooperrider, K., Slotta, J., & Núñez, R. (2018). Cooperrider, K., Slotta, J., & Núñez, R. (2018). The preference for pointing with the hand is not universal. *Cognitive Science*, 42(4), 1375-1390. <https://doi.org/10.1111/cogs.12585>
- Dalla Costa, E., Cannas, S., Minero, M., & Palestini, C. (2010). Influence of age in understanding human gestures in pet dogs. *Journal of Veterinary Behavior: Clinical Applications and Research*, 5(1), 40.
<https://doi.org/10.1016/j.jveb.2009.10.009>
- D’Aniello, B. Alterisio, A., Scandurra, A., Petremolo, E., Iommelli, M. R. & Aria, M. (2017) What’s the point? Golden and Labrador retrievers living in kennels do not understand human pointing gestures. *Animal Cognition*, 20, 777-787.
<https://doi.org/10.1007/s10071-017-1098-2>
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A. & Liaw, J. (2004) Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40(2), 271-283.
<https://doi.org/10.1037/0012-1649.40.2.271>.

- Dorey, N. R., Udell, M. A. R., & Wynne, C. D. L. (2010). When do domestic dogs, *Canis familiaris*, start to understand human pointing? The role of ontogeny in the development of interspecies communication. *Animal Cognition*, 79, 37- 41. <https://doi.org/10.1016/j.anbehav.2009.09.032>
- Duranton, C. & Gaunet, F. (2016) Effects of shelter housing on dogs' sensitivity to human social cues. *Journal of Veterinary Behaviour*, 14, 20-27. <https://doi.org/10.1016/j.jveb.2016.06.011>.
- Enfield, N.J. (2001)., "Lip-pointing": A discussion of form and function with reference to data from Laos. *Gesture*, 1, 185-212. <https://doi.org/10.1075/gest.1.2.06enf>.
- Essler, J.L., Schwartz, L.P., Rossettie, M. S. & Judge, P. G. (2017) Capuchin monkeys' use of human and conspecific cues to solve a hidden object-task. *Animal Cognition*, 20, 985-998. <https://doi.org/10.1007/s10071-017-1118-2>.
- Flack, Z., & Leavens, D. A. (2018) Communication and developmental milestones. In T. K. Shackelford & V. E. Weekes-Shackelford (Eds.), *Encyclopedia of evolutionary psychological science*. New York: Springer.
- Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, 23, 307-336. <https://doi.org/10.1017/S0305000900008813>
- Gácsi, M. Györi, B., Virányi, Z., Kubinyi, E., Range, F., Belényi, B. & Miklósi, A. (2009) Explaining dog wolf differences in utilising human pointing gestures: Selection for synergistic shifts in the development of some social skills. *PLoS ONE*, 4(8), 1-6. <https://doi.org/10.1371/journal.pone.0006584>

- Gácsi, M., Kara, E., Belényi, B., Topál, J. & Miklósi, A. (2009) The effect of development and individual differences in pointing comprehension in dogs. *Animal Cognition*, 12(3), 471-479. <https://doi.org/10.1007/s10071-008-0208-6>
- Gácsi, M., McGreevy, P., Kara, E., & Miklósi, A. (2009). Effects of selection for attention and cooperation in dogs. *Behavioural and Brain Functions*, 5, 31-38. <https://doi.org/10.1186/1744-9081-5-31>
- Hall, N. J., Udell, M. A. R., Dorey, N. R., Walsh, A. L. & Wynne, C. L. (2011) Megachiropteran bats (*Pteropus*) utilize human referential stimuli to locate hidden food. *Journal of Comparative Psychology*, 125(3), 341- 346. <https://doi.org/10.1037/a0023680>
- Hare, B. (2001) Can competitive paradigms increase the validity of experiments on primate social cognition? *Animal Cognition*, 4, 269-280. <https://doi.org/10.1007/s100710100084>
- 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>
- Hare, B., & Tomasello, M. (1999). Domestic dogs (*Canis familiaris*) use human and conspecific social cues to locate hidden food. *Journal of Comparative Psychology*, 113(2), 173-177. <http://dx.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. *Animal Behaviour*, 68, 571-581. <http://dx.doi.org/10.1016/j.anbehav.2003.11.011>

- Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *TRENDS in Cognitive Sciences*, 9(9), 439-444. <https://doi.org/10.1016/j.tics.2005.07.003>
- Hare, B., Plyusnina, I., Ignacio, N., Schepina, O., Stepika, A., Wrangham, R. & Trut, L. (2005) Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. *Current Biology*, 15, 226-230.
<https://doi.org/10.1016/j.cub.2005.01.040>
- Hattori, Y., Kuroshima, H., & Fujita, K. (2007). I know you are not looking at me: capuchin monkeys' (*Cebus apella*) sensitivity to human attentional states. *Animal Cognition*, 10, 141-148. <https://doi.org/10.1007/s10071-006-0049-0>
- Hegedüs, D., Bálint, A., Miklósi, A. & Pongrácz, P. (2013) Owners fail to influence the choices of dogs in a two-choice, visual pointing task. *Behaviour*, 150, 427-443.
<https://doi.org/10.1163/1568539X-00003060>
- Henrich, J., Heine, S. J. & Norenzaya, A. (2010) The weirdest people in the world? *Behavioural and Brain Sciences*, 33, 61-135.
<https://doi.org/10.1017/S0140525X0999152X>
- Hernádi, A., Kis, A., Turcsán, B. & Topál, J. (2012) Man's underground best friend: Domestic ferrets, unlike the wild forms, show evidence of dog-like socio-cognitive skills. *PLoS ONE*, 7(8), 1-6.
<https://doi.org/10.1371/journal.pone.0043267>
- Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317, 1360-1367.
<https://doi.org/10.1126/science.1146282>

- Herrmann, E., Melis, A. P., & Tomasello, M. (2005). Apes' use of iconic cues in the object-choice task. *Animal Cognition*, 9, 118-130.
<https://doi.org/10.1007/s10071-005-0013-4>
- Hill, A., Collier-Baker, E., & Suddendorf, T. (2011). Inferential reasoning by exclusion in great apes, lesser apes and spider monkeys. *Journal of Comparative Psychology*, 125(1), 91-103. <http://dx.doi.org/10.1037/a0020867>
- Hopkins, W. D., Russell, J. L., McIntyre, J., & Leavens, D. A. (2013). Are chimpanzees really so poor at understanding imperative pointing? Some new data and an alternative view of canine and ape social cognition. *PLoS ONE*, 8, 1-6.
<https://doi.org/10.1371/journal.pone.0079338>
- Inoue, Y., Inoue, E., & Itakura, S. (2004). Use of experimenter given cues by gibbon Use of experimenter-given directional cues by a young white-handed gibbon (*Hylobates lar*). *Japanese Psychological Research*, 46(3), 262-267.
<https://doi.org/10.1111/j.1468-5584.2004.00258.x>
- Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (*Pan troglodytes*), an orangutan (*Pongo pygmaeus*) and human infants (*Homo sapiens*). *Journal of Comparative Psychology*, 112(2), 119-126. <https://psycnet.apa.org/doi/10.1037/0735-7036.112.2.119>
- Itakura, S., Agnetta, B., Hare, B., & Tomasello, M. (1999). Chimpanzee use of human and conspecific cues to locate hidden food. *Developmental Science*, 2(4), 448-456. <https://doi.org/10.1111/1467-7687.00089>
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16(5), 367-371.
<https://doi.org/10.1111/j.0956-7976.2005.01542.x>

- Kaminski, J. & Nitzchner, M. (2013). Do dogs get the point? A review of dog-human communication ability. *Learning and Motivation*, 44(4), 294-302.
<https://doi.org/10.1016/j.lmot.2013.05.001>
- Kaminski, J., Riedel, J., Call, J. & Tomasello, M. (2005) Domestic goats, *Capra hircus*, follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, 69, 11-18. <https://doi.org/10.1016/j.anbehav.2004.05.008>
- Kaminski, J., Schulz, L. & Tomasello, M. (2011) How dogs know when communication is intended for them. *Developmental Science*, 15(2), 222-232.
<https://doi.org/10.1111/j.1467-7687.2011.01120.x>.
- Kirchhofer, K. C., Zimmermann, F., Kaminski, J., & Tomasello, M. (2012). Dogs (*Canis familiaris*), but not chimpanzees (*Pan troglodytes*), understand imperative pointing. *PLoS ONE*, 8(2), 1-7.
<https://doi.org/10.1371/journal.pone.0030913>
- Kopp, C. B. (1982) Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*, 18(2), 199-214. <https://doi.org/10.1037/0012-1649.18.2.199>
- Kraus, C., van Waveren, C. & Huebner, F. (2014) Distractible dogs, constant cats? A test of the distraction hypothesis in two domestic species. *Animal Behaviour*, 93, 173-181. <https://doi.org/10.1016/j.anbehav.2014.04.026>
- Krause, M. A., Udell, M. A. R., Leavens, D. A., & Skopos, L. (2018). Animal pointing: Changing trends and findings from 30 years of research. *Journal of Comparative Psychology*, 132, 326-345. <https://doi.org/10.1037/com0000125>

