https://doi.org/10.35662/unine-thesis-2893

Social learning and flexibility in vocal communication of wild vervet monkeys

PhD Thesis
Institute of Biology, Faculty of Science
University of Neuchâtel

By

Adwait Deshpande

Defended on the 19th May 2021

Thesis committee:

Prof Klaus Zuberbühler, University of Neuchâtel (Thesis director)
Prof Redouan Bshary, University of Neuchâtel
Dr Catherine Crockford, Max Planck Institute for Evolutionary Anthropology
Prof Robert Seyfarth, The University of Pennsylvania

UNIVERSITÉ DE NEUCHÂTEL

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

IMPRIMATUR POUR THESE DE DOCTORAT

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

Monsieur Adwait DESHPANDE

Titre:

“Social learning and flexibility in vocal communication of wild vervet monkeys”

sur le rapport des membres du jury composé comme suit:

- Prof. Klaus Zuberbühler, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Redouan Bshary, Université de Neuchâtel, Suisse
- Dr Catherine Crockford, Max Planck Institute of Evolutionary Anthropology, Leipzig, Allemagne
- Prof. Emeritus Robert M. Seyfarth, University of Pennsylvania, Philadelphia, USA

Neuchâtel, le 21 juin 2021

Le Doyen, Prof. A. Bangerter

Dedication

To my grandfather with whom conversations always blossomed

Summary of the thesis

Human language is probably the most complex communication system in the living world. It is investigated by various scientific disciplines, including linguistics, neuroscience, or cultural studies. However, despite this large and interdisciplinary effort, one key question has remained open and continues to perplex the scientific community; how could such an intricate system evolve? Comparative research on our extant evolutionary neighbours—the non-human primates—is often considered a good starting point to investigate the origins and evolution of human language. As humans communicate mainly with speech, primate vocal behaviour is the natural target of investigation, although this approach is not uniformly accepted. Behaviourist theories, in particular, characterise primate vocal behaviour as a predominantly hard-wired system, arguing that not much can be learned from it regarding language evolution. On the other hand, there is growing evidence for a considerable cognitive component in non-human primate communication, which often points to early signs of flexibility and indications of gradual evolutionary patterns more generally.

In this thesis, I ventured to further our understanding of the flexibility in non-human primate communication systems through series of field experiments on wild South African vervet monkeys. First, I assessed the capacity of vervet monkeys to socially learn novel call-context associations. Using unfamiliar animal models in conjunction with alarm call playbacks, I showed that monkeys rapidly associated alarm calls with these models, evidenced by high vigilance towards them in the subsequent encounter. Furthermore, some juveniles also produced alarm calls similar to the playbacks they heard during the first encounter, showing how rapid social learning could influence call comprehension and usage in this species.

In a second experiment, I tested the functional flexibility of vocalisations by providing wild vervet monkeys with opportunities to socially learn a novel usage of move-grunts to obtain food rewards. I worked with two groups that differed in the complexity of the learning stimuli provided during the experiments. For the first group of monkeys, I paired playback of move-grunt with a food dispenser providing a reward, such that the call predicted food as a simple conditioning stimulus. In the second group, I provided subjects with a demonstration video of a conspecific producing a move-grunt in order to activate the food dispenser and obtain a reward. While I did not find any evidence for relevant learning in the first group, a juvenile female from the second group started to produce her own move-grunts to obtain food rewards, suggesting that primates can learn to produce calls in completely novel circumstances if provided with the right social input.

The focus of the third experiment was on flexibility in call perception. I examined whether vervet monkeys, when confronted with referentially ambiguous calls, use contextual information to respond to them. I addressed this by probing them with male 'leopard' alarm calls, which can naturally be given to terrestrial predators or during intergroup encounters. In the experiment, I played back leopard alarms either during natural between-group encounters or in a control situation. The subjects showed anti-predator responses and looked for additional information in control but not inter-group situations, suggesting that call meaning in primate communication is subject to simple forms of pragmatic inference.

The results of these experiments indicate that non-human primate vocal communication rests on a primitive cognitive infrastructure that, within the human lineage, could have gradually evolved into the complex communication system seen in today's human languages. Finally, I conclude the thesis by proposing a classification scheme for non-human primate vocalisations, based on differences in underlying cognitive complexity, and briefly speculate about the future of primate research in light of emerging technologies that have the potential to revolutionise our understanding of the evolution of human and non-human communication.

Keywords

Evolution of language; Non-human primates; Comparative Cognition; Alarm calls

Résumé de la thèse

Le langage humain est probablement le système de communication le plus complexe du monde vivant. Il est étudié par diverses disciplines scientifiques, notamment la linguistique, les neurosciences ou les études culturelles. Cependant, malgré cet effort important et interdisciplinaire, une question clé est restée ouverte et continue de laisser perplexe la communauté scientifique : comment un système aussi complexe a-t-il pu évoluer ? La recherche comparative sur nos voisins évolutifs actuels - les primates non humains - est souvent considérée comme un bon point de départ pour étudier les origines et l'évolution du langage humain. Comme les humains communiquent principalement par la parole, le comportement vocal des primates est la cible naturelle des recherches, bien que cette approche ne soit pas uniformément acceptée. Les théories comportementalistes, en particulier, caractérisent le comportement vocal des primates comme un système essentiellement câblé, et soutiennent que l'on ne peut pas en tirer beaucoup d'enseignements sur l'évolution du langage. D'un autre côté, il existe de plus en plus de preuves d'une composante cognitive considérable dans la communication des primates non humains, ce qui indique souvent des signes précoces de flexibilité et, plus généralement, des indications de modèles évolutifs graduels.

Dans cette thèse, j'ai tenté d'approfondir notre compréhension de la flexibilité des systèmes de communication des primates non humains par le biais d'une série d'expériences de terrain sur des singes vervets sauvages d'Afrique du Sud. Tout d'abord, j'ai évalué la capacité des singes vervets à apprendre socialement de nouvelles associations appel-contexte. En utilisant des modèles animaux non familiers en conjonction avec des enregistrements de cris d'alarme, j'ai montré que les singes associaient rapidement les cris d'alarme à ces modèles, ce qui se traduisait par une grande vigilance à leur égard lors de la rencontre suivante. De plus, certains jeunes ont également produit des cris d'alarme similaires à ceux qu'ils avaient entendus lors de la première rencontre, ce qui montre comment un apprentissage social rapide peut influencer la compréhension et l'utilisation des cris chez cette espèce.

Dans une deuxième expérience, j'ai testé la flexibilité fonctionnelle des vocalisations en offrant aux singes vervets sauvages la possibilité d'apprendre socialement une nouvelle utilisation des grognements pour obtenir des récompenses alimentaires. J'ai travaillé avec deux groupes qui différaient par la complexité des stimuli d'apprentissage fournis pendant les expériences. Pour le premier groupe de singes, j'ai couplé la lecture du move-grunt avec un distributeur de nourriture fournissant une récompense, de sorte que l'appel prédisait la nourriture comme stimulus de conditionnement simple. Dans le second groupe, j'ai fourni aux sujets une vidéo de démonstration d'un congénère produisant un grognement de déplacement afin d'activer le distributeur de nourriture et d'obtenir une récompense. Alors que je n'ai trouvé aucune preuve d'un apprentissage pertinent dans le premier groupe, une femelle juvénile du second groupe a

commencé à produire ses propres grognements pour obtenir des récompenses alimentaires, ce qui suggère que les primates peuvent apprendre à produire des cris dans des circonstances complètement nouvelles s'ils reçoivent un apport social approprié.

La troisième expérience portait sur la flexibilité de la perception des cris. J'ai examiné si les singes vervets, lorsqu'ils sont confrontés à des cris ambigus sur le plan référentiel, utilisent des informations contextuelles pour y répondre. J'ai abordé cette question en les sondant avec des cris d'alarme de type " léopard ", qui peuvent naturellement être émis par des prédateurs terrestres ou lors de rencontres intergroupes. Dans l'expérience, j'ai fait écouter des cris d'alarme de léopard soit lors de rencontres naturelles entre groupes, soit dans une situation de contrôle. Les sujets ont montré des réponses anti-prédateurs et ont cherché des informations supplémentaires dans les situations de contrôle mais pas dans les situations intergroupes, ce qui suggère que la signification des cris dans la communication des primates est sujette à des formes simples d'inférence pragmatique.

Les résultats de ces expériences indiquent que la communication vocale des primates non humains repose sur une infrastructure cognitive primitive qui, au sein de la lignée humaine, aurait pu évoluer progressivement vers le système de communication complexe que l'on observe dans les langues humaines actuelles. Enfin, je conclus la thèse en proposant un schéma de classification des vocalisations des primates non humains, basé sur les différences de complexité cognitive sous-jacente, et je spécule brièvement sur l'avenir de la recherche sur les primates à la lumière des technologies émergentes qui ont le potentiel de révolutionner notre compréhension de l'évolution de la communication humaine et non humaine.

Mots clés

Évolution du langage ; primates non humains ; cognition comparée ; cris d'alarme.

Acknowledgements

First, I would like to thank my supervisor, Professor Klaus Zuberbühler, for giving me this incredible research opportunity. I am grateful for all the support he provided during this journey and for his stimulating supervision. I appreciate the freedom Klaus gave me to explore new ideas and work on them as I wished. Although I was free to explore and design my experiments, he was always there when I needed his guidance in the process of ideation, data collection, analyses and writing. I do not think I could have done the experiments this PhD without Klaus being my supervisor.

I would like to express gratitude to Professor Redouan Bshary for his continuous support and for agreeing to be an internal examiner of the thesis. I would also like to thank Professor Robert Seyfarth and Doctor Catherine Crockford for accepting to be part of my thesis committee as external examiners. Their constructive criticism has helped me immensely in improving the final draft of the thesis.

I collected the data at the Inkawu Vervet Project (IVP) run by Professor Erica van de Waal. IVP is one of the best field stations I know, and it is mainly because of Erica's leadership. She always supported my experiments and promptly solved any problems we encountered during the fieldwork. The ease with which I could work with monkeys and collect the data was impossible without all the volunteers who initially worked really hard and habituated the monkeys. I still cannot believe that I spent almost two years at the IVP and worked with more than forty passionate volunteers who came from all over the world. It was because of them I got a truly multicultural experience while living in the remote, isolated house in the middle of savanna. I would like to thank each volunteer and master's student for the data collected and for every moment living as housemates at IVP.

Other than volunteers, the pillars of IVP are the researchers and managers who work there. I didn't know what to expect when I first reach Mawana. However, from day one, fellow PhD student Axelle taught me everything I needed to learn. I would like to thank you for being such an awesome friend. Charlotte has been a fantastic person to live and work with and a terrific researcher whom I admire. Pooja and Tecla, who joined later as PhDs, were superb colleagues and friends. Thank you, Pooja, for your chilled out vibe, which always made us feel lighter and happy. Tecla, thank you for your fabulous cakes and for teaching me Baie Dankie along with Charlene. Also, I enjoyed working and supervising an ideal master's student and my co-author Bas.

Arend was the IVP manager for most of the time I was there. You taught me the skills needed to survive the bushveld and as well as the braai. I appreciate all the hard work you did to keep the project running, ranging from fixing the house to monitoring the potentially dangerous lions in reserve. Everyone always felt safe in such a remote place because of your efforts. I will always cherish your Potjiekos, Vetkoeks and Baboties. Thank you, Jenny, for always spreading positivity and sparking joy in the house. Later, Mike took over as the equally excellent manager. I appreciate all the help and support he provided during the last stretch of my fieldwork and initiating night drives, birding trails and some good times in St Lucia.

The stay at IVP was further made pleasurable by all the four-legged house members. First, Shadow, Bushman, Ghost and Axou. Later with Duke, Ilanga and Invula. All the hours spent with these furry loving creatures would remain in my memory for a long time. I would like to thank the owners of Mawana for letting me work there. Mainly, I am thankful to Kali, Barry sand Paul (managers at Mawana) and late Beyers for being there whenever needed. Sandy, along with the late Bruce of the vervet sanctuary, supported the project from the beginning. They are some of the loveliest people I met in my life. I cannot thank South African people enough for their help, support and hospitability during my stay in one of the most beautiful countries I ever visited. I witnessed the most magnificent sunsets of my life here in South Africa. I would like to thank Klaus again, who let me run the field course in Uganda, where I was mind blown after seeing the wild chimpanzee for the first time.

Back at the University, Fredy, Coco, Stephanie, Melissa, Yaelle, Alice, Quentin, thank you for a few laughs and the discussion we shared. Christof and Christoph, I appreciate all your support and the insightful discussions, which enriched me to become a better researcher. Julian, you have been an awesome lab-mate. I would also like to thank Andrea. Because of her, I was able to navigate administrative work related to this project smoothly from Africa. I also appreciate the constant assistance from the Biology secretariat and administration team of the University of Neuchatel.

Thank you, Louis and Odile, for letting me stay in the room with a great lake view where I always found peace and Theo for being an awesome housemate.

It was impossible to complete this adventure without having a buddy like Miguel on my side. Miguel, thank you for inspiring me on every step of this journey. You are the most hardworking person I have ever met, and I always aspired to work like you. I cherish all the field days, laughs, beers, hikes, dives, debates, conversations that we shared. I am looking forward to visiting together all the forests and mountains we spoke about.

My parents and family have always been supportive. I would like to thank them for all their love and encouragement.

Last ten years, Maithili has been my friend, support, guide, and partner. Because of her, I got the courage to seriously think about my passion and hobby of observing animals as a career option. Very few people in India get support for such an ‘unconventional’ career choice. She tolerated all my whims, tantrums during the PhD. Her focus and concentration while working motivated me to pick myself up and finish the writing. I could start, survive and complete the PhD only because of her being there with all the love and patience. I cannot be grateful enough.

Finally, this research was only possible because of the cooperating monkeys. I will always be thankful to them and all the other non-humans for letting me glimpse into their lives, which ironically helped me understand the human mind a bit more.

Ethics statement

Monkeys voluntarily participated in all the experiments and freely interacted with the experimental setups. All the experiments and observational data collection protocols were approved by Ezemvelo KZN Wildlife and the Ethics Commission of the University of Neuchâtel (permit number 28/2018).

Funding

This study primarily was funded by the Swiss National Science Foundation through grants (31003A_166458 and 310030_185324) awarded to Professor Klaus Zuberbuhler.

IVP is a registered Not-for-profit organisation in South Africa (Registration Number: 202-031NPO) established by a team of researchers from various Swiss Universities. Different funding agencies support the field station through the grants acquired by collaborating researchers.

Table of contents

Chapter 1: General introduction

1.1 Evolution of human language

"Monkeys are superior to (humans) in this: when a monkey looks into a mirror, (s)he sees a monkey."

The cynical Mauritian humourist Malcolm de Chazal eccentrically captures the difference between humans and non-humans in a single quote. On the other hand, the pursuit of comparing humans to non-human animals forms an extensive body of scientific literature. Researchers have found many similarities between humans and other animals in almost every aspect of biology. However, one human feature that exceptionally stands out as unique and probably 'superior' is the human language [1]. Human language is intricately complex in every sense. The complexity of language can be gauged just by the range of disciplines that directly study the language faculty [2]. For example, language has been a focus of research in biology, psychology, mathematics, computer science, philosophy [3]. Furthermore, within every discipline, language can be studied from different perspectives (Figure 1.1).

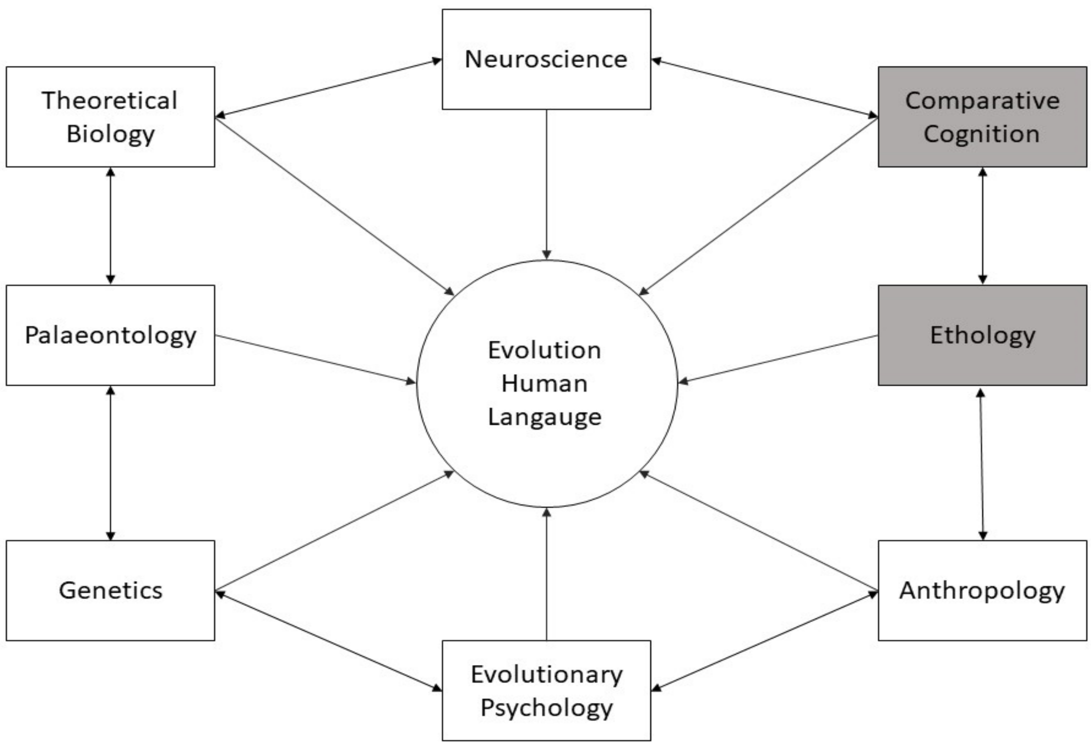

Figure 1.1: The inter-disciplinary nature of investigations on language evolution. The diagram shows a non-exhaustive list of sub-disciplines of biology that focus on different aspects of language evolution and often inform and overlap with each other. This thesis mainly uses methods and approaches from comparative cognition and ethology.

One of the hard open questions of this multidisciplinary subject is how such a complex communication system evolved from seemingly primitive communication systems that we see in other animals? One of the challenges in investigating this hard question is that, unlike some other biological traits, language and the associated organs, such as the larynx and brain, are not fossilised [4]. Thus, it is hard to directly compare the linguistic abilities of our extinct ancestors as it has been done for other anatomical traits. The next best approach is to compare human language to communication systems of extant animal species [3].

1.2 The comparative approach

The comparative approach might seem unreasonable or even forced, given the stark differences between human and non-human communication systems [4]. In fact, in 1866, any discussion on the evolutionary continuity of language was banned by the influential Société Linguistique de Paris [5]. However, it remains one of the powerful tools to study language evolution if used without excessive anthropomorphism. The comparison between communication systems of different species can help in understanding both homologues (similar biological traits found between the species that are derived from common ancestors) and analogues (independently evolved traits in distant species that share a common form or function) [6]. The logical starting point to look for homologues of human language is to study the communication systems of closest living relatives, the non-human primates.

Earlier studies involving non-human primates were mainly restricted to ape language experiments. These experiments were focused on training the human-reared apes in some form of human language. However, their success remained limited [7]. For example, after extensive training and effort, Viki, the human reared chimpanzee (*Pan troglodytes*), could utter only a few simple words like mama, papa, cup [8]. The success in the non-vocal domains such as sign language was not encouraging, and the approach faced extensive criticism (Reviewed in [9]).

Parallelly, behavioural ecologists looked at the animal communication systems through Tinbergen's framework [10]. In Tinbergen's view, the phylogenetic history and adaptive significance of any behaviour, including animal signals, offer the ultimate explanation about why particular behaviour evolved. Simultaneously, the proximate explanations elucidate how the particular behaviour develops during ontogeny and the underlying mechanisms [10]. Keeping up with the behavioural ecological view, Dawkins and Krebs [11] challenged the general model of animal communication, which assumed signals to convey information from the signaller to the receiver. They argued that there is usually a conflict of interest between signallers and receivers and proposed that communication evolves as an arms race between senders and receivers trying to manipulate each other. However, later Maynard Smith and

Harper [12] pointed out that evolutionarily stable signals would evolve only if, on average, both parties receive some benefits out of the signalling systems. In other words, animal signals, on average, must carry correct information for the receiver to be selected to respond to the sender [12].

Among the various perspectives towards a comparative approach of communication, the studies on wild Kenyan vervet monkeys (*Chlorocebus pygerythrus hilgerti*), first by Struhsaker [13], followed by Seyfarth, Cheney and Marler [14,15], became an important milestone. They discovered that vervet monkeys produce acoustically distinct alarm calls, which correspond to different classes of predators. Furthermore, receivers show appropriate and visibly different anti-predator response to each alarm call [14,15]. Authors interpreted it as a form of semantic communication similar to human words [15]. This interpretation has been updated and revised since then, which I will describe later in the thesis. Nevertheless, these studies are especially important as they drew a direct parallel between human and non-human primate communication systems with experimental evidence from the wild primates [16]. Furthermore, subsequent studies on vervet monkeys [17,18] brought the discussion about cognitive capacities and their relationship with communicative abilities to the forefront of animal behaviour research. These studies inspired a new line of comparative research that focused explicitly on finding roots and origins of human language in non-human vocal communication [16,19,20].

1.3 Non-human primate vocal communication: State of the art

Here, I will review selective literature and describe our current understanding of non-human primate vocal communication and its relationship with human language. However, it is essential to note that communication research is not limited to the auditory modality. Instead, many studies suggest that visual communication through gestures is far more flexible than vocal communication in apes and other simians [21–24]. Some researchers have even proposed a theory of the gestural origins of human language [25]. Several studies also highlight the importance of integrating different communication modalities to understand the origins of human communication [26,27]. However, henceforth in this thesis, I restrict my research to the vocal communication of non-human primates while acknowledging the importance of integrating different modalities.