- Lakatos, G., Dóka, A. & Miklósi, A. (2007). The role of visual clues in the comprehension of the human pointing signals in dogs. *International Journal of Comparative Psychology*, 20, 341-350.
- Lakatos, G., Soproni, K., Dóka, A. & Miklósi, A. (2009). A comparative approach to dogs' (*Canis familiaris*) and human infants' comprehension of various forms of pointing gestures. *Animal Cognition*, 12, 621-631.
<https://doi.org/10.1007/s10071-009-0221-4>
- Lazarowski, L., & Dorman, D. C. (2015). A comparison of pet and purpose-bred research dog (*Canis Familiaris*) performance on human-guided object-choice tasks. *Behavioural Processes*, 110, 60-67.
<https://doi.org/10.1016/j.beproc.2014.09.021>
- Lea, S. E. G. & Osthaus, B. (2018) In what sense are dogs special? Canine cognition in comparative context. *Learning and Behaviour*, 46(4), 335-363.
<https://doi.org/10.3758/s13420-018-0349-7>
- Leavens, D. A. (2012) Joint attention: Twelve myths. In A. Seeman (ed), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp.43-72) Cambridge, Massachusetts: MIT Press.
- Leavens, D. A. (2014). The plight of the sense-making ape. In M. Cappuccio & T. Froese (Eds.), *Enactive cognition at the edge of sense-making* (pp. 81-104). Basingstoke, U.K.: Palgrave Macmillan.
- Leavens, D. A. (2018) The cognitive implications of intentional communication: A multi-faceted mirror. In: Di Paolo L., Di Vincenzo F., De Petrillo F. (eds) *Evolution of Primate Social Cognition. Interdisciplinary Evolution Research*. Cham: Springer.

- Leavens, D. A., & Bard, K. A. (2011) Environmental influences on joint attention in great apes: Implications for human cognition. *Journal of Cognitive Education and Psychology*, 10, 9-31. <http://dx.doi.org/10.1891/1945-8959.10.1.9>
- Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2010) BIZARRE chimpanzees do not represent “the chimpanzee”. *Behavioural and Brain Sciences*, 33, 61-135. <https://doi.org/10.1017/S0140525X10000166>
- Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2017). The mismeasure of ape social cognition. *Animal Cognition*, 1-18. <https://doi.org/10.1007/s10071-017-1119-1>
- Leavens, D. A., & Clark, H. (2017) Joint attention. In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior*. New York: Springer.
- Leavens, D. A. & Hopkins, W. D. (1999) The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology*, 113(4), 417- 425. <http://dx.doi.org/10.1037/0735-7036.113.4.417>
- Leavens, D. A., Hopkins, W. D. & Bard, K. A. (2005) Understanding the point of chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in Psychological Science*, 14(4), 185-189. <https://doi.org/10.1111/j.0963-7214.2005.00361.x>
- Leavens, D. A. & Racine, T. P. (2009) Joint attention in apes and humans: are humans unique? *Journal of Consciousness Studies*, 16(6-8), 240-267.
- Leavens, D. A., Reamer, L. A., Marenco, M. C., Russell, J. L., Wilson, D., Schapiro, S. J. & Hopkins, W. D. (2015) Distal communication by chimpanzees (*Pan troglodytes*): evidence for common ground? *Child Development*, 86(5), 1623-1638. <https://doi.org/10.1111/cdev.12404>
- Leung, E. H. L. & Rheingold, H. L. (1981) Development of pointing as a social gesture.

Developmental Psychology, 17(2), 215-220. <http://dx.doi.org/10.1037/0012-1649.17.2.215>

Liszkowski, U., Schäfer, M., Carpenter, M. & Tomasello, M. (2009) Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*, 20(5), 654- 660. <https://doi.org/10.1111/j.1467-9280.2009.02346.x>

Lyn, H. (2010). Environment, methodology, and the object choice task in apes: Evidence for declarative comprehension and implications for the evolution of language. *Journal of Evolutionary Psychology*, 8(4), 333-349. <https://doi.org/10.1556/JEP.8.2010.4.3>

Lyn, H., Russell, J. L., & Hopkins, W. D. (2010). The impact of environment on the comprehension of declarative gestures in apes. *Psychological Science*, 21(3), 360-365. <https://doi.org/10.1177/0956797610362218>

Lyn, H., Russell, J. L., Leavens, D. A., Bard, K. A., Boysen, S. T., Schaeffer, J. A. & Hopkins, W. D. (2014) Apes communicate about absent and displaced objects: methodology matters. *Animal Cognition*, 17(1), 85-94. <https://doi.org/10.1007/s10071-013-0640-0>

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

Maclean, E. L., Herrmann, E., Suchindran, S. & Hare, B. (2017). Individual differences in cooperative communicative skills are more similar between dogs and humans than chimpanzees. *Animal Behaviour*, 126, 41-51. <https://doi.org/10.1016/j.anbehav.2017.01.005>

- Maclean, E. L., Krupenye, C. & Hare, B. (2014). Dogs (*Canis familiaris*) account for body orientation but not visual barriers when responding to pointing gestures. *Journal of Comparative Psychology*, 128(3), 285-297.
<https://doi.org/10.1037/a0035742>
- Maros, K., Gácsi, M. & Miklósi, A. (2008) Comprehension of human pointing gestures in horses (*Equus caballus*). *Animal Cognition*, 11, 457- 466.
<https://doi.org/10.1007/s10071-008-0136-5>
- Marsh, H. L. (2012). Orangutans' use of contiguous versus distal social and non-social cues in an object-choice task. *International Journal of Comparative Psychology*, 25, 299-308. <https://doi.org/10.5811/westjem.2011.5.6700>
- Masataka, N. (2003) From index-finger extension to index-finger pointing: The ontogenesis of pointing in preverbal infants. In S. Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 9-33). Mahwah, NJ: Lawrence Erlbaum Associates.
- McKinley, J., & Sambrook, T. D. (2000). Use of human-given cues by domestic dogs (*Canis familiaris*) and horses (*Equus caballus*). *Animal Cognition*, 3, 13-22.
<https://doi.org/10.1007/s100710050046>
- Miklósi, A., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z. & Csányi, V. (2003). A simple reason for a big difference: Wolves do not look back at humans, but dogs do. *Current Biology*, 13, 763- 766.
- Miklósi, A., Polgárdi, R., Topál, J. & Csányi, V. (1998) Use of experimenter-given cues in dogs. *Animal Cognition*, 1, 113-121. <https://doi.org/10.1007/s100710050016>

- Miklósi, A., Pongrácz, P., Lakatos, G., Topál, J. & Csányi, V. (2005) A comparative study of the use of visual communicative signals in interactions between dogs (*Canis familiaris*) and humans and cats (*Felis catus*) and humans. *Journal of Comparative Psychology*, 119(2), 179-186. <http://dx.doi.org/10.1037/0735-7036.119.2.179>
- Miklósi, A., & Soproni, K. (2006). A comparative analysis of animals' understanding of the human pointing gesture. *Animal Cognition*, 9(2), 81-93.
<https://doi.org/10.1007/s10071-005-0008-1>
- Miller, P. E. & Murphy, C. J. (1995). Vision in dogs. *Journal of the American Medical Veterinary Association*, 207(12), 1623-1634.
- Moll, H., & Tomasello, M. (2007). Cooperation and human cognition; The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society*, 362, 639-648. <https://doi.org/10.1098/rstb.2006.2000>
- Moore, R., Call, J., & Tomasello, M. (2015). Production and comprehension of gestures between orang-utans (*Pongo pygmaeus*) in a referential communication game. *PLoS ONE*, 10(6), e0129726.
<https://doi.org/10.1371/journal.pone.0129726>
- Moore, C. & Corkum, V. (1994) Social understanding at the end of the first year of life. *Developmental Review*, 14, 349- 372. <http://dx.doi.org/10.1006/drev.1994.1014>
- Mulcahy, N. J., & Call, J. (2009). The performance of bonobos (*Pan paniscus*), chimpanzees (*pan troglodytes*) and orangutans (*Pongo pygmaeus*) in two versions of an object-choice task. *Journal of Comparative Psychology*, 123(3), 304-309. <http://dx.doi.org/10.1037/a0016222>