Initially, it was generally thought that non-human primate vocal communication is innate and reflexive [28]. In other words, vocalisations merely reflect an individual's internal or affective state. Evidence for such characterisation mainly came from social isolation, deafening and cross-fostering experiments conducted in captive settings, and also from the failures of some ape language experiments [29]. First, Newman and Symmes [30] showed that social isolation could strongly affect primate vocal development, resulting in spectro-temporal abnormalities in specific calls. However, other studies failed to confirm these results and suggested that social isolation does not affect vocal development [31–33]. Also, it has been argued that social isolation could influence overall social behaviour and reported vocal abnormalities could result from overall social deprivation and not explicitly related to the deprivation in auditory experience [29]. This problem was addressed by the deafening studies where individuals are deprived only of the auditory experience while controlling for other social factors. Deafening experiments showed mixed results; some studies suggested that there is no effect on vocal development caused by deafening [31,34]. On the contrary, others suggested strong effects manifested in spectrotemporal abnormalities and the persistence of infant-like calls into adulthood in deafened individuals [35]. The less invasive method of cross-fostering in which individuals of one species are reared with a foster mother of other species showed some role of learning, at least in food calls [36]. However, again other studies failed to replicate these results [37].

The general failure to find flexibility in non-human primate vocal communication led some researchers to theorise how such a vocal communicative system could work without flexibility and cognitive input. First, Morton proposed that individuals' affective states are directly linked to the acoustic structure of their call. He termed it as 'motivation-structural' rules [38,39]. Owren and Rendall further developed this idea into a model of vocal communication in non-human primates termed as the "affect-conditioning" model [40]. The main tenets of this model are as follows (a) Non-human primate calls reflect their internal state, and their main function is to influence the behaviour of others (receivers) (b) Receiver responses are hard-wired, i.e., certain acoustic structures automatically evoke particular physiological responses. For example, high-pitched noisy screams are inherently perceived as obnoxious and could function to deter an aggressor (c) Receiver responses are also subject to Pavlovian conditioning, driven by the inter-individual interaction history, which produces anticipatory responses in individuals [40]. Overall, the affect-conditioning model does not ascribe any flexibility to the producer and attributes minimal flexibility to the receiver. Proponents of this view also argue that non-human primate vocalisations do not contain semantic information and that linguistic terms or metaphors to describe animal communication are unnecessary and misleading [41,42].

However, after the classic studies on vervet alarm calls, the characterisation of primate vocalisations as reflexive and inflexible was already changing. Many studies on different species in the wild and captivity attributed much more complexity and plasticity to non-human primate calls [43,44]. The growing evidence for the flexible nature of non-human primate vocalisations came from studies focused on different aspects of vocalisations such as call similarity and convergence, dialects, underlying intentionality, audience effect or strategic usage of calls.

The evidence for capacity for acoustic modification of calls and vocal production learning (learning to produce or modify specific call) [45] comes from different studies. Campbell's monkey (*Cercopithecus campbelli*) contact calls were acoustically more similar when exchanged between socially close individuals compared to other individuals. Furthermore, call similarity was also more correlated with the strength of the social bond than the genetic relatedness between individuals [46]. Similarly, the grunts of Guinea baboons (*Papio papio*) of the same gang were more similar compared to members of other gangs in the multi-level society. Again, the similarity was not explained by genetic relatedness [47]. Group or population-specific calls were documented in apes and other non-human primates [48–50]. In the wild, it has been shown that different chimpanzee communities have structurally specific pant hoots [51]. Similar differences in pant hoots were also recorded between captive colonies of chimpanzees [52]. Also, there is evidence of vocal production learning of food grunts after integrating two different groups of housed chimpanzees [53].

In a recent study on marmosets (*Callithrix jacchus*), using dense sampling and biomechanical modelling, researchers showed how dramatic changes in vocalisations could not be accounted for only by maturational changes and genetic relationships; and how contingent parental feedbacks influence the rate of vocal development in this species, emphasising the role of vocal production learning [54,55]. Although the mentioned studies suggest flexibility in vocal production, generally, these effects were small and only seemed to influence the fine structure of a stereotypical call [56]. Instead, much stronger evidence for control and flexibility in vocal production comparable to human speech has been found in distant taxa, such as cetaceans, pinnipeds and songbirds [57–59], suggesting convergent evolution.

Despite the limited control over the call structure, other studies suggest more flexibility in call usage—I.e., the ability to use or withhold different calls strategically [43]. For example, wild chimpanzees adjusted their alarm calls based on the attentional states of the audience. Furthermore, intentional alarm calls attracted more audience attention than those produced randomly [60,61]. Chimpanzees also strategically produced food calls to inform particular individuals [62,63]. Audience awareness has been found in chimpanzee greeting calls, and also, the victim chimpanzees are known to exaggerate the severity of aggression through screams if bystanders consist of high-ranking individuals [64,65]. Similar audience effects have been found in copulation calls of both chimpanzee and bonobo (*Pan paniscus*) females [66–68] and also in alarm calls of Thomas langurs (*Presbytis thomasi*) [69].

Control on call usage is also evident from studies on the combinatoriality of calls. Several non-human primate species have been documented to use different call combinations depending on external context. For example, male putty-nosed monkeys produced four call sequences using two basic calls, and each of the sequences had independent meaning [70]. Comparable call combinations have been recorded in different old world and new world monkeys as well as apes [71–73]

Do non-human primate receivers also show flexibility in terms of call comprehension? Studies that investigated how meaning is extracted from the call and how call-referent relationships are formed indeed attribute more flexibility to call receivers. After the discovery that vervet receivers appropriately respond to different alarm calls, similar phenomena have been seen in different non-human primates [74,75]. Also, there is good evidence that monkeys could relate particular alarm calls to the probable causal event and respond accordingly [76]. Field experiments have shown that recipients relied on contextual cues and prior knowledge to interpret ambiguous calls akin to linguistic pragmatics [77,78].

Early studies also attempted to look at the role of social learning (learning from other conspecifics [79]) in primate vocalisations. In their detailed study on the vocal development in vervet monkeys, Cheney and Seyfarth documented the changes in different vocalisations with respect to age. They also speculated the role of learning in the comprehension and usage of vocalisations [80,81]. Other studies also noted the importance of auditory experience and the maturational changes in vocal tract anatomy and increased motor control in the ontogeny of vocal signals in vervet monkeys [82,83]. Furthermore, it has been shown that Campbell monkeys pay more attention to adult calls [84]. These findings highlight the importance of social influence on the ontogeny of vocal communication.

Overall, there is growing evidence for flexibility and learning in all aspects of vocal communication. Consequently, some researchers theorised various concepts and frameworks to explain the origins and mechanisms of flexibility and cognitive component of the calls and their parallels to human language. For example, Macedonia and Evans [85,86] defined the category of 'Functionally referential' signals. Calls that are context-specific and evoke specific responses even in the absence of stimuli qualify as the functionally referential signals. According to Macedonia and Evans, the vervet alarm call system is a textbook example of functionally referential calls as different alarm calls evoke predator-specific responses in receivers even in the absence of a predator [86]. Referentiality is considered as a hallmark of human language. The functionally referential signals were then thought of as a precursor to referentiality in human communication [85]. Biologists have documented functionally referential calls in many non-human primate species and other taxa [87]. However, recently the concept of functional referentiality was criticised. Wheeler and Fischer [88] argued that, by definition, context-specific signals are cognitively less demanding for receivers, and they are not inherently different from any other animal signal. Furthermore, they pointed out that many of the so-called functionally referential signals are not exclusively produced in a single context. They suggested abandoning the concept and looking for cognitive sophistication in the receivers as signal production is essentially hard-wired (But see [89] for a response to Fischer and Wheeler's criticism).

Generally, no one argued that non-human primate vocal communication involves mental state attribution, as seen in human communication. However, at the same time, it also meets the criteria for intentional communication [21,60,90–92]. To resolve this confusion, researchers have also established a new framework to operationalise the notion of intentionality in non-human primate communication. The framework is less focused on mental state attribution, which is implicit in the traditional Gricean approach to intentional communication [93].

Price and Fischer's view is intermediate in terms of the cognitive complexity of vocalisations. They claim non-human primate communication can be best conceived as goal-directed, with no communicative or informative intention [94]. They also argue, similar to the affect-conditioning model, that complexity primarily lies in the receiver's side. Schamberg et al.'s synthesis focus on the caller's goal to elicit particular behaviour in the receiver as the primary explanation for vocal production in non-human primates. Furthermore, they also indicate social complexity as the main driver of the call repertoire size of species [95].

Seyfarth and Cheney also focus on sociality to explain the role of cognition in non-human primate communication by highlighting the importance of social bonds for individual fitness in group-living primates [96–98]. To put it simply, they argue that social knowledge and cognition play an important role in maintaining social relationships, and flexible call usage and comprehension was shaped by selection pressure to solve social challenges and to make social decisions [99–102]. Furthermore, according to them, computational and conceptual abilities required for human language are present (albeit in primitive forms) in non-human primates. However, non-human primates have constraints on communicating their internal representations of the world, which later evolved in primate lineage [103].

Despite the diverse and sometimes opposing viewpoints, there is a general consensus among researchers on the following points (1) Vocal production learning is limited in non-human primates, and better parallels to human language in this aspect can be seen in other taxa; (2) Recognition of individuals based on their vocalisations is widely spread across non-human primate species [104]; (3) The affective or internal state of an individual is reflected in certain parameters such as fundamental frequency, amplitude, duration, call rate together termed as call prosody [95].

Overall, the evidence varying in strengths is present for most viewpoints and arguments that broadly lie between two extreme positions that non-human primate communication is either hard-wired or cognitively complex. Thus, our current understanding of non-human primate vocal communication systems can be described as cloudy at best. Studies in ecologically relevant settings that empirically test the extent of flexibility and learnability in both signallers and receivers are urgently needed. To push the field forward, it is also necessary to elucidate precise mechanisms underlying the production and reception of different call types. Finally, a coherent conceptual framework could make our understanding clearer and help us comprehend the evolutionary journey of the human communication system.

1.4 Study species

The members of the genus *Chlorocebus* are Old World monkeys belonging to the family *Cercopithecidae* and are native to Africa. Recent taxonomic studies identify at least five subspecies that are geographically distributed across the continent. Subspecies differ subtly in their morphologies but largely share similar social structure and behaviours [105,106]. Here, I studied a particular subspecies, *Chlorocebus pygerythrus pygerythrus*, restricted in its distribution to southern Africa. Populations are found in several countries such as South Africa, Lesotho, Botswana and Swaziland [107].

The southern African vervet monkeys mainly inhabit savanna, riverine woodland, coastal forest and mountains up to 4,000 m. The species is known to be highly resilient to adverse environmental changes. Thus, it is also commonly seen in secondary forests, highly fragmented habitats, agricultural landscapes and urban areas within its geographic range. The annual home range size of a group has been observed to be as large as 1.76 km^2, with an average population density of 54.68 animals/km^2; however, these attributes can vary according to various ecological factors [105]. Vervet monkeys are considered 'least concerned' by International Union for Conservation of Nature (IUCN) because of the stable population trend and wide geographic distribution [108].

Furthermore, in some parts, the species is considered a nuisance or vermin as it tends to enter and destroy urban areas and raid agriculture plots, sometimes resulting in retaliatory culling by humans [108]. Hence, it is required to regularly monitor the species' population and ensure its long-term survival. However, given the status, long-term field studies and careful manipulative field experiments on a small population should not pose any apparent direct threat to this species.

The species shows sexual dimorphism; the males are larger in body weight and length. Adult females weigh between 3 and 5 kg and average 4.1 kg. Body length measure between 300 and 495 mm, averaging around 426 mm. Adult males weigh between 4 and 8 kg, averaging 5.5 kg and have a body length between 420 and 600 mm, averaging 490 mm from the top of the head to the base of the tail. Males have a bright turquoise-blue scrotum, which makes them easily distinguishable [109].

The social structure of South African vervet monkeys is similar to other *Cercopithecidae* members. Females are philopatric and stay in their natal groups for their whole lives, whereas males migrate to other, often adjacent groups after sexual maturity (around the age of 4-6 years) [110]. Group size typically consists of 10 to 70 individuals depending on habitat quality [106]. Female matrilines have a strict linear dominance hierarchy that rarely changes. Thus, female infants inherit their matrilineal dominance ranks by birth. All adult immigrant males are dominant over females. However, unlike females, between-male hierarchy keeps changing and is determined by age, tenure in the group, fighting abilities and alliances with other group members [111]. Dominance mainly allows individuals to gain priority when accessing resources [18]. Moreover, individuals generally support higher-ranking individuals during group aggression, probably to avoid self-injuries [112].

Similar to other group-living primate species, sociality plays a crucial role in achieving long-term fitness benefits [98,99,113]. The research has shown that vervet monkeys are quite 'Machiavellian' [114] and behave strategically during social interactions [112,115–117]. They actively keep track of relationships between other group members and are aware of the dominance rank of every group member [112]. Such social knowledge helps them maintain their dominance rank and form strategic coalitions and long-term social bonds [96].

1.5 Vocal communication in vervet monkeys

In this section, I will provide a brief overview of vocal communication in vervet monkeys. Throughout the subsequent chapters, different aspects of the vervet vocal communication and call types are described with more details in the context of specific hypotheses.

The vervet vocal communication is widely known for its alarm call system. Vervet monkeys produce at least three acoustically different alarm calls when they encounter different predator classes [15]. The 'leopard alarms' are produced when monkeys encounter terrestrial predators like leopards and jackals. 'eagle alarms' are elicited for the aerial predators like different birds of prey, and 'snake alarms' are produced when monkeys detect pythons and other snake species [118]. Receivers respond in an appropriate way to the different alarm calls. When the monkeys hear 'leopard alarms', they typically run up a tree, a safe refuge from a terrestrial predator. Whereas, after hearing the 'eagle alarms', monkeys typically run down into a bush and scan the sky to locate and monitor the attacking raptor. Finally, 'snake alarms' trigger bipedal vigilance towards the ground [15,118]. These are the findings of the early studies. Recent re-analysis showed that the vervet alarm call system could be more complex than previously thought. New insights suggested that three types of alarm calls are acoustically graded rather than strongly categorical as three distinct call types, and receivers might not be receiving semantic information from the different alarm calls [119]. Furthermore, alarm calls were observed to be produced in different non-predatory contexts, as noted by earlier studies [18,119]. The male 'leopard alarm' or male 'alarm bark' is an interesting call type in vervet monkeys. Adult males are known to produce these calls during encounters with terrestrial predators as well as during the within and between-group aggressive interactions. The exact function of this call type is not yet fully understood. This call may serve multiple functions, such as territorial defence and predator deterrence [119].

The second group of calls in the vocal repertoire of vervet monkeys is known as 'grunts', which are primarily social calls [17]. Grunts can be divided into subtypes based on the functional significance and acoustic differences, such as grunts to dominant or sub-ordinates. The function of social grunts is generally thought to be that of benign intent, to reduce the probability of aggression during social interactions (but see [120] for different hypotheses regarding the functions of grunts). Another sub-type of grunts is an acoustic variant used for the coordination of group movement. The move-grunts (or 'moving-into-an-open-area' grunt) are typically given by individuals while moving into open areas and observing other individuals move [81]. Move-grunts are also produced to initiate group movements, especially before crossing the rivers (AD, unpublished data).

The remaining call types include lost or contact calls when individuals are separated from the group or encounter other groups [18]. Finally, vervet monkeys produce high-pitched, high-amplitude screams during aggressive interactions, which differ between victims and aggressors and could also be directed towards bystanders [121]. Some of the calls are also described onomatopoeically as 'chutter', 'wrrs' and 'wraps' etc. [18]. However, I will not refer to call types onomatopoeically to avoid confusion.

Observational studies on vocal development suggest that infant vervet monkeys can produce calls that are acoustically similar to adult calls from the early days of their lives [80,81].

However, the usage and comprehension of call types show different developmental trajectories [122]. Most importantly, it has been experimentally shown that vervet monkeys can individually recognise other group members by their vocalisations and discriminate calls of kin and outgroup members [123]. Thus, competence in vocal communication has consequences for an individual's fitness during predator encounters and social interactions in this species [103].

1.6 Study site and groups

The data for this project were collected at the Inkawu Vervet Project (IVP) in South Africa. It is situated inside the 'Mawana Game Reserve', a privately-owned hunting farm that spreads across 12,000 hectares in the Kwazulu Natal province. One part of the farm is reserved as a research area where other activities such as trophy hunting or game viewing are restricted.

The study area consisted of mainly secondary wooded savanna, patches of primary forest and riverine vegetation. Other wild mammals included mostly herbivore species, like impala, wildebeest, giraffe, zebra, kudus or blesbok and a small population of African elephants. The main terrestrial predator species of the vervet monkeys recorded at this site were caracal (*Caracal caracal*), serval (*Leptailurus serval*), jackal (*Lupulella mesomelas*) and rock pythons (*Python sebae natalensis*). Aerial predators mainly include the marshal eagle (*Polemaetus bellicosus*) along with several other raptor species. Compared to the studies on Kenyan subspecies of vervet monkeys, predatory pressure on the IVP population appears to be less [18]. At the time of the study, six groups of wild vervet monkeys were continuously monitored at the IVP. The groups varied in their degree of habituation to humans. Group sizes ranged from 15-60 individuals. One adult female in each group was fitted with a VHF radio collar, which helped to locate the group daily. Each individual in the study group was recognised by their morphological features, and individual life histories were systematically recorded.

IVP aims to gain crucial insights into the biology of the species through long-term continuous observations and experiments. Thus, the systematic ecological, behavioural and demographic data on every group is continuously collected. These data are essential to understand and quantify home ranges, dietary preferences, genetic relatedness, dominance hierarchies and social networks. For the long-term data, different protocols such as scan, focal sampling, and ad-lib observations were appropriately used to record different behaviours [124]. Monkeys from all the groups were regularly provisioned with the food items for the various experiments (corn kernels and apples) [117,125,126]. The researchers announced food provisioning events using specific calls that prompt monkeys to move towards the provisioning site. All the data at IVP were collected by several researchers and volunteer research assistants who underwent rigorous training. Standardised individual identification and inter-observer reliability tests ensured consistency in data collection. At IVP, three groups were very well habituated and studied continuously since 2010, allowing close behavioural observations from an average distance of 2-5 m. I conducted the main experiments on these well-habituated groups, namely Baie Dankie, Ankhase and Noha.

1.7 Scope of the thesis

The objective of the thesis is to further our understanding of the differences and similarities between non-human and human communication systems, which could help shed light on the evolution of human language. The approach is to empirically test (and retest) distinct but interconnected concepts and mechanisms proposed to explain non-human primate vocal communication. Here, I exclusively focus on learning and flexibility in comprehension and usage of different call types in non-human primates in ecologically relevant settings [45].

I explore the topic through an interdisciplinary lens of ethology and comparative cognition and briefly touch upon neurobiology whenever necessary. However, the methods used are mainly inspired by a rich tradition of field experiments in ethology. The three empirical chapters use novel field experiments to test various hypotheses regarding social learning, functional flexibility, and pragmatic inference in the vocal communication of the wild vervet monkey population in Western South Africa. In a final chapter, I summarise the results and also attempt to conceptualise a new framework to study primate vocal communication that could advance our understanding of the origins and evolution of human language.

1.7.1 Thesis chapters

In the first empirical chapter, I investigated the role of social learning in the acquisition of novel call-referent associations. Social learning can be an effective mechanism to cope with constantly changing environments, especially when the call meaning changes with time and space [127]. Alarm calls are ideal for such investigation, considering the constant changes in predatory species [128]. To investigate the role of social learning in alarm call comprehension and usage, I exposed two groups of monkeys to unfamiliar animal models in conjunction with playback of alarm calls. I predicted that if monkeys learned to associate the alarm call with the novel stimulus, they would show increased vigilance towards the model in subsequent encounters. Furthermore, if comprehension learning also leads to usage learning, individuals would also produce alarm calls similar to what they have heard from the playbacks during their first exposures to the models.

The second chapter focuses on the functional flexibility of the calls—i.e., an ability to produce the same call in several different contexts [129]. In humans, the functional flexibility of phonemes allows for rapid social learning of a language [130]. Here, I investigated to what extent non-human primate calls are functionally flexible. I designed field experiments involving a manual food-dispenser, where monkeys can socially learn the novel usage of a call to obtain food rewards. In one experiment, I provided only an auditory stimulus in the form of a playback of a conspecific call paired with food rewards. In a second experiment, an audio-visual stimulus in the form of a demonstration video of a conspecific was used. The success in learning the novel usage can be considered direct evidence for functional flexibility in primate vocal communication.

The third chapter is aimed at assessing the flexibility on the part of signal receivers. Some researchers argue that pragmatics (more than semantics) forms a foundation of human language [131]. Sophisticated psychological capacities, such as intentionality and theory of mind, allow humans to decipher 'intended' meaning, which could be orthogonal to the semantic meaning of the words. Do non-human primates also use some form of pragmatics to decipher the meaning of ambiguous calls? There is some indication for the presence of pragmatic inference in primates [102]. However, empirical tests remain scarce. This experiment tested whether wild vervet monkeys use contextual information to assess the cause of the alarm calls. I used inter-group encounters as natural priming events, followed by playbacks of referentially ambiguous calls. In the control condition, no prior information regarding the probable cause of the call was available to the receivers. I predicted that if the subject monkeys make pragmatic inferences about the cause of the calls, they would respond in a context-appropriate way to the same call.

The final chapter provides a synthesis of the results and conclusions of all the experiments, along with their broader implications. Furthermore, in this chapter, I identify deficiencies in the current models of non-human primate vocal communication and propose a new classification scheme for studying primate vocalisations that can be tested empirically across species. Finally, I turn to upcoming technologies and discuss their potential to provide new insights into primate behaviour and cognition, and ultimately to the origins and evolution of human language.

1.8 References

1. Chomsky N. 2017 Language architecture and its import for evolution. *Neurosci. Biobehav. Rev.* **81**, 295–300. (doi:10.1016/j.neubiorev.2017.01.053)

2. Fitch WT. 2017 Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* (doi:10.3758/s13423-017-1236-5)

3. Fitch W. 2010 *The evolution of language*. Cambridge University Press.

4. Fitch WT. 2000 The evolution of speech: A comparative review. *Trends Cogn. Sci.* **4**, 258–267. (doi:10.1016/S1364-6613(00)01494-7)

5. Christiansen MH, Kirby S. 2003 Language evolution: Consensus and controversies. *Trends Cogn. Sci.* **7**, 300–307. (doi:10.1016/S1364-6613(03)00136-0)

6. Hauser MD, Chomsky N, Fitch WTS. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science (80-.).* **298**, 1569–79. (doi:10.1126/science.298.5598.1569)

7. Pepperberg IM. 2017 Animal language studies: What happened? *Psychon. Bull. Rev.* **24**, 181–185. (doi:10.3758/s13423-016-1101-y)

8. Hayes C. 1951 *The ape in our house.*

9. Krause MA, Beran MJ. 2020 Words matter: Reflections on language projects with chimpanzees and their implications. *Am. J. Primatol.* **82**, 1–14. (doi:10.1002/ajp.23187)

10. Tinbergen N. 1963 On aims and methods of Ethology. *Z. Tierpsychol.* **20**, 410–433. (doi:10.1111/j.1439-0310.1963.tb01161.x)

11. Dawkins R, Krebs JR. 1979 Arms races between and within species. *Proc. R. Soc. London. Ser. B. Biol. Sci.* **205**, 489–511. (doi:10.1098/rspb.1979.0081)

12. Smith J, Harper D. 2003 *Animal signals*. Oxford University Press.

13. Struhsaker TT. 1967 Auditory communication among vervet monkeys (Cercopithecus aethiops). In *Social communication among primates*, pp. 281–324.

14. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Anim. Behav.* **28**, 1070–1094. (doi:10.1016/S0003-3472(80)80097-2)

15. Seyfarth RM, Cheney DL, Marler P. 1980 Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801–803. (doi:10.1126/science.7433999)

16. Fischer J. 2020 Nonhuman primate alarm calls then and now. *Anim. Behav. Cogn.* **7**, 108–116. (doi:10.26451/abc.07.02.05.2020)

17. Cheney DL, Seyfarth RM. 1982 How vervet monkeys perceive their grunts: Field playback experiments. *Anim. Behav.* **30**, 739–751. (doi:10.1016/S0003-3472(82)80146-2)

18. Cheney DL, Seyfarth RM. 1992 *How monkeys see the world: Inside the mind of another species*. The University of Chicago Press.

19. Vonk J. 2020 Forty years on from the question of referential signals in nonhuman communication. *Anim. Behav. Cogn.* **7**, 82–86. (doi:10.26451/abc.07.02.01.2020)

20. Snowdon CT. 2020 Vervet monkey alarm calls: Setting the historical context. *Anim. Behav. Cogn.* **7**, 87–94. (doi:10.26451/abc.07.02.02.2020)

21. Deshpande A, Gupta S, Sinha A. 2018 Intentional communication between wild bonnet macaques and humans. *Sci. Rep.* **8**. (doi:10.1038/s41598-018-22928-z)

22. Gupta S, Sinha A. 2016 Not here, there! Possible referential gesturing during allogrooming by wild bonnet macaques, Macaca radiata. *Anim. Cogn.* , 1–6. (doi:10.1007/s10071-016-1012-3)

23. Cartmill EA, Byrne RW. 2010 Semantics of primate gestures: Intentional meanings of orangutan gestures. *Anim. Cogn.* **13**, 793–804. (doi:10.1007/s10071-010-0328-7)

24. Hobaiter C, Byrne RW, Zuberbühler K. 2017 Wild chimpanzees' use of single and combined vocal and gestural signals. *Behav. Ecol. Sociobiol.* **71**. (doi:10.1007/s00265-017-2325-1)

25. Tomasello M. 2010 *Origins of Human Communication*. MIT press.

26. Fröhlich M *et al.* 2019 Multimodal communication and language origins: integrating gestures and vocalisations. *Biol. Rev.* **94**, 1809–1829. (doi:10.1111/brv.12535)

27. Slocombe KE, Waller BM, Liebal K. 2011 The language void: The need for multimodality in primate communication research. *Anim. Behav.* **81**, 919–924. (doi:10.1016/j.anbehav.2011.02.002)

28. Liebal K, Waller B, Slocombe K, Burrows A. 2014 *Primate Communication: a multimodal approach*. Cambridge University Press.

29. Egnor SER, Hauser MD. 2004 A paradox in the evolution of primate vocal learning. *Trends Neurosci.* **27**, 649–654. (doi:10.1016/j.tins.2004.08.009)

30. Newman JD, Symmes D. 1974 Vocal Pathology in Socially Deprived Monkeys. *Dev. Psychobiol.* **7**, 351–358. (doi:10.1002/dev.420070412)

31. Ploog D, Schott D, Winter P, Handley P. 1973 Ontogeny of Squirrel Monkey Calls Under Normal Conditions and Under Acoustic Isolation. *Behaviour* **47**, 230–239. (doi:10.1163/156853973X00085)

32. Eisenberg JF. 1976 Communication mechanisms and social integration in the black spider monkey, Ateles fusciceps robustus, and related species. *Smithson. Contrib. to Zool.* , 1–108. (doi:10.5479/si.00810282.213)

33. Hammerschmidt K, Newman JD, Champoux M, Suomi SJ. 2000 Changes in rhesus macaque 'Coo' vocalisations during early development. *Ethology* **106**, 873–886. (doi:10.1046/j.1439-0310.2000.00611.x)

34. Talmage-Riggs G, Winter P, Ploog D, Mayer W. 1972 Effect of deafening on the vocal behavior of the squirrel monkey (Saimiri sciureus). *Folia Primatol.*

35. Roupe S, Pistorio A, X W. In press. Vocal plasticity induced by auditory deprivation. In *Program No. 627.7 2003 Abstract Viewer and Itinerary Planner*, Society for Neuroscience Online,2003.

36. Masataka N, Fujita K. 1989 Vocal learning of Japanese and rhesus monkeys. *Behaviour* **109**, 191–199. (doi:10.1163/156853989X00222)

37. Owren MJ, Dieter JA, Seyfarth RM, Cheney DL. 1992 'Food' calls produced by adult female rhesus (Macaca mulatta) and Japanese (M. fuscata) macaques, thier normally-raised offspring, and offspring cross-fostered between species. *Behaviour* **120**, 218–231. (doi:10.1017/CBO9781107415324.004)

38. Morton ES. 1982 Grading, Discreteness, Redundancy, and Motivation-Structural Rules. In *Acoustic Communication in Birds*, pp. 183–212. Elsevier. (doi:10.1016/B978-0-08-092416-8.50015-2)

39. Morton ES. 1977 On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869. (doi:10.1086/283219)

40. Owren MJ, Rendall D. 1997 An Affect-Conditioning Model of Nonhuman Primate Vocal Signaling. pp. 299–346. (doi:10.1007/978-1-4899-1745-4_10)

41. Owren MJ, Rendall D. 2001 Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol. Anthropol.* **10**, 58–71. (doi:10.1002/evan.1014)

42. Owren MJ, Rendall D, Ryan MJ. 2010 Redefining animal signaling: Influence versus information in communication. *Biol. Philos.* **25**, 755–780. (doi:10.1007/s10539-010-9224-4)

43. Snowdon C. 2018 Cognitive Components of Vocal Communication: A Case Study. *Animals* **8**, 126. (doi:10.3390/ani8070126)

44. Snowdon CT. 2009 *Plasticity of Communication in Nonhuman Primates*. 1st edn. Elsevier Inc. (doi:10.1016/S0065-3454(09)40007-X)

45. Janik VM, Slater PB. 2000 The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11. (doi:10.1006/anbe.2000.1410)

46. Lemasson A, Ouattara K, Petit EJ, Zuberbühler K. 2011 Social learning of vocal structure in a nonhuman primate? *BMC Evol. Biol.* **11**. (doi:10.1186/1471-2148-11-362)

47. Fischer J, Wegdell F, Trede F, Pesco FD, Hammerschmidt K. 2020 Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. (doi:10.1101/2020.03.24.005504)

48. Wich SA *et al.* 2012 Call cultures in orang-utans? *PLoS One* **7**. (doi:10.1371/journal.pone.0036180)

49. Briseño-Jaramillo M, Estrada A, Lemasson A. 2015 Behavioural innovation and cultural transmission of communication signal in black howler monkeys. *Sci. Rep.* **5**, 13400. (doi:10.1038/srep13400)

50. Lameira AR, Hardus ME, Nouwen KJJM, Topelberg E, Delgado RA, Spruijt BM, Sterck EHM, Knott CD, Wich SA. 2013 Population-Specific Use of the Same Tool-Assisted Alarm Call between Two Wild Orangutan Populations (Pongopygmaeus wurmbii) Indicates Functional Arbitrariness. *PLoS One* **8**, e69749. (doi:10.1371/journal.pone.0069749)

51. Crockford C, Herbinger I, Vigilant L, Boesch C. 2004 Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**, 221–243. (doi:10.1111/j.1439-0310.2004.00968.x)

52. Marshall A, Wrangham R, Arcadi A. 1999 Does learning affect the structure of vocalisations in chimpanzees? *Anim. Behav.* **58**, 825–830. (doi:10.1006/anbe.1999.1219)

53. Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE. 2015 Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* **25**, 495–499. (doi:10.1016/j.cub.2014.12.032)

54. Takahashi DY, Liao DA, Ghazanfar AA. 2017 Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Curr. Biol.* , 1–9. (doi:10.1016/j.cub.2017.05.004)

55. Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, Ghazanfar AA. 2015 The developmental dynamics of marmoset monkey vocal production. *Science* **349**, 734–8. (doi:10.1126/science.aab1058)

56. Fischer J, Hammerschmidt K. 2019 Towards a new taxonomy of primate vocal learning. *Philos. Trans. R. Soc. B Biol. Sci.* (doi:10.1098/rstb.2019.0045)

57. Garland EC, McGregor PK. 2020 Cultural transmission, evolution and revolution in vocal displays: insights from bird and whale song. *Front. Psychol.* **11**, 544929. (doi:10.3389/fpsyg.2020.544929)

58. Stansbury AL, Janik VM. 2019 Formant Modification through Vocal Production Learning in Gray Seals. *Curr. Biol.* , 1–6. (doi:10.1016/j.cub.2019.05.071)

59. Catchpole CK, Slater PJB. 2003 *Bird song: biological themes and variations*. Cambridge University Press.

60. Schel AM, Townsend SW, Machanda Z, Zuberbühler K, Slocombe KE. 2013 Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality. *PLoS One* **8**, e76674. (doi:10.1371/journal.pone.0076674)

61. Crockford C, Wittig RM, Zuberbühler K. 2015 An intentional vocalisation draws others' attention: A playback experiment with wild chimpanzees. *Anim. Cogn.* **18**, 581–591. (doi:10.1007/s10071-014-0827-z)

62. Schel AM, Machanda Z, Townsend SW, Zuberbühler K, Slocombe KE. 2013 Chimpanzee food calls are directed at specific individuals. *Anim. Behav.* **86**, 955–965. (doi:10.1016/j.anbehav.2013.08.013)

63. Slocombe KE, Kaller T, Turman L, Townsend SW, Papworth S, Squibbs P, Zuberbühler K. 2010 Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behav. Ecol. Sociobiol.* **64**, 1959–1966. (doi:10.1007/s00265-010-1006-0)

64. Laporte MNC, Zuberbühler K. 2010 Vocal greeting behaviour in wild chimpanzee females. *Anim. Behav.* **80**, 467–473. (doi:10.1016/j.anbehav.2010.06.005)

65. Slocombe KE, Zuberbühler K. 2007 Chimpanzees modify recruitment screams as a function of audience composition. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 17228–33. (doi:10.1073/pnas.0706741104)

66. Clay Z, Pika S, Gruber T, Zuberbühler K. 2011 Female bonobos use copulation calls as social signals. *Biol. Lett.* **7**, 513–516. (doi:10.1098/rsbl.2010.1227)

67. Clay Z, Zuberbühler K. 2012 Communication during sex among female bonobos: effects of dominance, solicitation and audience. *Sci. Rep.*

68. Townsend SWS, Deschner T, Zuberbühler K. 2008 Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS One* **3**. (doi:10.1371/journal.pone.0002431)

69. Wich S, Sterck E. 2003 Possible audience effect in Thomas langurs (Primates; Presbytis thomasi): an experimental study on male loud calls in response to a tiger model. *Am. J. Primatol.*

70. Arnold K, Zuberbühler K. 2006 Language evolution: semantic combinations in primate calls. *Nature* **441**, 303. (doi:10.1038/441303a)

71. Berthet M, Neumann C, Mesbahi G, Cäsar C, Zuberbühler K. 2018 Contextual encoding in titi monkey alarm call sequences. , 1–11. (doi:10.1007/s00265-017-2424-z)

72. Coye C, Ouattara K, Zuberbühler K, Lemasson A. 2015 Suffixation influences receivers' behaviour in non-human primates. *Proc. R. Soc. B Biol. Sci.* **282**, 20150265–20150265. (doi:10.1098/rspb.2015.0265)

73. Schamberg I, Cheney DL, Clay Z, Hohmann G, Seyfarth RM. 2016 Call combinations, vocal exchanges and interparty movement in wild bonobos. *Anim. Behav.* **122**, 109–116. (doi:10.1016/j.anbehav.2016.10.003)

74. Zuberbühler K. 2000 Referential labelling in Diana monkeys. *Anim. Behav.* **59**, 917–927. (doi:10.1006/anbe.1999.1317)

75. Zuberbühler K. 2001 Predator-specific alarm calls in Campbell's monkeys, Cercopithecus campbelli. *Behav. Ecol. Sociobiol.* **50**, 414–422. (doi:10.1007/s002650100383)

76. Zuberbühler K. 2000 Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* **76**, 195–207. (doi:10.1016/S0010-0277(00)00079-2)

77. Arnold K, Zuberbühler K. 2013 Female Putty-Nosed Monkeys Use Experimentally Altered Contextual Information to Disambiguate the Cause of Male Alarm Calls. *PLoS One* **8**, 2–8. (doi:10.1371/journal.pone.0065660)

78. Wheeler BC, Hammerschmidt K. 2013 Proximate factors underpinning receiver responses to deceptive false alarm calls in wild tufted capuchin monkeys: Is it counterdeception? *Am. J. Primatol.* **75**, 715–725. (doi:10.1002/ajp.22097)

79. Laland KN. 2004 Social learning strategies. *Learn. Behav.* **32**, 4–14.

80. Seyfarth RM, Cheney DL. 1980 The Ontogeny of Vervet Monkey Alarm Calling Behavior: A Preliminary Report. *Z. Tierpsychol.* **54**, 37–56. (doi:10.1111/j.1439-0310.1980.tb01062.x)

81. Seyfarth RM, Cheney DL. 1986 Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640–1658. (doi:10.1016/S0003-3472(86)80252-4)

82. Hauser MD. 1988 How infant vervet monkeys learn to recognise starling alarm calls: the role of experience. *Behaviour* **105**, 187–201. (doi:10.1163/156853988X00016)

83. Hauser MD. 1989 Ontogenetic Changes in the Comprehension and Production of Vervet Monkey (Cercopithecus aethiops) Vocalisations. *J. Comp. Psychol.* **103**, 149–158. (doi:10.1037/0735-7036.103.2.149)

84. Lemasson A, Gandon E, Hausberger M. 2010 Attention to elders' voice in non-human primates. *Biol. Lett.* **6**, 325–328. (doi:10.1098/rsbl.2009.0875)

85. Evans CS. 1997 Referential Signals. pp. 99–143. (doi:10.1007/978-1-4899-1745-4_5)

86. Macedonia JM, Evans CS. 1993 Essay on Contemporary Issues in Ethology: Variation among Mammalian Alarm Call Systems and the Problem of Meaning in Animal Signals. *Ethology* **93**, 177–197. (doi:10.1111/j.1439-0310.1993.tb00988.x)

87. Townsend SW, Manser MB. 2013 Functionally referential communication in mammals: The past, present and the future. *Ethology* **119**, 1–11. (doi:10.1111/eth.12015)

88. Wheeler BC, Fischer J. 2012 Functionally referential signals: A promising paradigm whose time has passed. *Evol. Anthropol.* **21**, 195–205. (doi:10.1002/evan.21319)

89. Scarantino A, Clay Z. 2015 Contextually variable signals can be functionally referential. *Anim. Behav.* **100**, 1–8. (doi:10.1016/j.anbehav.2014.08.017)

90. Graham K, Wilke C, Lahiff N, Slocombe K. 2019 Scratching Beneath the Surface : Intentionality in Great Ape Signal production. *Philos. Trans. R. Soc. B Biol. Sci.*

91. Canteloup C, Bovet D, Meunier H. 2015 Intentional gestural communication and discrimination of human attentional states in rhesus macaques (Macaca mulatta). *Anim. Cogn.* **18**, 875–883. (doi:10.1007/s10071-015-0856-2)

92. Dennett D. 1987 The intentional stance, 1987.

93. Townsend SW *et al.* 2016 Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biol. Rev.* **92**, 1427–1433. (doi:10.1111/brv.12289)

94. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)

95. Schamberg I, Wittig RM, Crockford C. 2018 Call type signals caller goal : a new take on ultimate and proximate influences in vocal production. (doi:10.1111/brv.12437)

96. Cheney DL, Seyfarth RM, Smuts B. 2016 Social Relationships and Social Cognition in Nonhuman Primates. *Science (80-.).* **234**, 325–336. (doi:10.4000/books.editionsmsh.1745)

97. Silk JB, Seyfarth RM, Cheney DL. 2018 Quality versus quantity : do weak bonds enhance the fi tness of female baboons ? *Anim. Behav.* **140**, 207–211. (doi:10.1016/j.anbehav.2018.04.013)

98. Silk JB, Beehner JC, Bergman TJ, Crockford C, Engh AL, Moscovice LR, Wittig RM, Seyfarth RM, Cheney DL. 2010 Strong and consistent social bonds enhance the longevity of female baboons. *Curr. Biol.* **20**, 1359–1361. (doi:10.1016/j.cub.2010.05.067)

99. Seyfarth RM, Cheney DL. 2015 Social cognition. *Anim. Behav.* **103**, 191–202. (doi:10.1016/j.anbehav.2015.01.030)

100. Cheney DL, Seyfarth RM. 2008 *Baboon metaphysics: the evolution of a social mind*. The University of Chicago Press.

101. Seyfarth RM, Cheney DL. 2018 Pragmatic flexibility in primate vocal production. *Curr. Opin. Behav. Sci.* **21**, 56–61. (doi:10.1016/j.cobeha.2018.02.005)

102. Seyfarth RM, Cheney DL. 2016 Precursors to language: Social cognition and pragmatic inference in primates. *Psychon. Bull. Rev.* **24**, 1–6. (doi:10.3758/s13423-016-1059-9)

103. Seyfarth RM, Cheney DL, Bergman TJ. 2005 Primate social cognition and the origins of language. *Trends Cogn. Sci.* **9**, 264–266. (doi:10.1016/j.tics.2005.04.001)

104. Candiotti A, Zuberbühler K, Lemasson A. 2013 Voice discrimination in four primates. *Behav. Processes* **99**, 67–72. (doi:10.1016/j.beproc.2013.06.010)

105. Wilson DE, Reeder DM. 2005 *Mammal Species of the World: A Taxonomic and Geographic Reference*. JHU Press.

106. Struhsaker TT. 1967 Ecology of vervet monkeys (Cercopithecus aethiops) in The Masai-Amboseli Game Reserve, Kenya. *Ecology* **48**, 891–904. (doi:10.2307/1934531)

107. Hall KRL, Gartlan JS. 1965 Ecology and Beaviour of the Vervet Monkey, Cercopithecus Aethiops, Lolui Island, Lake Victoria. *Proc. Zool. Soc. London* **145**, 37–56. (doi:10.1111/j.1469-7998.1965.tb01999.x)

108. Butynski TM, de Jong YA. 2019 Chlorocebus pygerythrus. The IUCN Red List of Threatened Species 2019: e.T136271A17957823. **2**, 277–283. (doi:https://dx.doi.org/10.2305/IUCN.UK.2019- 3.RLTS.T136271A17957823.en)

109. Delson E. 1982 Catalogue of primates in the British Museum (natural history) and elsewhere in the British isles. Part II: Family cercopithecidae, subfamily cercopithecinae.

110. Cheney DL, Seyfarth RM. 1983 Nonrandom dispersal in free-ranging vervet monkeys: social and genetic consequences. *Am. Nat.* **122**, 392–412. (doi:10.1086/284142)

111. Cheney DL, Seyfarth RM. 1990 The representation of social relations by monkeys. *Cognition* **37**, 167–196. (doi:10.1016/0010-0277(90)90022-C)

112. Seyfarth RM, Cheney DL. 1984 Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature* **308**, 541–543. (doi:10.1038/308541a0)

113. Silk JB, Alberts SC, Altmann J. 2003 Social Bonds of Female Baboons. *Science (80-.).* **302**, 1231–1235. (doi:10.1126/science.1088580)

114. de Waal FBM. 2007 *Chimpanzee politics: power and sex among apes*. JHU Press.

115. Arseneau-Robar TJM *et al.* 2018 Male monkeys use punishment and coercion to de-escalate costly intergroup fights. *Proc. R. Soc. B Biol. Sci.* **285**. (doi:10.1098/rspb.2017.2323)

116. Arseneau-Robar TJM, Taucher AL, Müller E, Van Schaik CP, Bshary R, Willems EP. 2016 Female monkeys use both the carrot and the stick to promote male participation in intergroup fights. *Proc. R. Soc. London B Biol. Sci.* **283**. (doi:10.1098/rspb.2016.1817)

117. Borgeaud C, Bshary R. 2015 Wild Vervet Monkeys Trade Tolerance and Specific Coalitionary Support for Grooming in Experimentally Induced Conflicts. *Curr. Biol.* **25**, 3011–3016. (doi:10.1016/j.cub.2015.10.016)

118. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey (Cercopithecus aethiops) alarm calls: semantic communication in a free ranging primate. *Anim. Behav.* **28**, 1070–1094.