- Mulcahy, N. J., & Hedge, V. (2012). Are great apes tested with an abject object-choice task? *Animal Behaviour*, 83, 313-321.
<http://dx.doi.org/10.1016/j.anbehav.2011.11.019>
- Mulcahy, N. J., & Suddendorf, T. (2011). An obedient orangutan (*Pongo abelii*) performs perfectly in peripheral object-choice tasks but fails the standard centrally presented versions. *Journal of Comparative Psychology*, 125(1), 112-115. <http://dx.doi.org/10.1037/a0020905>
- Mundy, P., Sigman, M., Ungerer, J. & Sherman, T. (1987) Nonverbal communication and play correlates in the language development of autistic children. *Journal of Autism and Developmental Disorders*, 17(3), 349-364.
<http://dx.doi.org/10.1007/BF01487065>
- Mundy, P. (2013) *RJA/ IJA (Initiating/ responding to joint attention)*. In F. R. Volkmar (eds.) *Encyclopaedia of Autism Spectrum Disorders*. Springer: Cham.
<https://doi.org/10.1007/978-1-4419-1698-3>.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V. & Parlade, M. V. (2007) Individual differences and the development of joint attention in infancy. *Child Development*, 78(3), 938-954. doi:10.1111/j.1467-8624.2007.01042.x.
- Nakajima, S., Fukuoka, T., Takamatsu, Y., & Chin, H. (2009). Do dogs follow human social cues? A replication with family dogs in Japan. *The Japanese Journal of Animal Psychology*, 59(1), 77- 84. <https://doi.org/10.2502/janip.59.1.6>
- Nawroth, C., Ebersbach, M. & von Borrell, E. (2014) Juvenile domestic pigs (*Sus scrofa domestica*) use human-given cues in an object choice task. *Animal Cognition*, 17(3), 701-713. <http://dx.doi.org/10.1007/s10071-013-0702-3>

- Neiworth, J. J., Burman, M. A., Basile, B. M., & Lickteig, M. T. (2002). Use of experimenter-given cues in visual co-orienting and in an object-choice task by a New World monkey species, cotton top tamarins (*Saguinus Oedipus*). *Journal of Comparative Psychology*, 116(1), 3-11. <http://dx.doi.org/10.1037/0735-7036.116.1.3>
- Okamoto, S., Tomonaga, M., Iishi, K., Kawai, N., Tanaka, M. & Matsuzawa, T. (2002). An infant chimpanzee (*Pan troglodytes*) follows human gaze. *Animal Cognition*, 5, 107-114. <http://dx.doi.org/10.1007/s10071-002-0138-7>.
- Okamoto-Barth, S., Tomonaga, M., Tanaka, M. & Matsuzawa, T. (2008). Development of using experimenter-given cues in infant chimpanzees: longitudinal changes in behaviour and cognitive development. *Developmental Science*, 11(1), 98-108. <https://doi.org/10.1111/j.1467-7687.2007.00658.x>
- Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, 24(3), 247-257. <https://doi.org/10.1007/BF02172225>
- Pack, A. A. & Herman, L. M. (2004) Bottlenose dolphins (*Tursiops truncatus*) comprehend the referent of both static and dynamic human gazing and pointing in an object-choice task. *Journal of Comparative Psychology*, 118(2), 160-171. <http://dx.doi.org/10.1037/0735-7036.118.2.160>
- Pedersen, J., Segerdahl, P., & Fields, W. M. (2009). Why apes point: pointing gestures in spontaneous conversation of language-competent *Pan/Homo* bonobos. In E. Potocki & J. Krasinski (eds), *Primateology: Theories, methods, and research* (pp. 53-74). Hauppauge, N.Y.: Nova.

- Peignot, P., & Anderson, J. R. (1999). Use of experimenter-given manual and facial cues by Gorillas (*Gorilla gorilla*) in an object-choice task. *Journal of Comparative Psychology*, 113(3), 253-260. <http://doi.org/10.1037//0735-7036.113.3.253>.
- Petitto, L. (1988) “*Language*” in the prelinguistic child. In E. Kessel (eds) Development of language and language researchers. (p. 187- 222). Hillsdale, NJ: Erlbaum.
- Pettersson, H., Kaminski, J., Herrmann, E., & Tomasello, M. (2011). Understanding of human communicative motives in domestic dogs. *Applied Animal Behaviour Science*, 133, 235-245. <https://doi.org/10.1016/j.applanim.2011.05.008>
- Pflandler, E., Lakatos, G., & Miklósi, A. (2013). Eighteen-month-old human infants show intensive development in comprehension of different types of pointing gestures. *Animal Cognition*, 16, 711-719. <https://doi.org/10.1007/s10071-013-0606-2>
- Plaude, V. & Fiset, S. (2013) Pointing gestures modulate domestic dogs’ search behaviour for hidden objects in a spatial rotation problem. *Learning and Motivation*, 44, 282-293. <https://doi.org/10.1016/j.lmot.2013.03.004>
- Plotnik, J. M., Pokorny, J. J., Keratimanochaya, T., Webb, C., Beronja, H. F., Hennessy, A., Hill, J., Hill, V. J., Kiss, R., Maguire, C., Melville, B. L., Morrison, V. M. B., Seecoomar, D., Singer, B., Ukehaxhaj, J., Vlahakis, S. K., Ylli, D., Clayton, N. S., Roberts, J., Fure, E. L., Duchatelier, A. P., & Getz, D. (2013). Visual cues given by humans are not sufficient for Asian elephants (*Elephas maximus*) to find hidden food. *PLoS ONE*, 8(4), 1-7. <https://doi.org/10.1371/journal.pone.0061174>

- Pongrácz, P., Gácsi, M., Hegedüs, D., Péter, A., & Miklósi, A. (2013). Test sensitivity is important for detecting variability in pointing comprehension in canines. *Animal Cognition*, 16, 721-735. <https://doi.org/10.1007/s10071-013-0607-1>
- Povinelli, D. J., Bierschwale, D. T., & Čech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, 17, 37- 60. <https://doi.org/10.1348/026151099165140>
- Povinelli, D. J. & Eddy, T. J. (1996) What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development*, 247(61), 1-152. <https://doi.org/10.2307/1166159>
- Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 104(3), 203- 210. <http://dx.doi.org/10.1037/0735-7036.104.3.203>
- Povinelli, D. J., Parks, K. A., & Novak, M. A. (1991). Do Rhesus monkeys (*Macaca Mulatta*) attribute knowledge and ignorance to others? *Journal of Comparative Psychology*, 105(4), 318- 325. <http://dx.doi.org/10.1037/0735-7036.105.4.318>
- Povinelli, D. J., Reaux, J. E., Bierschwale, D. T., Allain, A. D., & Simon, B. B. (1997). Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees. *Cognitive Development*, 12, 423-461. [https://doi.org/10.1016/S0885-2014\(97\)90017-4](https://doi.org/10.1016/S0885-2014(97)90017-4)
- Proops, L., Rayner, J., Taylor, A. M. & McComb, K. (2013) The responses of young domestic horses to human-given cues. *PLoS ONE*, 8(6), e67000. <https://doi.org/10.1371/journal.pone.0067000>

- Racine, T. P., Leavens, D. A., Susswein, N., & Wereha, T. J. (2008). Conceptual and methodological issues in the investigation of primate intersubjectivity. In F. Morganti, A. Carassa & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions* (pp. 65-79). Amsterdam: IOS Press.
- Riedel, J., Schumann, K., Kaminski, J., Call, J., & Tomasello, M. (2008). The early ontogeny of human-dog communication. *Animal Behaviour*, 75(3), 1003-1014.
<https://doi.org/10.1016/j.anbehav.2007.08.010>
- Rumbaugh, D. M., Washburn, D. A. & Pate, J. L. (1984) Discrimination learning set and transfer. In G. Greenburg & M. M. Haraway (Eds.), *Comparative Psychology: A Handbook* (pp. 562- 565), New York: Garland Publishing Inc.
- Russell, J. L., Lyn, H., Schaeffer, J. A., & Hopkins, W. D. (2011). The role of socio-communicative rearing environments in the development of social and physical cognition in apes. *Developmental Science*, 14, 1459–70.
<https://doi.org/10.1111/j.1467-7687.2011.01090.x>
- Scheel, M. H., Shaw, H. L., & Gardner, R. A. (2016). Incomparable methods vitiate cross-species comparisons: A comment on Haun, Rekers, and Tomasello (2014). *Psychological Science*, 27, 1667-1669.
<https://doi.org/10.1177/0956797615595229>
- Scheumann, M. & Call, J. (2004) The use of experimenter-given cues by South African fur seals (*Arctocephalus pusillus*). *Animal Cognition*, 7, 224-230.
<http://dx.doi.org/10.1007/s10071-004-0216-0>