119. Price T, Wadewitz P, Cheney DL, Seyfarth RM, Hammerschmidt K, Fischer J. 2015 Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* **5**, 1–11. (doi:10.1038/srep13220)

120. Mercier S, Neumann C, van de Waal E, Chollet E, Meric de Bellefon J, Zuberbühler K. 2017 Vervet monkeys greet adult males during high-risk situations. *Anim. Behav.* **132**, 229–245. (doi:10.1016/j.anbehav.2017.07.021)

121. Mercier S, Déaux EC, van de Waal E, Bono AEJ, Zuberbühler K. 2019 Correlates of social role and conflict severity in wild vervet monkey agonistic screams. *PLoS One* **14**, e0214640. (doi:10.1371/journal.pone.0214640)

122. Seyfarth RM, Cheney DL. 2010 Some general features of vocal development in nonhuman primates. *Soc. Influ. Vocal Dev.* , 249–273. (doi:10.1017/cbo9780511758843.013)

123. Cheney DL, Seyfarth RM. 1980 Vocal recognition in free-ranging vervet monkeys. *Anim. Behav.* **28**, 362–367. (doi:10.1016/S0003-3472(80)80044-3)

124. Altmann J. 1974 Observational study of behavior: sampling methods. *Behavior* **49**, 227–267. (doi:10.1163/156853974X00534)

125. van de Waal E, Borgeaud C, Whiten A. 2013 Potent Social Learning and Conformity Shape a Wild Primate's Foraging Decisions. *Science (80-.).* **340**, 483–485. (doi:10.1126/science.1232769)

126. Botting J, Whiten A, Grampp M, van de Waal E. 2018 Field experiments with wild primates reveal no consistent dominance-based bias in social learning. *Anim. Behav.* **136**, 1–12. (doi:10.1016/j.anbehav.2017.11.025)

127. Kendal RL, Boogert NJ, Rendell L, Laland KN, Webster M, Jones PL. 2018 Social Learning Strategies: Bridge-Building between Fields. *Trends Cogn. Sci.* **22**, 651–665. (doi:10.1016/j.tics.2018.04.003)

128. Griffin AS. 2004 Social learning about predators: A review and prospectus. *Learn. Behav.* **32**, 131–140. (doi:10.3758/BF03193188)

129. Jhang Y, Oller DK. 2017 Emergence of functional flexibility in infant vocalisations of the first 3 months. *Front. Psychol.* **8**. (doi:10.3389/fpsyg.2017.00300)

130. Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, Bakeman R. 2013 Functional flexibility of infant vocalisation and the emergence of language. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6318–23. (doi:10.1073/pnas.1300337110)

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

Chapter 2: One-trial social learning of vervet monkey alarm calling

Adwait Deshpande [1,2,*], Bas Van Boekholt [2,3] and Klaus Zuberbuhler [1,2,4]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchatel, Neuchatel, Switzerland, [2] Inkawu Vervet Project, Mawana Game Reserve, KwaZulu-Natal, South Africa, [3] Animal Ecology, Department of Biology, Utrecht University, Utrecht, Netherlands, [4] School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom

(Manuscript under review: Royal Society Open Science)

2.1 Abstract

How do non-human primates learn the usage of their alarm calls? Social learning is a promising candidate, but its role in the acquisition of meaning and call usage has not been studied systematically, neither during ontogeny nor in adulthood. To investigate the role of social learning in alarm call comprehension and use, we exposed groups of wild vervet monkeys to two unfamiliar animal models in the presence or absence of conspecific alarm calls. To assess the learning outcome of this experience, we presented the models again to individual monkeys without additional information. In subjects previously exposed with alarm calls, we found heightened predator inspection levels compared to control subjects, indicating one-trial social learning of 'meaning', i.e., call-context associations. Moreover, juveniles (but not adults) gave alarm calls like the conspecific alarm calls they heard during the first exposure. However, this effect was only found for one animal model, revealing how the subject's age and core knowledge of potential predators moderate the efficacy of social learning in acquiring novel call-referent associations.

2.2 Keywords

Primate vocal communication; Predator recognition; Call comprehension; Meaning attribution; Playback experiment; Fast mapping

2.3 Introduction

Few mammalian species are known for highly flexible vocal communication systems, such as cetaceans and humans [1]. Whereas our evolutionary neighbours—the non-human primates—are constrained by small call repertoires and limited modification of the call structure [2]. However, non-human primates have more control and flexibility in call usage, especially in their alarm calls. For example, several simian species flexibly use various alarm call types and combinations in different contexts [3–7]. Wild chimpanzees (*Pan troglodytes*) produce more alarm calls in the presence of ignorant than knowledgeable group members, suggesting socially aware alarm call production [8,9]. Furthermore, non-human primates sometimes use a particular type of alarm call in multiple contexts, such as during predator encounters or between-group aggression, requiring receivers to make a correct inference, and adjust their responses appropriately [10,11]. The adaptive value of correct inference and production of calls is self-evident, especially during risky situations. Yet, it is largely unclear how individuals acquire the necessary competence, both in alarm call comprehension and production.

A first possibility is that alarm calling behaviour is mainly hard-wired and develops without significant environmental input. This is, however, highly unlikely because the variability in environments and contexts is simply too large for any kind of hard-wired system to be sufficient [12]. A second way to acquire competence in alarm calling behaviour is through individual learning by trial-and-error. However, committing errors during predator encounters can be very costly, and it is unclear what kind of feedback would foster the formation of novel call-context associations. Finally, social learning, that is, learning from observing conspecifics [13], provides the third major mechanism. As the predation risk can change over time and space, with novel predators suddenly appearing or existing ones disappearing (e.g. [14]), social learning can be an efficient mechanism to avoid the high fitness costs in continuously changing environments [15–17]. In line with this, social learning has already been established as an important mechanism for predator recognition and avoidance in non-primate taxa [Reviewed in [15]]. For example, in a classic study, European blackbirds began to mob an innocuous bird model or even plastic bottles when observing other conspecifics mobbing [18].

Although the role of social learning has been explored in non-human primates, current social learning literature is mainly concerned with how animals acquire arbitrary food preferences or learn new manipulative actions [19,20], despite the fact that learning from others is likely to be important in all major biological functions. A few early primate studies have investigated socially acquired predator recognition, but they were mainly conducted in captive conditions [21,22] with little known about the wild (but see [23]). In the context of communication, especially for alarm calls, an unresolved problem is whether and how non-human primates socially learn the meaning of calls, that is, the kind of inferences that can be made from them [24]. We attempted to close this gap by experimentally testing whether conspecific alarm calls facilitate social learning in wild vervet monkeys and whether this leads to new call-context associations.

South African vervet monkeys (*Chlorocebus pygerythrus pygerythrus*) are an ideal model system for such investigations. The classic studies on the Kenyan sub-species (*C. P. hilgerti*) showed that vervet monkeys produce at least three acoustically distinct alarm calls for corresponding predator categories of aerial threats, terrestrial predators and snakes [25,26]. A recent study revealed that alarm calls for terrestrial predators are also used in other contexts like intra- and inter-group aggression, at least in the South African sub-species [27]. Furthermore, based on their observations, Cheney and Seyfarth have proposed an influential model for the development of alarm call use in vervet monkeys [28], which suggests that infants can produce adult-like alarm calls from early on but need to learn their correct use. In particular, aerial-like alarm calls are given from the beginning to a broad range of flying or falling objects, including non-predatory birds, but not to terrestrial predators, such as leopards, suggesting a hard-wired link between basic predator types or event categories and types of alarm calls (see [29] for similar patterns in green monkeys (*Chlorocebus sabaeus*) for a novel predator). Cheney and Seyfarth also observed that the within-category usage gets refined over time such that juveniles produce aerial alarm calls only for specific predatory bird species, such as martial eagle (*Polemaetus bellicosus*). Anecdotally, the authors also observed that juveniles who looked at adults before responding to alarm calls tended to respond correctly, indicating a role of social learning [28,30].

We capitalised on this well studied system and exposed groups of free-ranging South African vervet monkeys to two novel animal models (commercially available soft toys for toddlers). In a crossed experimental design, we either provided the information about the two models by broadcasting alarm calls of a known adult group member during the first encounter with the model or withheld the information by not playing any call or playback of a non-alarm call. We predicted that if social learning took place, individuals from the informed group should show increased predator inspection during a subsequent encounter with the model compared to individuals from the control group. We also predicted that socially acquired comprehension would lead to competence in production, such that individuals from the informed group should be more likely to produce alarm calls during a subsequent encounter with the model compared to individuals from the control group. We took into account the effects of the predator model type and subject age since both variables could impact acquisition.

2.4 Methods

2.4.1 Study site and groups

The study was conducted at the Inkawu Vervet Project (IVP) on the Mawana Game reserve, Kwazulu Natal, South Africa (S 28°00.327; E 031°12.348) from November 2018 to July 2019. At the IVP, six groups of wild vervet monkeys are well habituated to human presence and allow close observations by multiple observers. All the observers go through standard training and pass a test for individual identification before participating in data collection. Individuals were fed regularly on corn kernels for various experiments. Food provisioning events were announced using specific calls by observers, which are understood by monkeys attracting them to the location of the food [20,31,32]. The experiments presented here were conducted on the following three groups (1) BD (Group size= 29 adults, 18 Juveniles), (2) AK (Group size= 5 adults, 14 Juveniles), (3) NH (Group size= 17 adults, 20 Juveniles). As per IVP protocols, females are considered juveniles until they give birth to an infant for the first time, and males are considered juveniles until they immigrate from their natal group. Individuals are classified as infants till they are weaned. Thus, approximately the age of individual classified as infants = 0-12 months, Juveniles = 1 year to 4 years.

2.4.2 Experimental setup

We conducted four blocks of experiments on the two groups of monkeys, the BD and the AK group. Each experimental block consisted of two parts: Group exposures and Individual assessment trials. Group exposures consisted of four different treatments; two animal models (caterpillar vs horse, see below) by two information states (alarm call playback vs no playback) (see below for additional control treatment with Grunt playback). We cross-balanced the treatments of alarm playback (Alarm) and no playback (Control) with the two animal models (Table 2.1).

Alarm calls used for the playback were recorded previously from the adult males of each group, produced naturally for approaching jackals (*Lupulella mesomelas*) or feral dogs, which are common terrestrial mammalian predators in the study area. Acoustically, these alarm calls were very similar to the 'leopard' alarm calls of Amboseli vervet monkeys and henceforth labelled as such for convenience [26]. Playback stimuli comprised 10-14 leopard alarm calls at natural intervals clustered into 3-4 bouts. Stimuli were presented with a Nagra Kudelski DSM-Monitor speaker played through a Huawei ALE-L21 phone. Speaker was hidden under a blanket as well as behind a bush approximately 10-15m away from the cylindrical hide (see below). We pre-adjusted the speaker amplitude similar to a natural call at the appropriate distance. Original stimuli were edited such that the amplitude of other background sounds was reduced, and all the call tracks were padded at both ends by 2 seconds of silence using Audacity 2.3.3[33].

Table 2.1 Sample Size for individual assessment trials across treatments for the study groups AK, BD and NH

Treatment	Predator	Sample size (n) for individual assessments for each study group
Control (No playback)	**Caterpillar**	**AK** n=6
Alarm (Leopard alarm playback)	**Caterpillar**	**BD** n=8
Alarm (Leopard alarm playback)	**Horse**	**AK** n=6
Control (No playback)	**Horse**	**BD** n=6
Grunt (Grunt playback)	**Caterpillar**	**NH** n=5

We were interested in how novel call-context associations are formed and how this leads to competence in comprehension and production of alarm calls. For this, we used unfamiliar animal-like models as stimuli, which were completely alien to the monkeys. We chose two commercially available soft toys for toddlers. An artificially coloured, oversized caricature of an insect larva ('caterpillar') and a miniaturised caricature of an ungulate ('horse', Figure 2.1). Models were comparable in size (Length = Caterpillar (~100cm), Horse (~70cm), Breadth/diameter of both the models ~20 cm) but differed in their animacy. When pulled, the 'caterpillar' model made an animate movement with mobile legs, whereas the 'horse' model toppled sideways and could only be dragged inanimately.

The experimental arena consisted of a blanket and a camouflaged cylindrical hide placed approximately 10m apart on the ground. The predator model was hidden under the blanket, attached with a fishing line controlled by the experimenter, positioned at least 2m behind the hide. A trial began with the experimenter pulling the predator model out of the blanket and dragging it in a straight line across the arena to disappear inside the cylindrical hide. A tripod with a video camera was placed to record all the events in the arena (Figure 2.2 A). We accustomed both groups to the experimental arena (without models) by repeatedly (5-8 times) placing the setup in the group's travelling path until all individuals ignored the setup.

Figure 2.1: Experimental setups showing two different predator models (A) Caterpillar (B)

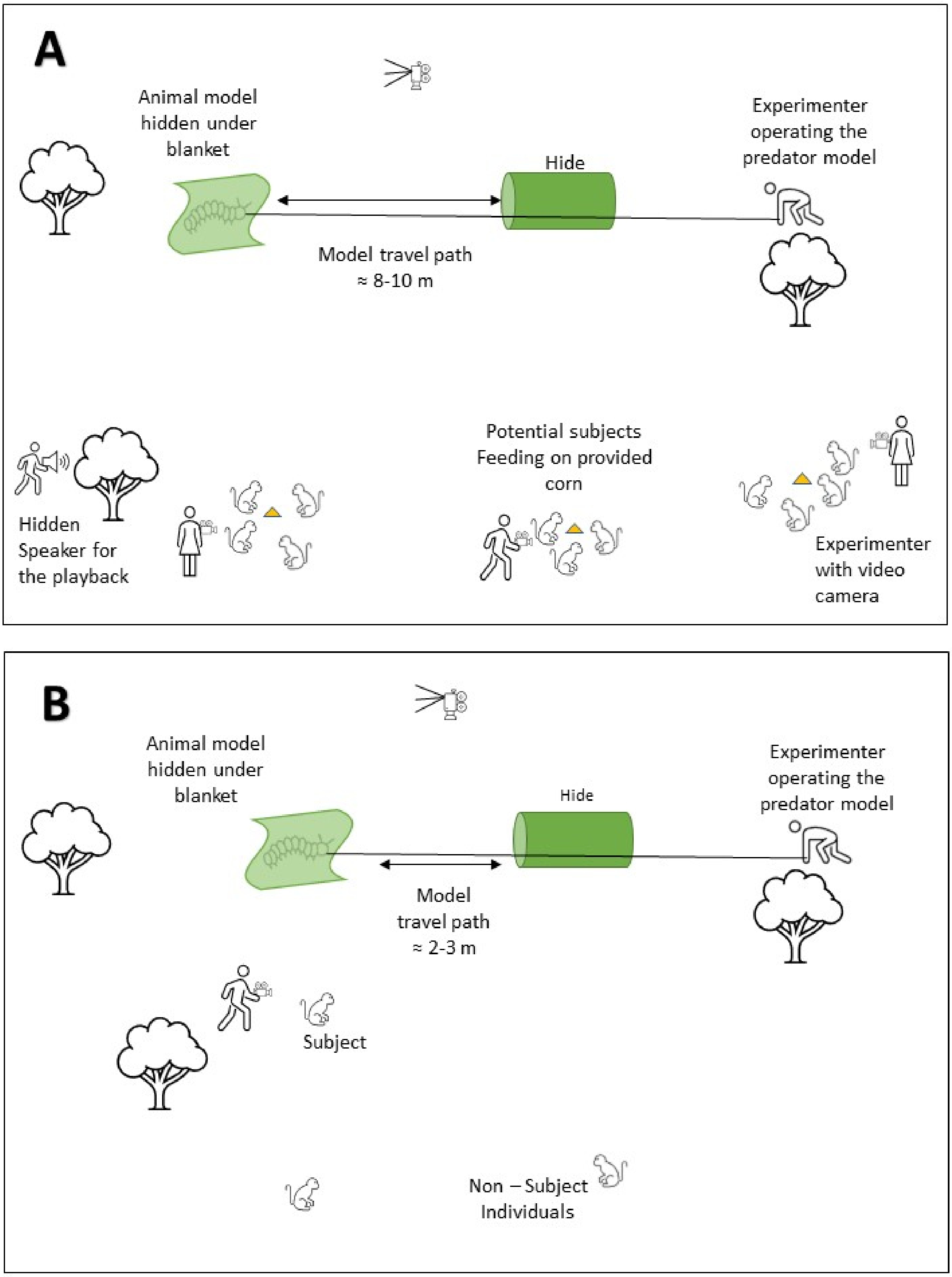

Figure 2.2: Heuristic illustrations (not to scale) of experimental setups for different trials

2.4.3 Experimental procedure

An experimental block started with a group exposure on the first day, followed by individual assessment trials on subsequent days (see Appendix). Group exposures were conducted after the monkeys left their sleeping sites in the morning. We set up the experimental arena in advance on the group's most likely travel path and attracted individuals with food calls to small mounds of corn kernels placed 5-10m away from the setup. For the Alarm treatment, once most of the monkeys were present at the mounds, we played the alarm call stimulus from the hidden speaker. We simultaneously moved one of the predator models across the arena as described above (Figure 2.2 A). The exposure lasted approximately 20-30 seconds. Group exposures in the control condition were precisely the same, except there were no playbacks. All group exposure trials were recorded on video using a JVC quadproof EverioR camcorder. The experimenter was assisted by 2-4 observers who identified individuals present during the first seconds of the exposure and recorded which individuals certainly saw the predator model by determining their clear head orientation towards the model. These individuals were considered as 'informed' and thus became subjects for the subsequent assessment trials. Twenty-six individuals opportunistically qualified for subject status across the treatments (Adults females = 4, Adult males =4, Juveniles females = 8, Juvenile males= 10). Once the predator model was not visible anymore, some individuals of the group usually returned to forage on the provided food.

In each block, we conducted one assessment trial per subject per day starting on the day after a group exposure. This design was adapted to minimise the confounding effects of natural predator encounters and other natural alarm call events. The frequency of natural terrestrial predator encounters at this site can vary significantly from once a day to one per week (AD personal obs). During individual assessment trials, we set up the arena and presented the model in the same way, but this time only to one subject, albeit with no playback and reduced predator exposure time (Mean 9.6s; SE ±1.1s) by reducing the distance between the blanket and the hide from 10m to about 2.5m (Figure 2.2 B). We exposed the model only when the subject was looking in the general direction of the model and the hide, which made sure the model was easily noticed by the subject for each trial. For one trial, a model was not noticed, and we hid the model in the cylinder immediately. A re-trial on the same subject was conducted later the same day. Individual assessment trials on the subjects were conducted opportunistically, making sure that no other subject was present in the vicinity at the time of the trial. We noted the distance between the subject and the experimental arena. Subjects were recorded for 10 minutes, starting with the model exposure. We chose a longer time window to monitor the behaviours because in a pilot exposure to another group, monkeys inspected, and an alarm called at the hide for more than 5 minutes. We also audio recorded all alarm calls and other vocalisations elicited by the subject using a solid-state audio recorder (Marantz PMD 661) and a directional microphone (Sennheiser MKH 416 P48) with a sampling frequency of 44.1 kHz (see supplementary videos).

Our initial study design included another control condition, which consisted of a (non-alarm call) grunt playback treatment (Grunt). Here, with the same experimental procedure, we played back a grunt to a third group (NH) paired with the 'caterpillar' model, followed by five individual assessments. The playback call was elicited by a subordinate male to the dominant male of the NH group [34]. However, due to speaker malfunction during exposure, the group exposure on another group with the 'horse' model was compromised as the individuals were exposed to the model without any playback. The treatment could not be repeated in the other lesser habituated groups due to time constraints, creating an imbalance in the experimental design. Hence, we did not include the Grunt treatment results in the primary analyses but presented them as a subsidiary analysis for comparison (see statistical analysis).

2.4.4 Behaviour coding

All the individual assessment videos were coded using the video coding software BORIS v. 6.3.9 [35]. We coded all the behaviours elicited during the 10 minutes from the exposure to the predator model. We were mainly interested in the following behaviours (1) Predator inspection (s): total time spent looking directly at the predator during exposure and/or time spent looking directly at the cylindrical hide after the exposure; (2) number of alarm calls (N).

To assess the inter-observer reliability of the behaviour, a second coder, experienced in primate behaviour but blind to the aims and objectives of the study, coded 20% of the videos, which were randomly selected. For predator inspection (a continuous variable), we found a strong Spearman correlation ($\rho = 0.83$, $p = 0.015$), suggesting that data were coded in sufficiently comparable ways.

2.4.5 Acoustic analysis and classification

We extracted a subject's alarm calls during the individual assessments from the audio recordings using Audacity 2.3.3 [33] for acoustic analyses after cross-checking them on the videos. Overlapping calls were removed from further analyses. We also included previously recorded alarm calls by AD. These calls were recorded during natural predator encounters of jackals (*Lupulella mesomelas*), python (*Python sebae natalensis*) and different species of raptors and corresponded to leopard, snake and eagle alarms, respectively, as described in earlier studies and henceforth labelled as such (see Appendix) [25–27]. A call bout was defined as a coherent acoustic utterance, separated by at least 500ms from another. Alarm call bouts could contain several calls separated by less than 500ms from each other, marked in Raven Pro 1.5 [36] using Hamming window at 1024 DFT and 93.8% overlap. Acoustic parameters were extracted at the call level and submitted for subsequent statistical modelling. We extracted a total of 117 calls from 48 call bouts and initially extracted 23 acoustic parameters from each call.

To avoid potential coding biases, we used Linear Discriminant Analysis (LDA) to classify the calls elicited by the subjects during the individual assessments. First, we scaled the data and reduced the dimensionality by removing highly correlated (>0.95) acoustic parameters and used the remaining 16 parameters for the LDA (see Appendix). Naturally recorded alarm calls from (N = 40) were used as training data for the LDA model with equal prior probabilities. We then predicted the call types for each call recorded in the individual assessment trials based on the LDA model. LDA was performed using the lda function in the package MASS [37] function in R v. 3.6.3 [38].