- Schmidjell, T., Range, F., Huber, L. & Virányi, Z. (2012) Do owners have a Clever Hans effect on dogs? Results of a pointing study. *Frontiers in Psychology*, 3, 1-15. <https://doi.org/10.3389/fpsyg.2012.00558>.
- Schmitt, V., Schloegl, C. & Fischer, J. (2014) Seeing the experimenter influences the response to pointing cues in long-tailed macaques. *PLoS ONE*, 9(3), 1-6. <https://doi.org/10.1371/journal.pone.0091348>
- Shearin, A. L., & Ostrander, E. A. (2010). Canine morphology: Hunting for genes and tracking mutations. *PLoS Biology*, 8(3), e1000310. <https://doi.org/10.1371/journal.pbio.1000310>
- Smet, A. F., & Byrne, R. W. (2013). African elephants can use human pointing cues to find hidden food. *Current Biology*, 23(20), 2033-2037. <https://doi.org/10.1016/j.cub.2013.08.037>
- Takaoka, A., Maeda, T., Hori, Y. & Fujita, K. (2015) Do dogs follow behavioural cues from an unreliable human? *Animal Cognition*, 18, 475-483. <https://doi.org/10.1007/s10071-014-0816-2>
- Tan, J., Tao, R. & Su, Y. (2014) Testing the cognition of the forgotten Colobines: A first look at Golden Snub-Nosed monkeys (*Rhinopithecus roxellana*). *International Journal of Primatology*, 35, 376-393. <https://doi.org/10.1007/s10764-013-9741-5>
- Tomasello, M. (1995) Joint attention as social cognition. In C. Moore & P. J. Dunham (eds.) *Joint Attention: Its Origins and Role in Development* (p. 103-130). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

- Tomasello, M. (1999) *The cultural origins of human cognition*. Cambridge, Massachusetts: Harvard University Press.
- Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, 68(6), 1067- 1080. <https://doi.org/10.1111/j.1467-8624.1997.tb01985.x>
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10(1), 121-125. <https://doi.org/10.1111/j.1467-7687.2007.00573.x>
- Tomasello, M. & Moll, H. (2010) The gap is social: Human shared intentionality and culture. In P. Kappeler & J. Silk (Eds.), *Mind the gap: Tracing the origins of human universals* (pp. 331-349). Berlin: Springer.
- Triesch, J., Teuscher, C., Deák, G. O. & Carlson, E. (2006) Gaze following: Why (not) learn it? *Developmental Science*, 9(2), 125- 157. <https://doi.org/10.1111/j.1467-7687.2006.00470.x>
- Udell, M. A. R., Dorey, N. R., & Wynne, C. D. L. (2008). Wolves outperform dogs in following human social cues. *Animal Behaviour*, 76, 1767-1773. <https://doi.org/10.1016/j.anbehav.2008.07.028>
- Udell, M. A. R., Dorey, N. R., & Wynne, C. D. L. (2010). The performance of stray dogs (*Canis familiaris*) living in a shelter on human-guided object-choice tasks. *Animal Behaviour*, 79, 717-725. <https://doi.org/10.1016/j.anbehav.2009.12.027>
- Udell, M. A. R., Dorey, N. R. & Wynne, C. D. L. (2010b). What did domestication do to dogs? A new account of dogs' sensitivity to human actions. *Biological Reviews*, 85, 327- 345. <https://doi.org/10.1111/j.1469-185X.2009.00104.x>