2.4.6 Statistical analyses

We fitted generalised linear models with the lme4 package [39] to test the effects of the Alarm and Control treatments on the two behaviours elicited during individual assessments, predator inspection (s) and the total number of alarm calls (N) [40–42]. For predator inspection, we log-transformed the data to use a linear model; for alarm calls, we accounted for zero inflation and used negative binomial models. We controlled for type of predator model, group identity, age of subject (juvenile or adult) as categories; and the order of trial, the distance of the subject from the predator at the start of exposure and exposure time as Z transformed variables by adding them as fixed effects in all the models. To test for the significance of each fixed effect, we used the analysis of variance or deviance depending on the type of full model. We visually checked for heteroscedasticity and then verified that variance-inflation factors were less than four for all the fixed effects in the full models [43]. Effect sizes were determined using the MuMIn package [44]. To assess the effects of the Alarm treatment on alarm calls produced by juveniles to the caterpillar model, we did a post hoc analysis by fitting a negative binomial model on the number of leopard and snake calls on a subset of data. The fixed effects were the same as earlier models except for the predator model, and group identity were removed as fixed effects.

We ran a separate non-parametric analysis to compare the difference between Alarm and Grunt treatments with Wilcoxon tests to address the possibility that changes in predator inspection could be caused by any vocalisation uttered during the model exposure. Finally, the

Grunt treatment did not lead to any alarm calls, so no further comparisons were made. All analyses were conducted in R v.3.6.3[38].

2.5 Results

2.5.1 Predator inspection

Mean predator inspection observed across the treatments was 57.9 sec (SE ± 10.5, Range 6.6 – 228.2). Irrespective of age, monkeys inspected both animal models for significantly more duration in the individual assessments following Alarm compared to Control treatments (N = 26, F = 8.36, df = 1,18; p = 0.009, R^2 = 0.39; Figure 2.3). There was a trend of 'caterpillar' model triggering more inspection than the 'horse' model, but this was not statistically significant (N = 26, F = 3.91, df = 1,18; p = 0.063). In a subsidiary analysis, we also compared predator inspection following Alarm and Grunt treatments which revealed that predator inspection was significantly higher in the Alarm treatment (N=19, W = 12, p= 0.033; Wilcoxon rank-sum test, see Appendix).

2.5.2 Alarm calls

Subjects produced N=77 alarm calls during individual assessments. N=37 were classified as snake alarms, N=3 as eagle alarms and N=37 as leopard alarms by the LDA (Table 2.2).

Regarding the total number of alarm calls we found no effect of treatment but differences in model type and subject age: the ‘caterpillar’ elicited more alarms than the ‘horse’ model and juveniles produced more alarms than adults (N = 26, df = 1, Model type: X^2 = 4.92, p = 0.026; Age: X^2 = 4.54, p = 0.032; Treatment: X^2 = 0.01, p = 0.89, R^2 = 0.65; Figure 2.4). Most of the alarm calls were given to the 'caterpillar' model by the juveniles (68% of all alarm calls; Table 2.2).

In a post-hoc analyses, only considering the ‘caterpillar’ model and juvenile subjects, we found that the Alarm treatment elicited more leopard alarm calls than the Control during individual assessments (N = 12, df = 1, X^2 = 8.69, p = 0.003, R^2 = 0.47; Figure 2.5). Leopard alarms also increased with the subjects’ distance from the model (N = 12, df = 1, X^2 = 9.18, p = 0.002; Figure 2.5). Snake calls, on the other hand, were not significantly affected, neither by treatment type nor by any other factor (N = 12, df = 1, X^2 = 0.11, p = 0.73, R^2 = 0.18; see Appendix).

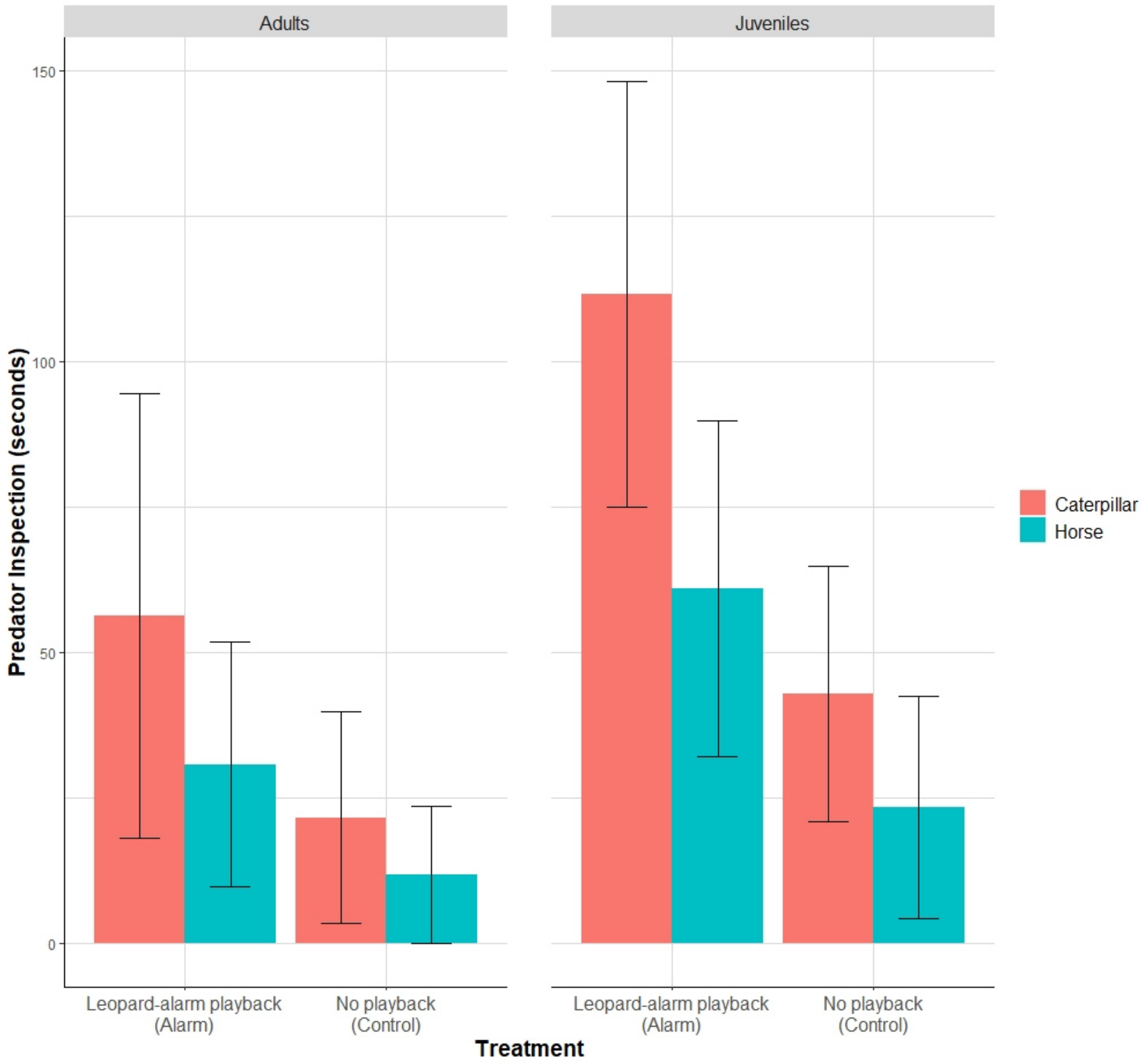

Figure 2.3Model predictions (± SE) of the effects of treatments on the predator inspection of the subjects grouped by age categories and predator model type

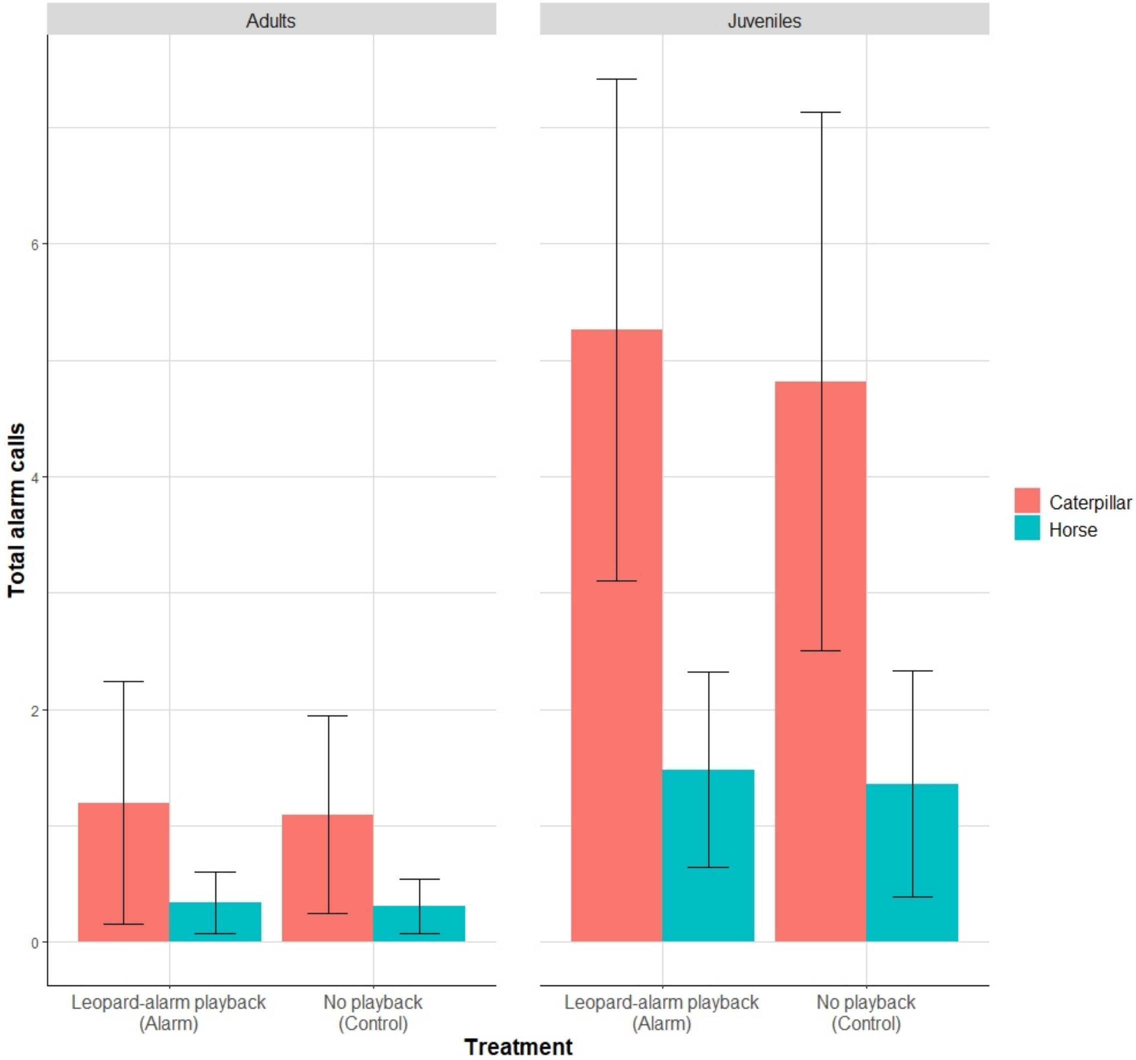

Figure 2.4: Model predictions (± SE) of the effects of treatments on total alarm calls of the subjects grouped by age categories and model type

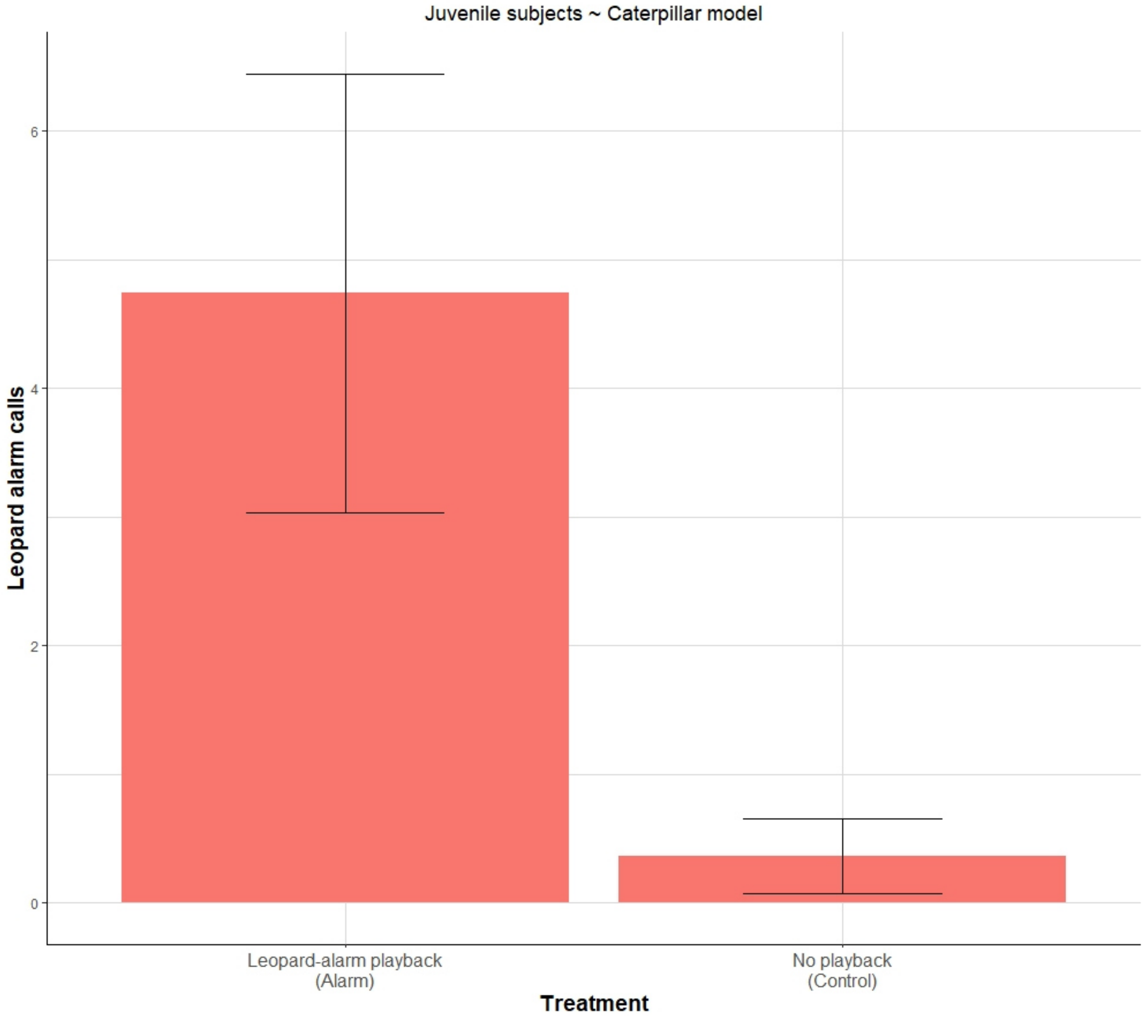

Figure 2.5: Model predictions (Mean ± SE) of the effects of treatments on the leopard alarms for the juvenile subjects exposed to caterpillar model

Table 2.2 Table showing Mean ±Standard Error (SE) of the behaviours for the data collected during individual assessment trials across treatments (except grunt treatment) and age categories

Treatment & sample size (n) of individual assessments per treatment		**Alarm Leopard playback** n=14		**Control No playback** n=12	
Animal Model		Caterpillar	Horse	Caterpillar	Horse
Predator inspection (Seconds) Mean ± SE, Sample size (n)	Juveniles	114.21 ±26.7, n=7	42.17±11.3, n=4	40.18 ±11.5, n=5	28.53±9.2, n=2
	Adults	20.27, n=1	66.9±37.7, n=2	52.2, n=1	18.13±6.5, n=4
Leopard alarms Mean ± SE	Juveniles	2.71±1.0	1.50±0.5	1.8 ± 1.0	1.00 ±1.0
	Adults	0	0	1	0
Snake alarms Mean ± SE	Juveniles	1.85±0.6	1.75±1.1	2.20± 0.8	1.00±1.0
	Adults	0	0	4	0
Eagle alarms Mean ± SE	Juveniles	0.14±0.1	0.25±0.2	0	0
	Adults	0	0	1	0
Total alarms Mean ± SE	Juveniles	4.71±1.4	3.50± 1.7	4.00±1.4	2.00±2.0
	Adults	0	0	6	0

2.6 Discussion

Humans acquire language socially by observing and interacting with others [45]. We investigated whether social learning was equally important in how non-human primates acquire their communicative competence. We regarded this as a non-trivial problem due to the fact that primate communication and human language differ drastically in scope, flexibility and power of the signal repertoire. While non-human primate communication is highly constrained to small repertoires of largely inflexible, species-specific signals, humans have evolved a second communication system based on the ability to precisely control and shape their vocal tracts and acoustic output in highly intricate ways [1]. Despite these differences, when acquiring competence, do primates employ the same underlying social learning processes? How do primates learn to associate existing signals with the appropriate referents (here: external stimulus like a predator model)?

We found that wild vervet monkeys that heard another's alarm calls whilst encountering an unfamiliar animal model exhibited a number of behavioural changes in how they reacted to the model in a future encounter. Compared to control subjects, monkeys showed increased predator inspection towards the model, while the effect on alarm call production was only seen in juveniles. We tested two models and found that one (the 'caterpillar') model was more effective than the other (the 'horse') in eliciting alarm calls. There was also a similar trend in predator inspection behaviour. This may be due to several factors, including the fact that the 'caterpillar' was more animate and also resembled an already familiar class of predators (snakes) more than the 'horse'. Nevertheless, juvenile subjects in the Alarm condition for the caterpillar model produced leopard alarm calls during individual assessments.

What type of social learning took place? The current taxonomies of social learning theory (e.g. local enhancement, emulation, imitation) are mainly helpful when animals acquire manipulation skills but are less useful in addressing the acquisition of communication skills and knowledge more generally. Here, a nomenclature established to differentiate processes of vocal learning [46] appears to be more effective. At a basic level, conspecific alarm calls can help naïve individuals to associate a novel context with an extant alarm call, a process called comprehension learning. At the higher level, similar to the manual food-extraction tasks from artificial fruits, individuals can also learn to produce pre-existing alarm calls in a new context as a result of experiencing the novel usage by other conspecifics, termed usage learning [46]. It should be noted that definitions of the contexts can vary between the cases. It can be a broader social context like aggression, or a more specific environmental context, like a novel predator.

In line with this terminology, the significant increase in predator inspection in subjects from the Alarm treatment group, irrespective of age, can be considered as evidence for rapid comprehension learning. In other words, subjects seem to attribute novel meaning to existing alarms through social learning in a single exposure [24,47]. This capacity of non-human primates to form links between alarm calls and novel visual stimuli, even after a single exposure, is akin to fast mapping seen in children and also recorded in one domestic dog [48,49]. Similarly, a recent study in captive settings demonstrated that great apes encode events into long-term memory after a single experience [50]. Theoretical models predict that such rapid formations of call-context associations through social learning could be especially important in continually changing environments, where more standard models of learning requiring repeated exposures are less useful [17,51]. In the real world, changes in predatory species are not always a slow or rare process [52]. For example, in the Mawana game reserve, where we conducted this study, common terrestrial predators of monkeys are caracals (*Caracal caracal*) and jackals (*Lupulella mesomelas*). However, a pack of African wild dogs (*Lycaon pictus*) and subadult male lions (*Panthera leo*) have transiently visited Mawana over the years (AD personal obs). Increased inspection of novel models could help learn about potential predator behaviour and avoid extreme costs associated with the mistakes when responding to a predator. If it is true, such rapidly developed irrational fears should also extinct quickly with repeated unthreatening exposures with the same models. This prediction needs to be tested in future studies. Our results are consistent with studies on birds, marsupials and other primate species [22,53,54]. Rapid learning is also essential during development. Infant primates are born into a local fauna that can vary drastically between locations, requiring individuals to categorise each animal species it encounters as either dangerous or harmless. Given the scale of the task, it is clear that acquiring this competence is highly unlikely to obtain without some form of social learning.

Furthermore, in a subsidiary analysis, we compared the effect of alarm call with grunt call treatments which revealed that increased predator inspection was specific to the alarm call treatment. Unlike the grunts used in affiliative social interactions [34], leopard alarms might contain some information such as the urgency or nature of the threat [55]. Alternatively, leopard alarms may not contain deterministic information per se but probabilistic information about a threat that triggers a visual search. Once the potential threat is located, individuals form an association. The latter explanation of making pragmatic inference seems plausible since leopard alarm calls in this species are used in other multiple, potentially dangerous contexts such as inter and intra-group aggression [27]. However, due to imbalance with the Grunt treatment, pragmatic inference interpretation should be considered suggestive. Nevertheless, it should be noted that recent computational models also support a similar mechanism of social learning through associative processes for predator avoidance [56].

Alternatively, a simpler social learning process of stimulus enhancement because of the conspecific call can also explain the increased inspection towards the models in the 'alarm' treatment [17,57]. Either way, individuals seem to acquire meaning and retain socially acquired knowledge about a potential predator with a predator's single exposure.

Our post-hoc analysis on the subset of juveniles exposed to the 'caterpillar' model shows that subjects elicited more leopard alarms if they have heard leopard alarms during the group exposure. This suggests usage learning of alarm calls by juveniles. However, not all juveniles equally produced leopard alarms or even elicited an alarm call. This might reflect individual differences in learning [58]. On the other hand, it is also possible that call usage learning maybe not as rapid as formations of call-context associations. For example, grey seals (*Halichoerus grypus*) learned to produce calls on command in captive settings; however, it took multiple trials [59]. This result suggests that social learning as a mechanism is not limited to the manipulative task but also helps to refine the usage of pre-existing alarm calls, at least in juveniles. Leopard alarm calls by juveniles for the 'caterpillar' model were also affected by the distance between the subject and the model at the start. We speculate that it reflects the perceived safety of the individuals before the alarm call. Alarm calling is a conspicuous behaviour and can attract predator attention; thus, a safe distance from the predator can reduce the associated risk [60]. Also, the juveniles elicited fewer alarm calls when they saw the model for a longer time. More model exposure might have provided the subjects with cues about the harmlessness of the model, which could have resulted in a few alarm calls. Though, these speculations need to be examined empirically.

It is also important to note that the 'caterpillar' model consistently elicited snake alarms across treatments, making it the 'default' alarm call for this type of model. Perhaps, monkeys perceived a resemblance with predators they were already familiar with (in this case, pythons) in terms of shape, movement pattern and encounter location, which might have biased categorisation and alarm call type [61,62]. However, this suggestion is in contrast with the recent findings on congeneric green monkeys (*Chlorocebus sabaeus*). These primates were never observed alarm calling to birds of prey but instantly produced eagle-like alarm calls to an artificial flying drone, suggesting a hard-wired link between eagle alarms and aerial threat [29]. It is not possible to reject either of the two explanations from our results. Future experiments using ambiguous predator models of different predator categories in combination with playback of different alarm call types are needed to examine both possibilities.

The difference in alarm calls between two predator models across the treatments and some indication of more inspection towards the 'caterpillar' model is noteworthy. The 'caterpillar' model consistently elicited heighten responses. Although both the models have eyes and are comparable in sizes, they differed in the aspects of animacy. The 'caterpillar' model made more enliven movements with moving legs and was fluffier in texture, whereas the 'horse' model impassively dragged on the ground during exposure. It is possible that these differences in the animacy of the models could account for the differences in the anti-predatory responses. Similar results were found in captive blackbirds (*Turdus merula*), which were more scared of realistic-looking non-predatory bird model than the plastic bottles and wild jackdaws (*Corvus monedula*), which were more attentive to the unexpected animacy of the objects [18,63]. In primates, including humans, the categorisation of objects as animate or inanimate develops early in life or could even be innate. It constitutes what is known as 'initial' or 'core knowledge' [64,65]. Our results show that anti-predatory responses to the models seem to be based on integrating core knowledge and socially acquired knowledge [66].

We faced several limitations related to fieldwork. Notably, the subject selection was opportunistic. Hence, although part of the same age category, the age of juveniles ranged from one to four years and subjects were not distributed equally across the sexes. These factors are known to affect alarm calls [67,68]. Thus, future replication of our experimental paradigm with more groups while controlling these factors would be important.

2.7 Conclusion

We provide experimental evidence that conspecific alarm calls facilitate social learning about novel predators in wild primates and confirm Cheney and Seyfarth's observations that vervet monkeys socially acquire some aspects of anti-predatory behaviour [28]. Our results also extend recent findings of rapid comprehension learning in congeneric green monkeys to the social learning realm [29]. Our study also highlights the influence of other factors on social learning, such as age or core knowledge. Overall, competence in alarm call comprehension and production appears to be acquired by refining pre-existing cognitive biases through social learning. Future studies should examine multiple factors to get a nuanced picture of mechanisms and cognitive mapping involved in the vocal communication of non-human primates.

2.8 Data, code and material

Data and the code for the analyses and supplementary videos are available here, or copy the following link into the browser:

https://osf.io/ycnhj/?view_only=ade1f400289a46d18b6ef540bc7cbf8a

2.9 Competing interests

The authors declare no competing interests.

2.10 Author contributions

AD and KZ conceived the study. AD and BB collected the data. BB coded the videos. AD conducted analyses and wrote the first draft, which KZ reviewed. All the authors reviewed the final version of the manuscript.

2.11 Acknowledgements

We would like to thank Jean Robar for letting the cylindrical hide, José van der Geest for volunteering as a second coder, Christof Neumann for suggestions on the experimental design and Radu Slobodeanu for the help with the statistical analyses. Finally, we would like to thank Dr Erica van de Waal for her continuous support and all the volunteers, researchers, and managers at the Inkawu vervet project for their help during data collection.

2.12 References

1. Hauser MD, Chomsky N, Fitch WTS. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science (80-.).* **298**, 1569–79. (doi:10.1126/science.298.5598.1569)
2. Hammerschmidt K, Fischer J. 2013 Constraints in Primate Vocal Production. In *Evolution of Communicative Flexibility*, pp. 92–119. Cambridge, MA, USA: The MIT Press. (doi:10.7551/mitpress/9780262151214.003.0005)
3. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey (Cercopithecus aethiops) alarm calls: semantic communication in a free ranging primate. *Anim. Behav.* **28**, 1070–1094.
4. Zuberbühler K. 2000 Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* **76**, 195–207. (doi:10.1016/S0010-0277(00)00079-2)
5. Arnold K, Zuberbühler K. 2008 Meaningful call combinations in a non-human primate. *Curr. Biol.* **18**, 202–203. (doi:10.1016/j.cub.2008.01.040)
6. Berthet M, Mesbahi G, Pajot A, Casar C, Neumann C, Zuberbühler K. 2018 Titi monkeys combine alarm calls to create probabilistic meaning. *Sci. Adv.* , 1–10.
7. Cheney DL, Seyfarth RM. 2018 Flexible usage and social function in primate vocalizations. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 1974–1979. (doi:10.1073/pnas.1717572115)
8. Crockford C, Wittig RM, Mundry R, Zuberbühler K. 2012 Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* **22**, 142–146. (doi:10.1016/j.cub.2011.11.053)
9. Schel AM, Townsend SW, Machanda Z, Zuberbühler K, Slocombe KE. 2013 Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality. *PLoS One* **8**, e76674. (doi:10.1371/journal.pone.0076674)
10. Seyfarth RM, Cheney DL. 2016 Precursors to language: Social cognition and pragmatic inference in primates. *Psychon. Bull. Rev.* **24**, 1–6. (doi:10.3758/s13423-016-1059-9)
11. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)
12. Lima SL, Dill LM. 1990 Behavioral decisions made under the risk of predation: a review and prospectus. *Can. J. Zool.* **68**, 619–640. (doi:10.1139/z90-092)
13. Hoppitt W, Laland KN. 2008 Chapter 3 Social Processes Influencing Learning in Animals: A Review of the Evidence. *Adv. Study Behav.* **38**, 105–165. (doi:10.1016/S0065-3454(08)00003-X)
14. Jenny D, Zuberbühler K. 2005 Hunting behaviour in West African forest leopards. *Afr. J. Ecol.* **43**, 197–200. (doi:10.1111/j.1365-2028.2005.00565.x)

15. Griffin AS. 2004 Social learning about predators: A review and prospectus. *Learn. Behav.* **32**, 131–140. (doi:10.3758/BF03193188)

16. Hollén LI, Radford AN. 2009 The development of alarm call behaviour in mammals and birds. *Anim. Behav.* **78**, 791–800. (doi:10.1016/j.anbehav.2009.07.021)

17. Laland KN. 2004 Social learning strategies. *Learn. Behav.* **32**, 4–14.

18. Curio E, Ernst U, Vieth W. 1978 Cultural transmission of enemy recognition: One function of mobbing. *Science (80-.).* **202**, 899–901. (doi:10.1126/science.202.4370.899)

19. Gunhold T, Whiten A, Bugnyar T. 2014 Video demonstrations seed alternative problem-solving techniques in wild common marmosets. *Biol. Lett.* **10**. (doi:10.1098/rsbl.2014.0439)

20. van de Waal E, Borgeaud C, Whiten A. 2013 Potent Social Learning and Conformity Shape a Wild Primate's Foraging Decisions. *Science (80-.).* **340**, 483–485. (doi:10.1126/science.1232769)

21. Mineka S, Cook M. 1993 Mechanisms involved in the observational conditioning of fear. *J. Exp. Psychol. Gen.*

22. Herzog M, Hopf S. 1984 Behavioral responses to species-specific warning calls in infant squirrel monkeys reared in social isolation. *Am. J. Primatol.* **7**, 99–106. (doi:10.1002/ajp.1350070204)

23. Dezecache G, Crockford C, Zuberbühler K. 2019 The development of communication in alarm contexts in wild chimpanzees. *Behav. Ecol. Sociobiol.* **73**. (doi:10.1007/s00265-019-2716-6)

24. Price T, Fischer J. 2014 Meaning attribution in the West African green monkey: Influence of call type and context. *Anim. Cogn.* **17**, 277–286. (doi:10.1007/s10071-013-0660-9)

25. Struhsaker TT. 1967 Ecology of vervet monkeys (Cercopithecus aethiops) in The Masai-Amboseli Game Reserve, Kenya. *Ecology* **48**, 891–904. (doi:10.2307/1934531)

26. Seyfarth RM, Cheney DL, Marler P. 1980 Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801–803. (doi:10.1126/science.7433999)

27. Price T, Wadewitz P, Cheney DL, Seyfarth RM, Hammerschmidt K, Fischer J. 2015 Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* **5**, 1–11. (doi:10.1038/srep13220)

28. Seyfarth RM, Cheney DL. 1986 Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640–1658. (doi:10.1016/S0003-3472(86)80252-4)

29. Wegdell FL, Hammerschmidt K, Fischer J. 2019 Conserved alarm calls but rapid auditory learning in monkey responses to novel flying objects. *Nat. Ecol. Evol.* (doi:10.1038/s41559-019-0903-5)

30. Seyfarth RM, Cheney DL. 1980 The Ontogeny of Vervet Monkey Alarm Calling Behavior: A Preliminary Report. *Z. Tierpsychol.* **54**, 37–56. (doi:10.1111/j.1439-0310.1980.tb01062.x)

31. Botting J, Whiten A, Grampp M, van de Waal E. 2018 Field experiments with wild primates reveal no consistent dominance-based bias in social learning. *Anim. Behav.* **136**, 1–12. (doi:10.1016/j.anbehav.2017.11.025)

32. Canteloup C, Hoppitt W, van de Waal E. 2020 Wild primates copy higher-ranked individuals in a social transmission experiment. *Nat. Commun.* **11**, 459. (doi:10.1038/s41467-019-14209-8)

33. In press. Audacity Team. The name Audacity® is a registered trademark of Dominic Mazzoni.

34. Cheney DL, Seyfarth RM. 1982 How vervet monkeys perceive their grunts: Field playback experiments. *Anim. Behav.* **30**, 739–751. (doi:10.1016/S0003-3472(82)80146-2)

35. Friard O, Gamba M. 2016 BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol.*

36. Center for Conservation Bioacoustics. 2014 Raven Pro: Interactive Sound Analysis Software.

37. Venables WN, Ripley BD. 1999 *Modern Applied Statistics with S-Plus*.

38. Team RDC. 2020 R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna, Austria. ISBN 3-900051-07-0*.

39. Bates D, Maechler M, Bolker B, Walker S. 2015 Package lme4. *J. Stat. Softw.* **67**, 1–91. (doi:http://lme4.r-forge.r-project.org)

40. Carey G. 2013 The general linear model (GLM): A gentle introduction. *Quant. Methods Neurosci.* , 129–142.

41. Bolker BM *et al.* 2013 Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods Ecol. Evol.* **4**, 501–512. (doi:10.1111/2041-210X.12044)

42. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS. 2009 Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135. (doi:10.1016/j.tree.2008.10.008)

43. Fox, John ; Weisberg S. 2018 *An R Companion to Applied Regression*. Sage Publications.

44. Barton K. 2020 MuMIn: Multi-model inference.

45. Tomasello M. 2010 *Origins of Human Communication*. MIT press.

46. Janik VM, Slater PB. 2000 The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11. (doi:10.1006/anbe.2000.1410)

47. Seyfarth RM, Cheney DL. 2016 The origin of meaning in animal signals. *Anim. Behav.* , 1–8. (doi:10.1016/j.anbehav.2016.05.020)

48. Markson L, Bloom P. 1997 Evidence against a dedicated system for word learning in children. *Nature* **385**, 813–815. (doi:10.1038/385813a0)

49. Kaminski J, Call J, Fischer J. 2004 Word learning in a domestic dog: Evidence for 'fast mapping'. *Science (80-.).* **304**, 1682–1683. (doi:10.1126/science.1097859)

50. Kano F, Hirata S. 2015 Great apes make anticipatory looks based on long-term memory of single events. *Curr. Biol.* **25**, 2513–2517. (doi:10.1016/j.cub.2015.08.004)

51. Rescorla R, Wagner A, Black A, Prokasy W. 1972 Classical conditioning II: current research and theory.

52. Berger J *et al.* 2001 Recolonizing carnivores and naïve prey: Conservation lessons from pleistocene extinctions. *Science* **291**, 1036–9. (doi:10.1126/science.1056466)

53. Griffin AS, Evans CS, Blumstein DT. 2001 Learning specificity in acquired predator recognition. *Anim. Behav.* **62**, 577–589. (doi:10.1006/anbe.2001.1781)

54. Potvin DA, Ratnayake CP, Radford AN, Magrath RD, Potvin DA, Ratnayake CP, Radford AN, Magrath RD. 2018 Birds Learn Socially to Recognize Heterospecific Alarm Calls by Acoustic Association. *Curr. Biol.* , 1–6. (doi:10.1016/j.cub.2018.06.013)

55. Manser MB, Seyfarth RM, Cheney DL. 2002 Suricate alarm calls signal predator class and urgency. *Trends Cogn. Sci.* **6**, 55–57. (doi:10.1016/S1364-6613(00)01840-4)

56. Lind J, Ghirlanda S, Enquist M. 2018 Social learning through associative processes: A computational theory. *bioRxiv* (doi:10.1101/446906)

57. Whiten A, van de Waal E. 2018 The pervasive role of social learning in primate lifetime development. *Behav. Ecol. Sociobiol.* **72**. (doi:10.1007/s00265-018-2489-3)

58. Mesoudi A, Chang L, Dall SRXX, Thornton A. 2016 The Evolution of Individual and Cultural Variation in Social Learning. *Trends Ecol. Evol.* **31**, 215–225. (doi:10.1016/j.tree.2015.12.012)

59. Shapiro AD, Slater PJB, Janik VM. 2004 Call usage learning in gray seals (Halichoerus grypus). *J. Comp. Psychol.* **118**, 447–454. (doi:10.1037/0735-7036.118.4.447)

60. Collier TC, Blumstein DT, Girod L, Taylor CE. 2010 Is alarm calling risky? Marmots avoid calling from risky places. *Ethology* **116**, 1171–1178. (doi:10.1111/j.1439-0310.2010.01830.x)

61. Sigala N, Logothetis NK. 2002 Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* **415**, 318–320. (doi:10.1038/415318a)

62. Sigala N, Gabbiani F, Logothetis NK. 2002 Visual Categorization and Object Representation in Monkeys and Humans. *J. Cogn. Neurosci.* , 187–198. (doi:10.1162/089892902317236830)

63. Greggor AL, McIvor GE, Clayton NS, Thornton A. 2018 Wild jackdaws are wary of objects that violate expectations of animacy. *R. Soc. Open Sci.* **5**. (doi:10.1098/rsos.181070)

64. Tsutsumi S, Ushitani T, Tomonaga M, Fujita K. 2012 Infant monkeys' concept of animacy: The role of eyes and fluffiness. *Primates* **53**, 113–119. (doi:10.1007/s10329-011-0289-8)

65. Spelke E. 1994 Initial knowledge: six suggestions. *Cognition* **50**, 431–445. (doi:10.1016/0010-0277(94)90039-6)

66. Shettleworth SJ. 2012 Modularity, comparative cognition and human uniqueness. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2794–2802. (doi:10.1098/rstb.2012.0211)

67. Lea AJ, Blumstein DT. 2011 Ontogenetic and sex differences influence alarm call responses in mammals: A meta-analysis. *Ethology* **117**, 839–851. (doi:10.1111/j.1439-0310.2011.01947.x)

68. Zuberbühler K. 2008 Audience effects. *Curr. Biol.* **18**, 189–190. (doi:10.1016/j.cub.2007.12.041)

2.13 Chapter 2: Appendix

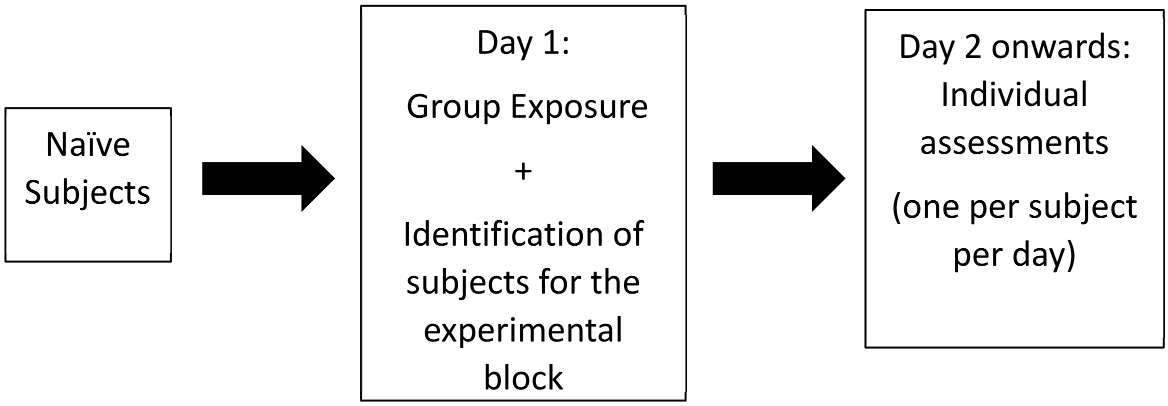

Figure A2.1: Design and timeline of the experimental blocks of different treatments

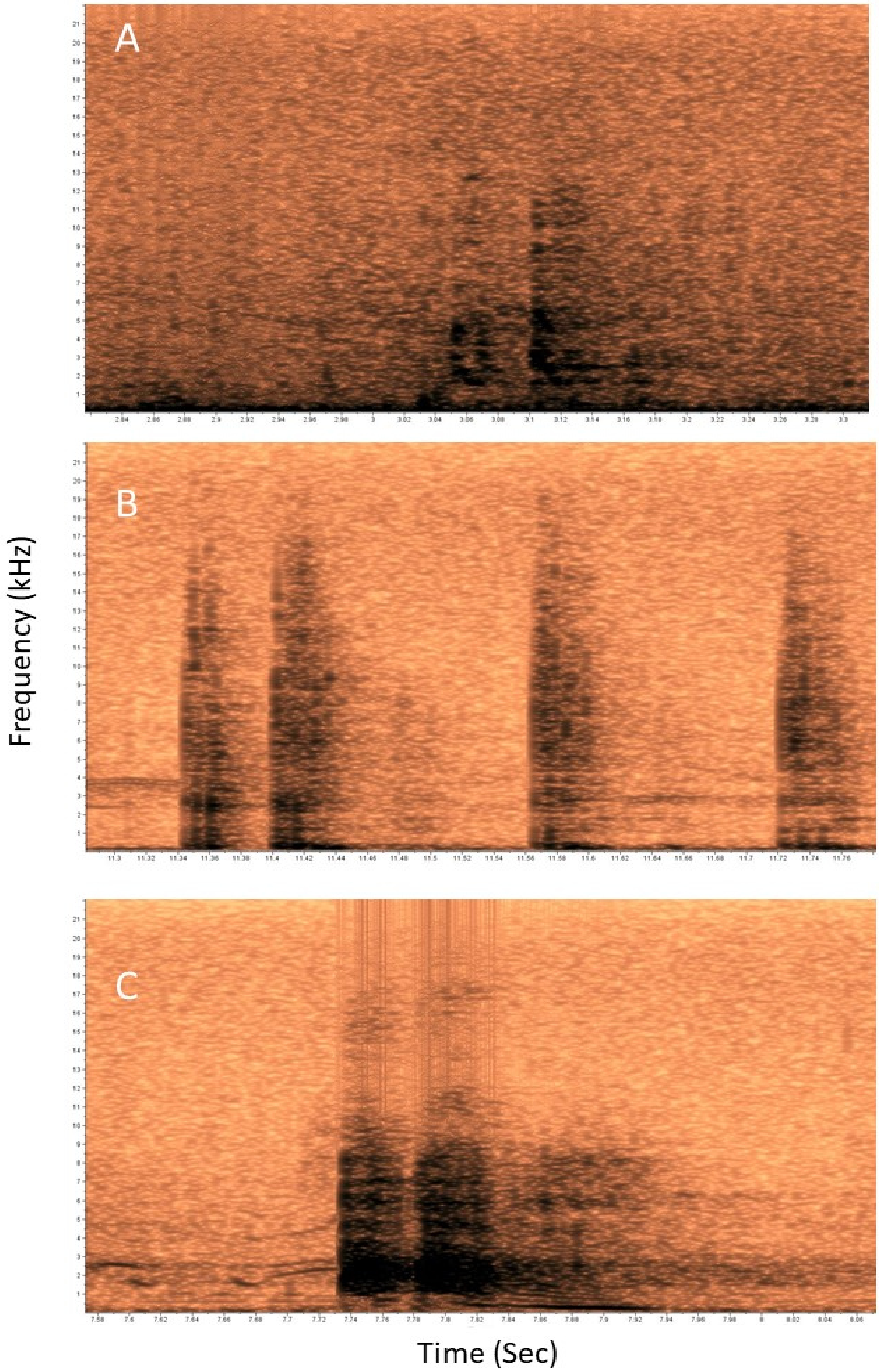

Figure A2.2: Spectrograms (Hamming window at 1024 DFT and 93.8% overlap) of different types of alarm calls elicited by the subjects (A) Eagle alarms (B) Snake alarms (C) Leopard alarms as described in the previous studies