- Udell, M. A. R., Dorey, N. R. & Wynne, C. D. L. (2012) Inter-species social learning in dogs: The inextricable roles of phylogeny and ontogeny. In T. R. Zentall & E. A. Wasserman (eds) *The Oxford Handbook of Comparative Cognition* (2nd ed. pp. 819-831). New York, N.Y.: Oxford University Press.
- Udell, M. A. R., Ewald, M., Dorey, N. R., & Wynne, C. D. L. (2014). Exploring breed differences in dogs (*Canis lupus familiaris*): Does exaggeration or inhibition of predatory response predict performance on human-guided tasks? *Animal Behaviour*, 89, 99-105. <https://doi.org/10.1016/j.anbehav.2013.12.012>
- Udell, M. A. R., Giglio, R. F., & Wynne, C. D. L. (2008). Domestic dogs (*Canis familiaris*) use human gestures but not nonhuman tokens to find hidden food. *Journal of Comparative Psychology*, 122(1), 84-93.
<http://dx.doi.org/10.1037/0735-7036.122.1.84>
- Udell, M. A. R., Hall, N. J., Morrison, J., Dorey, N. R. & Wynne, C. D. L. (2013). Point topography and within-session learning are important predictors of pet dogs' (*Canis lupus familiaris*) performance on human-guided tasks. *Revista Argentina de Ciencias de Comportamiento*, 5(2), 3-20.
- Udell, M. A. R., Spencer, J. M., Dorey, N. R., & Wynne, C. D. L. (2012). Human-socialized wolves follow diverse human gestures...and they may not be alone. *International Journal of Comparative Psychology*, 25, 97-117.
- Udell, M. A. R. & Wynne, C. D. L. (2010) Ontogeny and phylogeny: Both are essential to human-sensitive behaviour in the genus *Canis*. *Animal Behaviour*, 79, e9-e14.
<https://doi.org/10.1016/j.anbehav.2009.11.033>

- van der Goot, M. H., Tomasello, M. & Liszkowski, U. (2014) Differences in the nonverbal requests of great apes and human infants. *Child Development*, 85(2), 444-455. <https://doi.org/10.1111/cdev.12141>
- van Schaik, C. P. & Burkart, J. M. (2011) Social learning and evolution: The cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society B.*, 366, 1008-1016. <https://doi.org/10.1098/rstb.2010.0304>
- Vick, S., & Anderson, J. R. (2000). Learning and limits of use of eye gaze by capuchin monkeys (*Cebus apella*) in an object-choice task. *Journal of Comparative Psychology*, 114(2), 200-207. <http://dx.doi.org/10.1037/0735-7036.114.2.200>
- Vick, S., & Anderson, J. R. (2003). Use of human visual attention cues by olive baboons (*Papio anubis*) in a competitive task. *Journal of Comparative Psychology*, 117(2), 209-216. <http://dx.doi.org/10.1037/0735-7036.117.2.209>
- Virányi, Z., Gácsi, M., Kubinyi, E., Topál, J., Belényi, B., Ujfalussy, D., & Miklósi, A. (2008). Comprehension of human pointing gestures in young human-reared wolves (*Canis lupus*) and dogs (*Canis familiaris*). *Animal Cognition*, 11, 373-387. <https://doi.org/10.1007/s10071-007-0127-y>
- Whiten, A. & van Schaik, C. P. (2007) The evolution of animal ‘cultures’ and social intelligence. *Philosophical Transactions of the Royal Society B.*, 362, 603-620. <https://doi.org/10.1098/rstb.2006.1998>
- Wilkins, D. (2003) Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: Where Language, Culture, and Cognition Meet* (pp. 171–215) Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wobber, V., Hare, B., Koler-Matznick, J., Wrangham, R. & Tomasello, M. (2009) Breed differences in domestic dogs’ (*Canis familiaris*) comprehension of human

communicative signals. *Interaction Studies*, 10(2), 206-224.

<https://doi.org/10.1075/is.10.2.06wob>

Wynne, C. D. L., Udell, M. A. R. & Lord, K. A. (2008). Ontogeny's impact of human-dog communication. *Animal Behaviour*, 76, e1-e4.

<http://dx.doi.org/10.1016/j.anbehav.2008.03.010>

Zaine, I., Domeniconi, C. & Wynne, C. D. L. (2015) The ontogeny of human point following in dogs: When younger dogs outperform older. *Behavioural Processes*, 119, 76-85. <https://doi.org/10.1016/j.beproc.2015.07.004>

Zlatev, J., Madsen, E. A., Lenninger, S., Persson, T., Sayehli, S., Sonesson, G., & van de Weijer, J. (2013). Understanding communicative intentions and semiotic vehicles by children and chimpanzees. *Cognitive Development*, 28, 312- 329. <https://doi.org/10.1016/j.cogdev.2013.05.001>