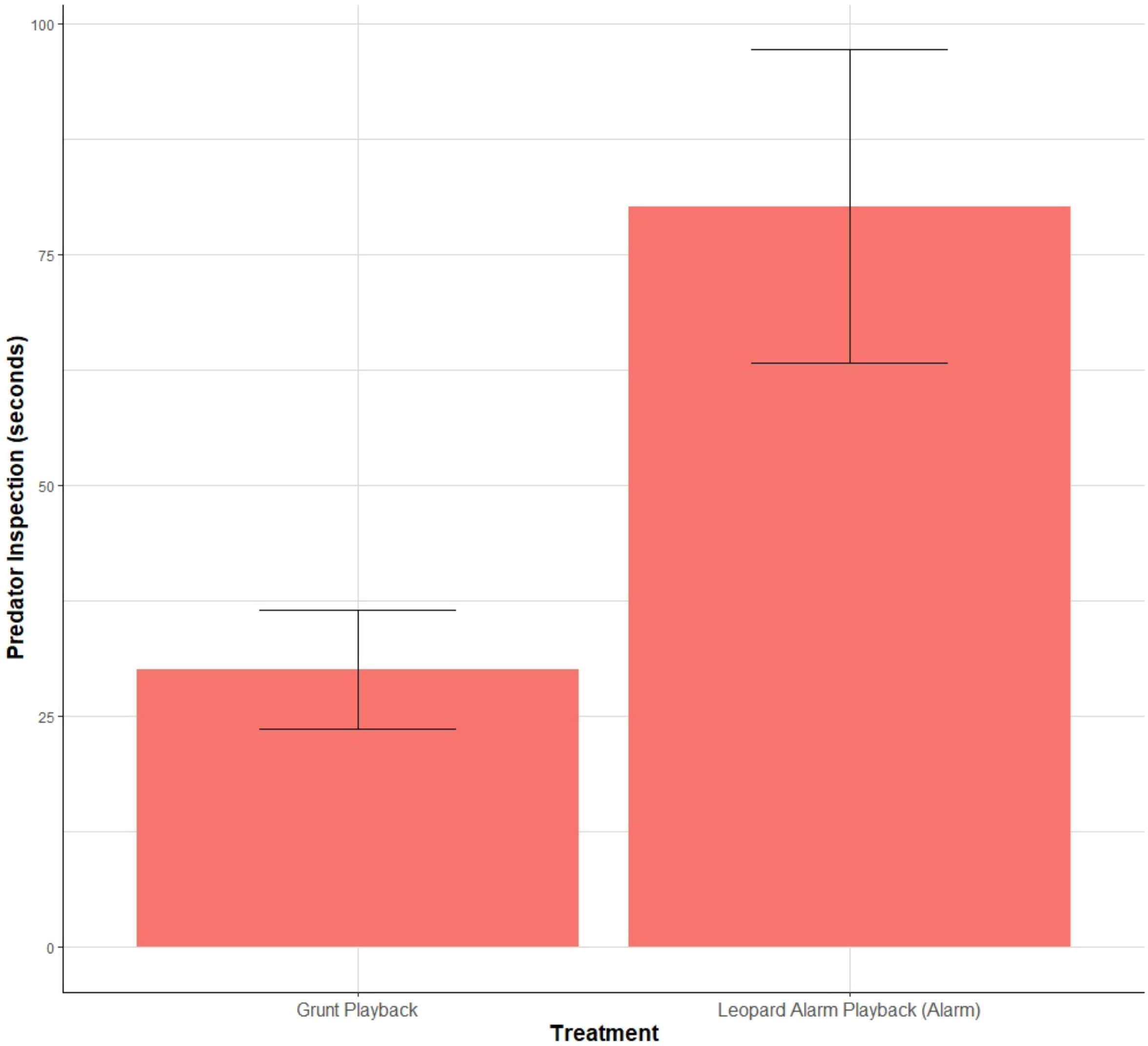

Figure A2.3: Mean (± SE) predator inspection of the subjects for the Grunt playback and leopard alarm playback (Alarm) treatment

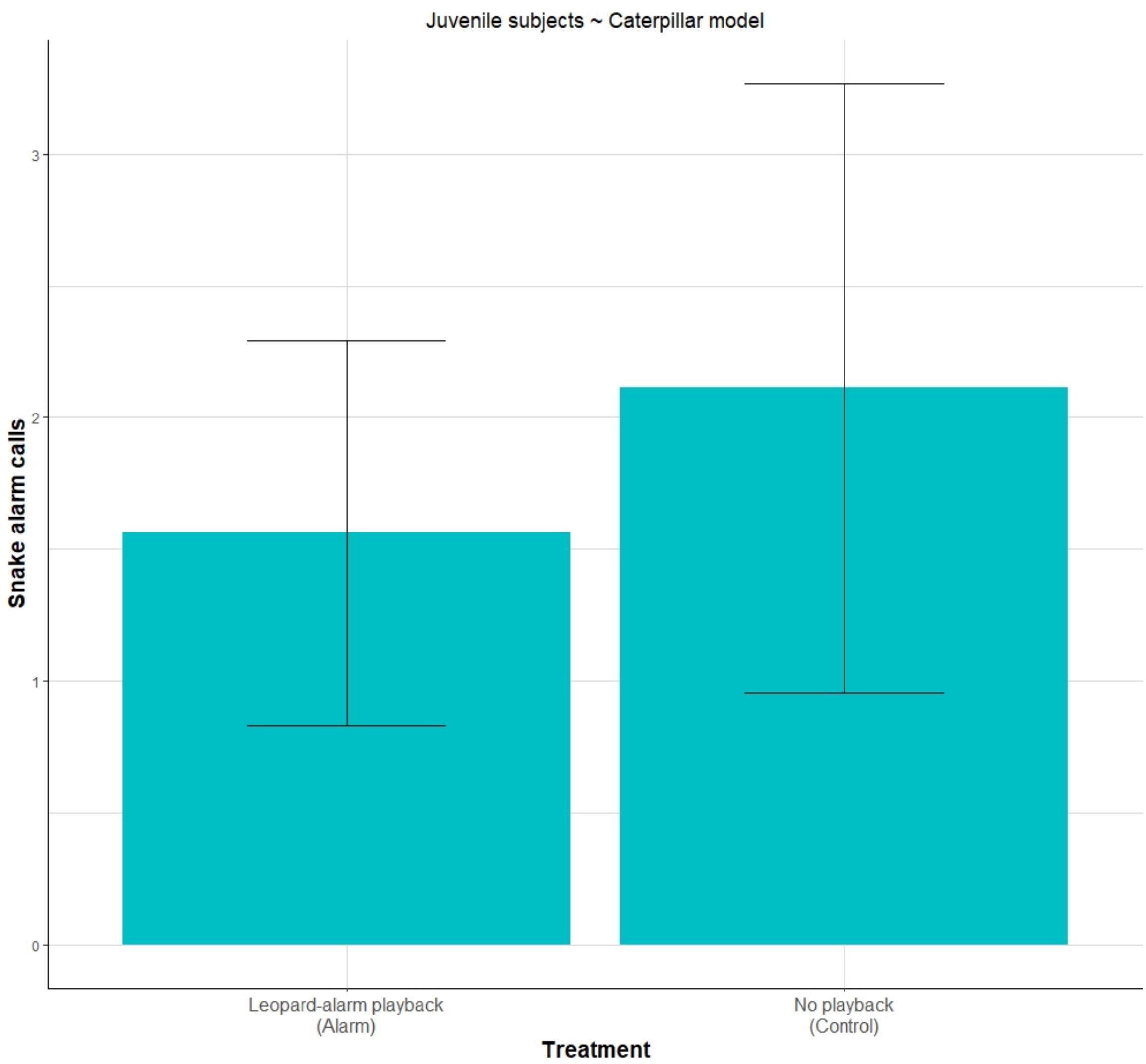

Figure A2.4: Model predictions (± SE) of the effects of treatments on the leopard snake for the juvenile subjects exposed to caterpillar model

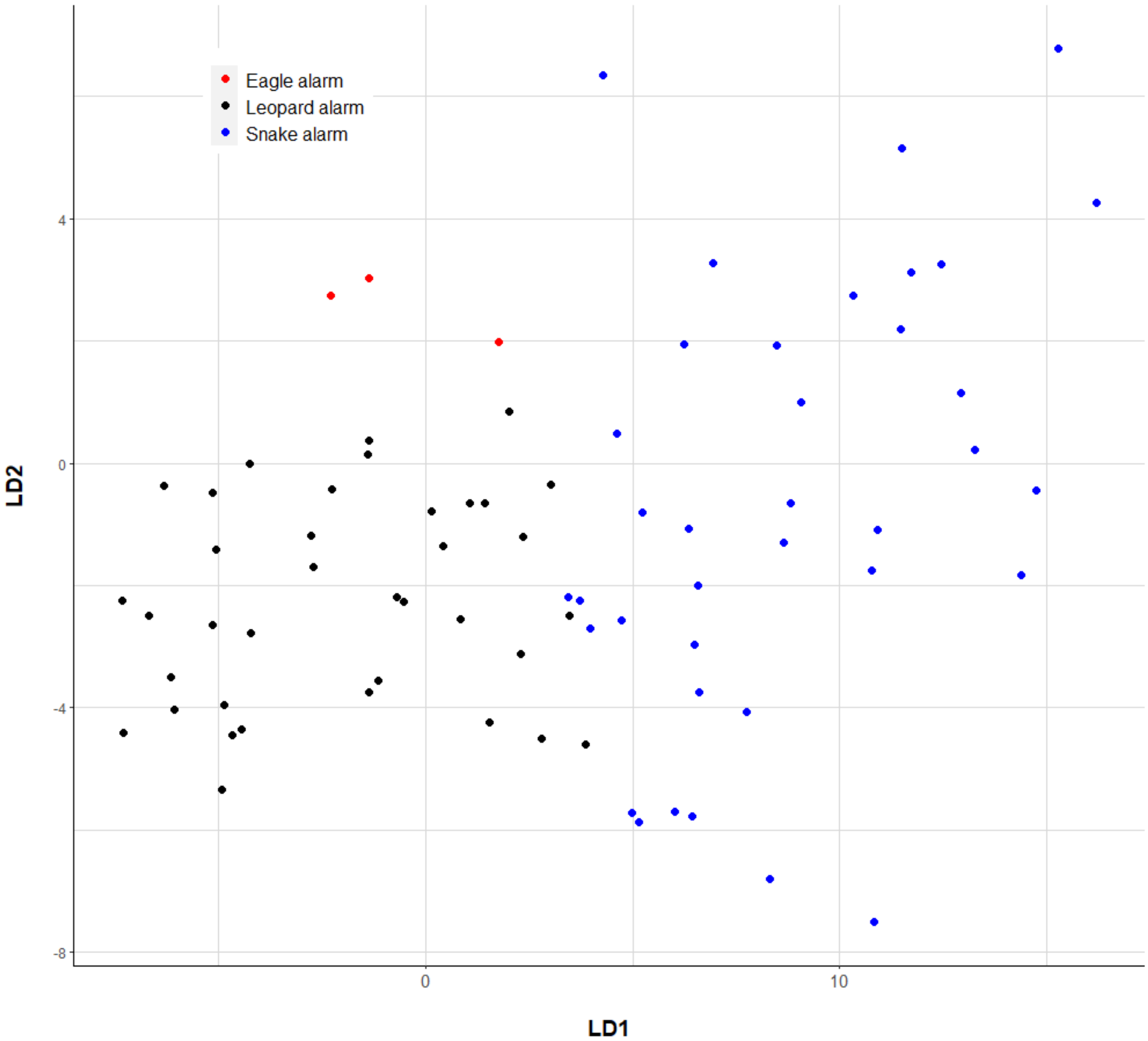

Figure A2.5: Predicted classification of the alarm calls elicited by the subjects during individual assessments into the three different alarm call types. The axes represent coefficients of two linear discriminants (LD) as per the LDA model.

Table A2.1: List of acoustic parameters extracted from the alarm calls. Shaded parameters were used for the LDA classification.

Acoustic parameter (unit)	Group Means from LDA Model (Scaled Data)		
	Eagle	Leopard	Snake
Number of elements	-0.0350025	-0.3522358	1.24903709
Low Frequency (Hz)	0.0736336	-0.3854407	1.3558483
High Frequency (Hz)	NA	NA	NA
Delta Time (s)	-0.0970311	1.89210804	- 0.3294874
Peak Frequency (Hz)	NA	NA	NA
1st Quartile Frequency (Hz)	-0.908022	-0.755146	1.005444
3rd Quartile Frequency (Hz)	-1.298510	-1.112019	1.152288
Delta Frequency (Hz)	-0.4686237	0.1796256	0.5083690
Average Entropy (bits)	-1.5461260	-1.2993005	0.9196337
Center Frequency (Hz)	-1.130832	-0.958378	1.084521
Max Entropy (bits)	-1.458680	-1.389708	0.633118
Average Power (dB)	-0.187820	1.892933	-0.514142
Average Amplitude (U)	-0.07130897	0.34673157	-0.07008640
Max Amplitude (U)	NA	NA	NA
Min Amplitude (U)	0.3045821	-1.5637610	0.3033251
Min Entropy (bits)	-0.9943584	-1.1550906	1.0071179
Energy (dB)	NA	NA	NA
Peak Amplitude (U)	NA	NA	NA
Peak Power (dB)	NA	NA	NA
Max Frequency (Hz)	-0.7533995	-0.6241039	1.0150978
Max Power (dB)	NA	NA	NA
Peak Frequency Contour Max Frequency (Hz)	-1.0853231	-0.8660863	-0.8660863
Peak Frequency Contour Average Slope (Hz/ms)	-0.038994987	-0.008381332	-0.008381332

Table A2.2: Coefficients of two linear discriminants (LD1, LD2) for the acoustic parameters extracted from the naturally elicited alarm calls.

Acoustic parameter (unit)	LD1	LD2
Number of elements	2.020090825	0.5037565
Low Frequency (Hz)	-0.134230415	1.0417247
Delta Time (s)	0.852853764	0.1073697
1st Quartile Frequency (Hz)	0.584751462	-1.9499961
3rd Quartile Frequency (Hz)	10.412549719	2.7626240
Delta Frequency (Hz)	-0.272457243	-0.1311624
Average Entropy (bits)	-2.655621226	-2.3014794
Center Frequency (Hz)	-3.181253834	-1.9464141
Max Entropy (bits)	0.489397479	-0.7991997
Average Power (dB)	-0.375999373	-3.4318099
Average Amplitude (U)	-0.158447347	-0.1527833
Min Amplitude (U)	-0.241407422	-1.0681072
Min Entropy (bits)	0.836528452	0.1623082
Max Frequency (Hz)	0.512005842	0.7755586
Peak Frequency Contour Max Frequency (Hz)	-0.392835201	0.7873859
Peak Frequency Contour Average Slope (Hz/ms)	-0.003812237	-0.5417000

Table A2.3: Alarm call types elicited by the subjects predicted by the LDA model.

Subject	Age	Treatment	Model type	Natural Eagle	Natural Leopard	Natural Snake
Add	Juvenile	Alarm	Caterpillar	0	1	0
Ati	Adult	Control	Caterpillar	1	1	4
Gaa	Juvenile	Alarm	Caterpillar	0	0	2
Ghid	Juvenile	Control	Caterpillar	0	1	1
Gubh	Juvenile	Control	Caterpillar	0	6	2
Nak	Juvenile	Alarm	Horse	1	2	5
Ndaw	Juvenile	Control	Caterpillar	0	1	3
Ndi	Juvenile	Control	Caterpillar	0	1	5
Ndik	Juvenile	Alarm	Horse	0	2	0
Nge	Juvenile	Alarm	Horse	0	2	2
Obe	Juvenile	Alarm	Caterpillar	0	7	1
Pom	Juvenile	Alarm	Caterpillar	0	3	5
Puol	Juvenile	Alarm	Caterpillar	0	2	2
Rat	Juvenile	Alarm	Caterpillar	1	6	3
Sirk	Juvenile	Control	Horse	0	2	2

Chapter 3: Socially learned arbitrary call use in a wild primate

Adwait Deshpande 1,2, and Klaus Zuberbuhler 1,2,3

1 Department of Comparative Cognition, Institute of Biology, University of Neuchatel, Neuchatel, Switzerland, 2 Inkawu Vervet Project, Mawana Game Reserve, KwaZulu-Natal, South Africa, 3 School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom

3.1 Abstract

Human speech is an unusually flexible form of primate communication. A limited repertoire of phonemes is assembled into morphemes whose combinations are linked to referents by social learning in culturally specific ways. Non-human primate vocal utterances, in contrast, are drastically less flexible at every level, i.e., acoustic diversity, combinatoriality and referential function. Here, we were interested in whether primate vocal behaviour, despite its inflexible nature, is still amenable to social learning, a well-documented, powerful mechanism in many other aspects of primate behaviour. We provided wild vervet monkeys with social learning opportunities to operate a food dispensing apparatus by uttering specific vocalisations, 'move-grunts', that were contextually inappropriate. In two separate experiments, subjects witnessed how another group member produced move-grunts whilst operating a food dispenser. In experiment 1, subjects only heard playbacks of the calls, followed by food release; in experiment 2, they witnessed a complete demonstration video of conspecific including the calls, followed by food release. None of the subjects learned the task in experiment 1, but 1 of 39 subjects, a juvenile female, succeeded in experiment 2 by producing move-grunts at the food dispenser to obtain food rewards. We discuss our results and behavioural observations in relation to the current theory and its eventual implications for reconstructing the evolutionary pathway to human speech.

3.2 Keywords

Vocal usage learning; Functional Flexibility; Social Learning; Evolution of language

3.3 Introduction

Human language is the capacity to represent and communicate events as hierarchically structured, recursively nested mental structures via a linear signal stream [1]. The natural vehicle for this process is speech, which consists of language-specific phoneme repertoires that range from less than a dozen (Piraha) to over a hundred (!Khong). Each language can be phonetically transcribed in terms of its speech sounds [2] and in how sound combinations contribute to linguistic structures, making human speech arguably the most flexible and complex natural vocal communication system known [3,4]. The underlying ability to generate speech sounds emerges early during human ontogeny, around the end of the first month, with infants producing proto-phones decoupled from any affective state and basic biological function, the 'babbling' behaviour [3,5]. Proto-phones are produced in addition to a more standard repertoire of functionally fixed vocalisations such as laughing, crying or protesting [6], used in specific social interactions similar to the calls of non-human primates. Proto-phones, however, do not have any obvious biological function but appear to be the foundation for subsequent speech development. During this process, infants socially learn from other group members to use these sounds and their combinations to interact socially and as referents for external events [7,8]. Thus, to reconstruct the evolutionary history of speech and language, it is vital to shed light on how human evolved the capacity to produce functionally flexible sounds amenable to social learning.

One promising way to address evolutionary questions of human behaviour is to make direct comparisons with closely related non-human primates [9]. According to most theories, primate communication is characterised by small call repertoires, with signallers prevented from majorly modifying the acoustic structure of existing calls and from producing novel sounds [10,11]. Moreover, apart from reports on marmosets (Reviewed in [12]), non-human primate infants do not show human-like babbling behaviour but gradually acquire an adult-like vocal repertoire with no compelling evidence of social learning [13]. Flexibility is somewhat higher in primate call usage [14], with examples of signallers using the same call in different contexts. For example, bonobos (*Pan paniscus*) produce 'peep' calls across several contexts, encompassing various emotional valences [15]. The same has been found for the 'grunt' calls in juvenile chimpanzees (*Pan troglodytes*) [16] and 'coo' calls in different macaque species [17–19].

While these findings confirm that primates can use the same call in different contexts, it is still possible that call production is the direct result of the same underlying psychological state experienced by callers. For example, many primates produce acoustically distinct alarms to a wide range of terrestrial disturbances, including non-predatory events [20], suggesting the same basic inner state evokes the calls [11,21]. However, other primate calls, such as bonobo 'peeps' or macaque 'coo' calls, are given to such a wide range of events that they appear to be affectively decoupled, i.e., functionally flexible [15,17,18]. In an early study, Sutton et al. were indeed able to condition rhesus macaques (*Macaca mulatta*) to produce 'coo' calls following arbitrary visual stimuli [22], although the test window was very long (5 min after stimulus presentation), suggesting that coo calls might have been produced for other reasons (see [22,23] for similar concerns). In other studies, conditioning has been possible with food calls. However, here, the issue is that callers may produce the calls in anticipation of the food rewards when completing a trial successfully, not as a response to the conditioned stimuli [17,24]. A recent study with rhesus macaques attempted to overcome these limitations by showing that subjects could switch between two call types based on different visual stimuli [25].

While such experiments demonstrate that non-human primates have some degree of volitional control over call use, they all required extensive training in captive settings with operant conditioning paradigms, raising questions about ecological validity. In natural conditions, the conditions for operant conditioning are rarely met, as both social and physical environments change continuously while consistent reward patterns are unlikely to be found. Social learning, i.e., learning by observing others, is a much more probable learning mechanism, which is well-documented and highly adaptive in wild animals, especially in foraging or antipredator contexts [26,27]. However, its importance in call use and communication is not well documented [28].

In this study, we addressed this issue by providing social learning opportunities to wild South African vervet monkeys (*Chlorocebus pygerythrus pygerythrus*), a species with an established capacity for social learning [29–31] and a rich vocal repertoire of calls with potential for functional flexibility [32,33]. We conducted two experiments on two different groups of monkeys that introduced subjects to the fact that a food dispenser could be activated by producing an out-of-context vocalisation, a 'move-grunt'. In the experiments, subjects witnessed another group member obtaining food from the dispenser by producing move-grunts (or 'moving-into-an-open-area' grunt), a call type typically given by individuals that move into open areas, observe other group members do so [34] or initiate group movements (AD, unpublished data). The call emerges relatively early during ontogeny, with 8-12 weeks old juveniles already producing the call in the correct context [35]. Due to these observations and the fact that this call type does not provoke any visible positive or negative affect in callers or receivers [36], we speculated that vervet monkey 'move-grunts' could be functionally flexible call, similar to bonobo 'peep' calls or macaque 'coo' calls. And possibly more amenable to social learning than other vocalisations.

In the first experiment (conducted on the first group of monkeys), subjects heard a playback of a move-grunt by a familiar group member directly followed by a food reward delivered from a food dispenser (Figure 3.1). We predicted that if monkeys were able to learn the association between call type and food reward, they would start producing move-grunt themselves to get the rewards. In the second experiment (conducted on the second group of monkeys), subjects watched a demonstration video of a conspecific obtaining a food reward following playback of a move grunt. This second experiment is based on findings in both wild and captive settings, showing that non-human primates can learn to extract food from an apparatus [37,38] and that they can learn from watching video clips. Again, we predicted that subjects could learn to operate the food dispenser from observing others producing move-grunts.

Figure 3.1: A photograph of the manual food dispenser set up for the experiments

3.4 Methods

3.4.1 Study site and groups

The study was conducted at the Inkawu Vervet Project (IVP) on the Mawana Game Reserve, Kwazulu Natal, South Africa (S 28° 00.327; E 031° 12.348). At the IVP, several neighbouring groups of wild South African vervet monkeys (*Chlorocebus pygerythrus pygerythrus*) are well habituated to human presence and allow close observations by multiple observers. All observers go through standard training followed by inter-observer reliability tests for individual recognition and ethogram recordings before participating in data collection. Individuals were fed regularly on corn kernels and apple pieces for various experiments. Food provisioning events were announced using specific calls by observers, which are understood by monkeys [27,39].

3.4.2 Experiment 1: Social learning of auditory demonstrations

In the first experiment, we provided monkeys with a playback call of another group member's move-grunts immediately followed by a food reward delivered manually from a dispenser. This experiment was conducted on group Noha (NH). At the time of the experiment, the group's size was 34 individuals, including juveniles. We used a naturally recorded move-grunt of a middle-ranking adult female of the group as a playback stimulus. The call was recorded when the female was moving out from the sleeping site in the morning (see Appendix). To record the call, we used an audio recorder (Marantz PMD 661) and a directional microphone (Sennheiser MKH 416 P48) with a sampling frequency of 44.1 kHz. We padded the call with 2 s of silence before and after using the Raven Pro 1.5 for the playback [40].

For the delivery of rewards, we used a custom-made, manually operated food dispenser. It was made of a handheld (1.5 m long and 7.6 cm diameter) polyvinyl chloride (PVC) pipe and a translucent plastic container. The experimenter sat on a field chair holding pipe at one end while the other end of the pipe was positioned inside the plastic container 1-2 m away from the experimenter (Figure 3.1). The experimenter was holding hidden food rewards (red apple pieces approx. 2.5 cm cubes). To reward a subject, the experimenter dropped an apple piece from his/her end of the pipe. The apple piece slid down the pipe and came out from another end in the plastic container to become accessible for the subject.

The final experimental setup consisted of the food dispenser handled by experimenter A1 (see Appendix). The hidden speaker (Nagra Kudelski, DSM-Monitor) was placed 5m away from the food dispenser, and the playback call was played from a speaker through a Motorola XPlay phone. We pre-adjusted the speaker amplitude similar to a natural call which was kept constant throughout the experiment. The speaker was operated by experimenter B1 sitting close to it. Experimenter C1 who recorded the subjects' behaviours, was positioned 3-8m from the food dispenser. We used two cameras (JVC quadproof EverioR camcorder) on tripods to record all the experimental sessions. One camera was positioned to cover the entire setup with a wide-angle view, whereas the other camera was focused on the food dispenser. We also recorded all the vocalisations elicited around the food dispenser using a solid-state audio recorder (Marantz PMD 661) and a directional microphone (Sennheiser MKH 416 P48) with a sampling frequency of 44.1 kHz.

Each experimental session started early in the morning near the sleeping site of the monkeys just after sunrise. Once the experimental setup was ready, we attracted the monkeys to the setup using conventionalised human 'food-provisioning calls' (see study site and groups) and provided them with a small number of corn kernels for the first three sessions. Once more than one monkey was present in the setup's proximity (5-10m), we started an experimental session. We conducted seven experimental sessions, which lasted from 10 to 105 minutes (Mean 56.8 minutes). A session ended when the group moved away from the setup.

Each experimental session consisted of training and contrast trials. For the training trials, experimenter B1 played a move-grunt from the hidden speaker in response to which the experimenter A1 immediately (< 1s) dispensed a reward usually picked up by the subject. If no individual approached the dispenser, the rewards were recovered by experimenter A1; such trials were considered trials without subjects. We gradually increased the time between two consecutive trials with each experimental session (mean: 54s, range: 1s-698s). Food rewards were also dispensed when a monkey in the proximity of the setup (5-8m) naturally elicited a move-grunt.

For the contrast trials, the experimenter dispensed non-edible items (small stones) from the dispenser without a playback. Contrast trials were designed to ensure that monkeys did not associate subtle, unintentional cues from the experimenter with the appearance of food rewards, i.e., to prevent Clever-Hans effects [41]. Contrast trials were conducted between two training trials randomly decided by the experimenter A1.

For each trial, the monkey closest to the food dispenser was considered the subject. If multiple monkeys were present, the experimenter considered the first individual looking at the dispenser as the subject.

For each trial, the experimenter C1 recorded the following behaviours by the subject: (1) Look at the speaker (yes/no: head orientation towards speaker after playback); (2) Look at the dispenser (yes/no: head orientation towards food dispenser before the reward was dispensed); (3) Gaze alternation (yes/no: head orientation towards speaker followed by rapid (<2s) reorientation to the dispenser (or vice-versa; see supplementary Video); (4) Startle response (yes/no: abrupt, sudden and brief locomotor response. For each trial, the second observer noted the identities of the audience members within a 5-8m radius around the setup.

We considered social learning to be accomplished if a subject elicited a move-grunt in the proximity of the food dispenser with a head orientation towards the dispenser and without any apparent group movement immediately before and after the call.

3.4.3 Experiment 2: Social learning of audio-visual demonstrations

In the second experiment, we provided monkeys with an opportunity to learn the novel usage of move-grunt from a demonstration video. This experiment was conducted on the 'Baie Dankie' group (BD) with 64 individuals, including juveniles. The demonstration video was edited from footage of two different events from Experiment 1. In the videos, a subject monkey from group NH can be seen stationed at the food dispenser and rewarded with an apple piece after the playback of move-grunt. We then removed the original audio track from both the footages and manually inserted the move-grunt (same as playback of Experiment 1) just before the food reward was dispensed in both the footages. Editing made both video events appear like the stationed monkey is eliciting the move-grunt and getting rewarded immediately after the call (at least to the human observer). In both the footages, the experimenter providing the food rewards was only partially visible. We then stitched one footage after another in a loop to make a 30 min long demonstration video showing 120 demonstrations (60 demonstrations of each event) (see Supplementary Video). The demonstration video was played on a 15-inch LCD laptop screen. The laptop was placed inside a protective opaque box with transparent plexiglass on one side for video watching (henceforth a video box) (Figure 3.2).

The final experimental setup contained two components; the video box place on the ground and the food dispenser (same as Experiment 1) held by the experimenter A2. The food dispenser was placed in front of the video box, approximately 3-5 m away. We also placed two different video cameras (JVC quadproof EverioR camcorder) on tripods. One camera recorded the whole setup with a wide-angle view. The second camera, along with the audio recorder (Marantz PMD 661) and a directional microphone (Sennheiser MKH 416 P48), was placed to record the behaviours and vocalisations of the monkeys around the food box (see Appendix). Each experimental session was 45 mins. The video was played for 30 mins, and for the last 15 min, we waited at the setup without video playing. Monkeys were free to interact with the food dispenser throughout the entire session. We did 2-5 playbacks of the move-grunt from a hidden speaker (Same as Experiment 1) in the last fifteen minutes of the seven experimental sessions. The experimenter dispensed the reward after each playback which was picked up by a subject monkey. The objective was to reinforce that the reward is independent of the videos and contingent solely on call elicitation. Two other experimenters B2 and C2, were present during each experimental session. Each experimental session was conducted in the early morning near the sleeping site of the monkeys. We chose a relatively open space near the sleeping site to set up the experiment. Before starting the videos, we attracted the monkeys to the video box by putting a few corn kernels in front of the box. The experimenter A2 played the video on the laptop and assumed the food dispenser position, which marked the start of the experimental session. During eight sessions, the monkeys left the area within 45 mins. We waited for 10-15 min for the monkeys to come back to the area; otherwise abandoned the session before 45 min. We conducted one experimental session every week over three months, with a total of 20 experimental sessions.

During the session, the experimenter A2 rewarded a subject upon the successfully eliciting move-grunt. I.e. if the subject elicited a move-grunt in the proximity of the food dispenser with a head orientation towards the dispenser and without any apparent group movement immediately before and after the call (see supplementary Video). The experimenter was highly trained in identifying call types in this species; however, as required to dispense the reward immediately after the call, there was a margin for human error (See results).

We also recorded the following behaviours during each experiment session. (1) Video-watching: Sitting/standing directly in front of the video box with the direct, clear head orientation towards the laptop screen when the video was playing. We recorded all the individuals looking at the video during the scan taken every minute by experimenter B2 [42]; (2) Look at dispenser: Direct gaze which inferred through a clear head orientation towards food dispenser while video-watching. The dispenser was placed opposite the video box, making it easier to infer clear head orientation as the subject had to turn around the head more than 90 degrees (Figure 3.3). (3) Approach the dispenser: The monkey was considered approaching the dispenser when it walked directly towards the dispenser. The last two behaviours were recorded ad-libitum by experimenter C2. Any monkeys that elicited one of the three behaviours were considered subjects for the analyses. We used a long-term dataset of the BD group's vocalisations to determine the natural frequency of move-grunt calls. These call recordings were collected over 36 days spread across six months, with a sampling effort of 8 hrs per day. We recorded all thevocalisations elicited by BD group members during the sampling, which the observer could hear. For both experiments, descriptive statistical analyses were conducted in R v 4.0.3 [43].

Figure 3.3: Photograph of a subject monkey watching a demonstration video during Experiment 2: Social learning of audio-visual demonstrations

Figure 3.2: Photograph showing the experimental setup for Experiment 2. Video-box and the dispenser held by the experimenter are visible. A subject monkey can be seen watching the demonstration video

3.5 Results

3.5.1 Experiment 1: Social learning of auditory demonstrations

We carried out a total of 440 trials (392 training, 48 contrast trials) distributed over seven experimental sessions (training: median= 56 , mean = 62.8 ± 17.92, range 11-133 trials per session; contrast: median= 6.5, mean = 8.0 ± 2.73, range = 2-22 trials per session) between sessions two and seven. In 344 trials a subject was present (range 3-11; median= 4, mean= 5.57 ± 1.28 subject per session); the remaining 96 trials were without subjects. A total of 13 subjects participated in one or several trials (range 1-79, median= 15 trials, mean= 26.46 ± 1.64 trials per subject) (Table 3.1).

No subject completed the task successfully, i.e., no one produced a move-grunt between the two trials in any session while stationed at the dispenser with a direct head orientation towards the dispenser.

Table 3.1: Summary statistics for Experiment 1: Social learning of auditory demonstrations

Total number of trials in seven sessions	Trials with subjects/ trials without subject	Total number of subjects	Number of times task was successfully completed
440 (392 training trials, mean= 62.8 ± 17.92 trials per session and 48 contrast trials, mean= 8.0 ± 2.73)	344/96	13 (mean= 5.57 ± 1.28 subject per session)	0

The participation by subjects was highly skewed. Three subjects participated in 229 trials (66.5% of trials). The first subject, Gran, was a high-ranking juvenile (4-year-old) female. The second subject, Wolf, was an adult male, whereas the third subject, Ula, was a juvenile (3-year-old) male. Due to this participation skew, we focused our subsequent behavioural analyses only on these three subjects. (See Appendix). We observed some trends in our raw data during the first session in terms of looking at the dispenser, looking at the speaker, no reaction and a startle response. From the second session, we never observed a startle response and looking at the speaker. The quick sequence of looking at the box and speaker or vice-a-versa was observed in session two, three and four. We only observed looking at the box or no reactions from the subjects in the last three sessions (Figure 3.4).

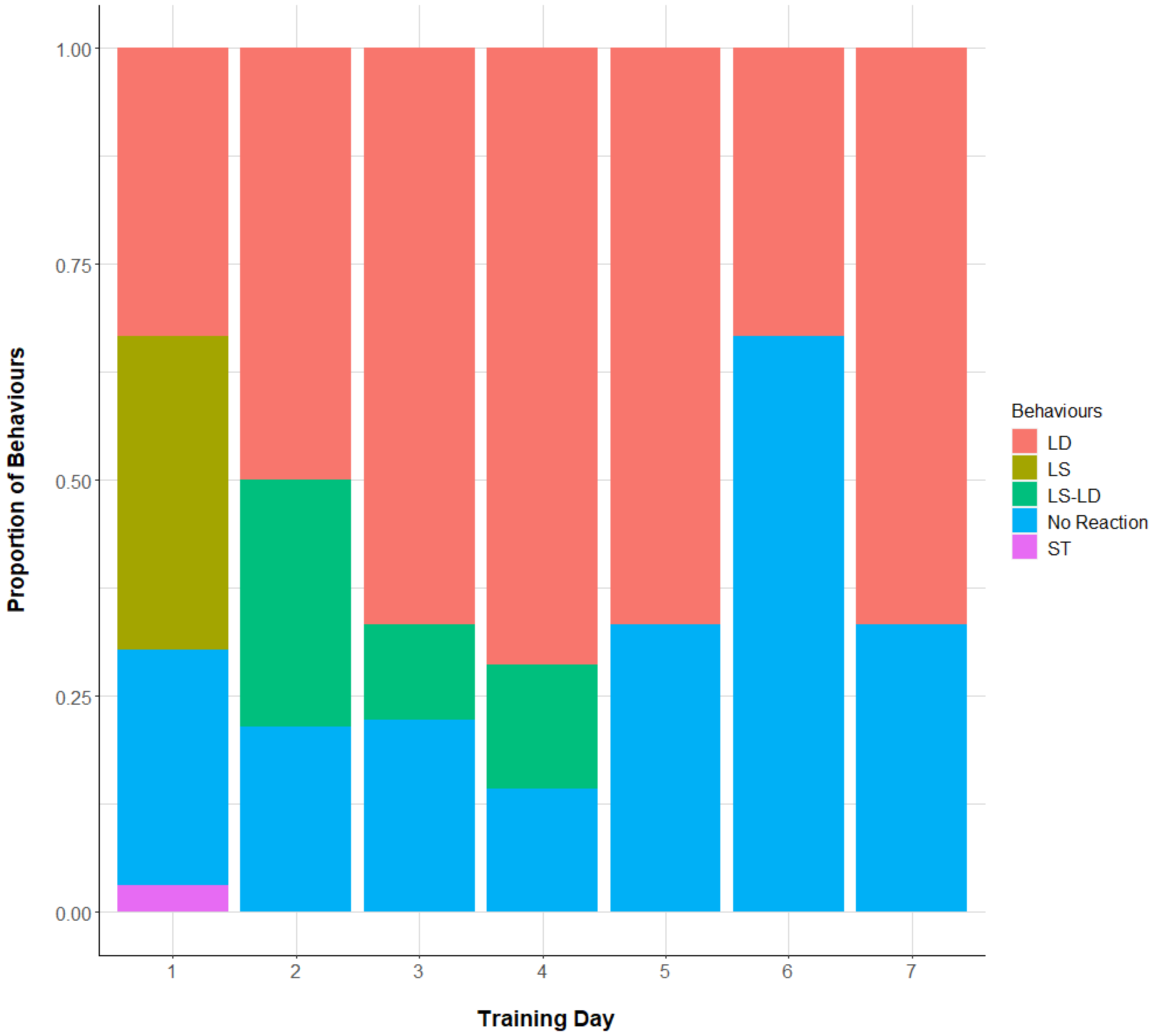

Figure 3.4: Proportions of cumulative behaviours of three subjects during participated trials across sessions of Experiment 1: Social learning of auditory demonstrations. LD (look at dispenser), LS (look at speaker), LS-LD (Gaze alternation), ST (Startle response)

3.5.2 Experiment 2: Social learning of audio-visual demonstrations

We carried out a total of 20 sessions, which lasted from 20-57 min (Mean = 39m 51s ±2m). Ten experimental sessions were 45 min long as per the experimental design. The number of subjects per session ranged from 3 – 27 (median= 9, mean = 9.70 ±1.29 subjects per session). We analysed three behaviours elicited by the subjects in every session. Video-watching ranged from 4 – 138 times per session (median= 9.50, mean = 18.25 ±6.53 times per session). Looks at the dispenser ranged from 1- 60 per session (median= 4.50, mean = 11.95 ±4.09 times per session). Approaches to the dispenser per session ranged from 1- 34 (median= 10, mean= 11.35 ±1.77 times per session). Total behaviours (sum of the three behaviours) per session ranged from 11-197 (median= 25.5, mean= 42.1 ±10.20 behaviours per session) (See supplementary behaviour). Thirty-nine individuals participated in the experiment (adults: females N=9; males N=9; juveniles: females N=11; males N=10) (Table 3.2).

Table 3.2: Summary statistics for Experiment 2: Social learning of audio-visual demonstrations

Total number of sessions	Mean duration of session ± Standard error	Total number of subjects	Total number of subjects who successfully completed the task	Total number of times task was completed successfully
20	Mean = 39m 51s ±2m	39 (mean = 9.70 ±1.29 subjects per session)	1	33 (All the successful completions between session 12 to 19)

One subject, Oorties (OT), a 4-year-old juvenile female, completed the task successfully by eliciting move-grunts in the previously specified way (see Appendix). A successful performance was observed from session 12 to 19, during which she produced 42 successful calls when stationed at the dispenser, for which she was rewarded every time by experimenter A2. Thirty-three calls were move-grunts produced in the correct way (Mean 4.12 correct calls per session between session 12 to 19). Further, three calls were also move-grunts but produced while OT was looking away from the speaker. The remaining six calls were possibly eagle alarm calls (for which she was wrongly rewarded by the experimenter; see supplementary Video).

OT watched the video 56 times across all sessions, which was the highest count compared to the other subjects (Figure 3.5). There was a significant Spearman's correlation between cumulative successful calling and video-watching by OT (r = 0.91, p<0.001; Figure 3.6). In 10 cases, OT did not produce any audible vocalisations, but her body posture was as if she were to produce move-grunts (sitting and bending body forward towards the ground) with noticeable mouth and lip movements (see supplementary Video). These failed attempts were only observed during sessions 12 to 19, where she also produced at least one successful call. During 288 hrs of the sampling effort before this experiment, we recorded 189 natural vocalisations, out of which 17 were move-grunt calls. The natural frequency of move-grunt for a BD group was 0.05 move-grunt/hr.

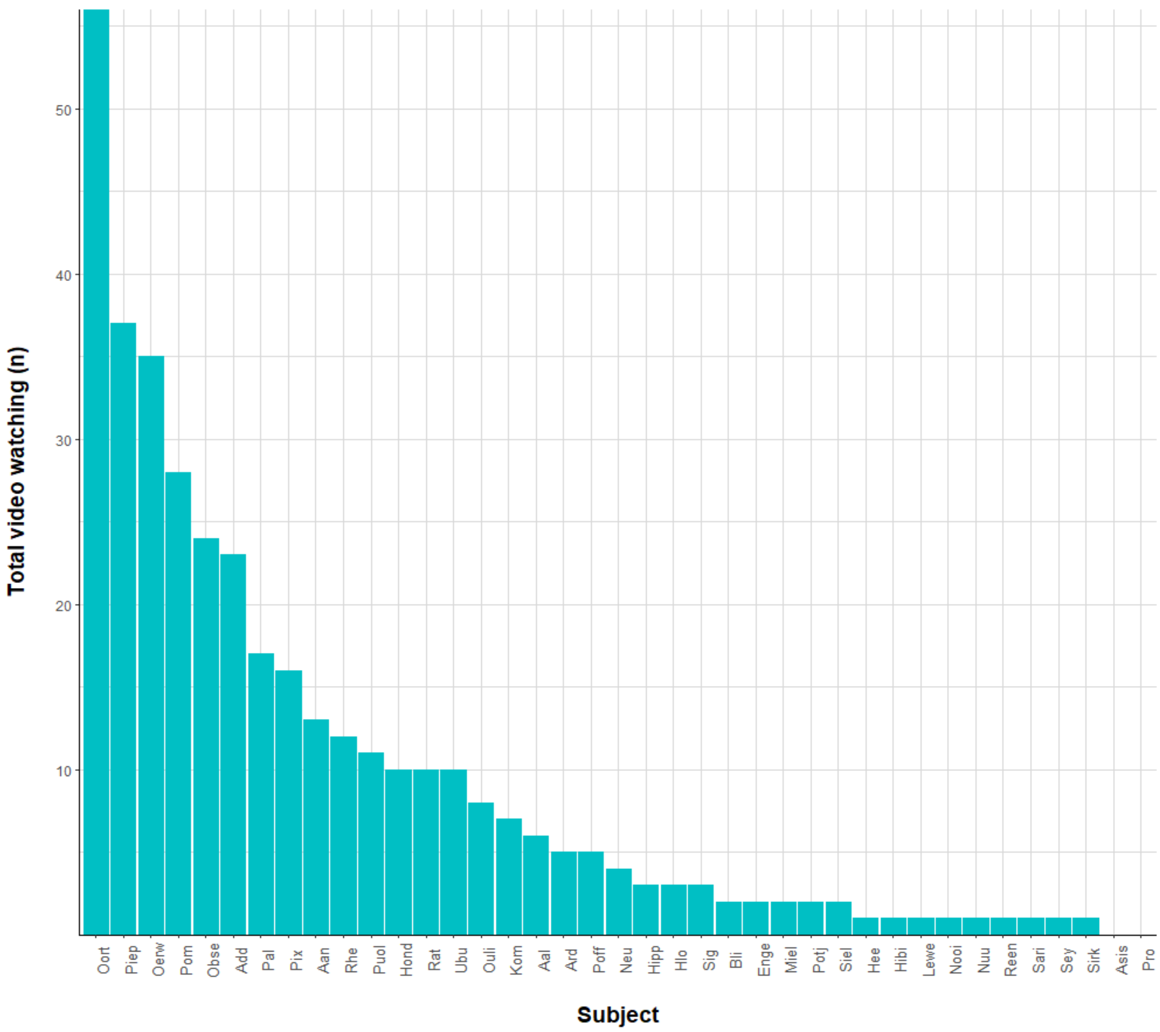

Figure 3.5 : Total number of times subjects watched the video in all the experimental session of Experiment 2: Social learning of audio-visual demonstrations

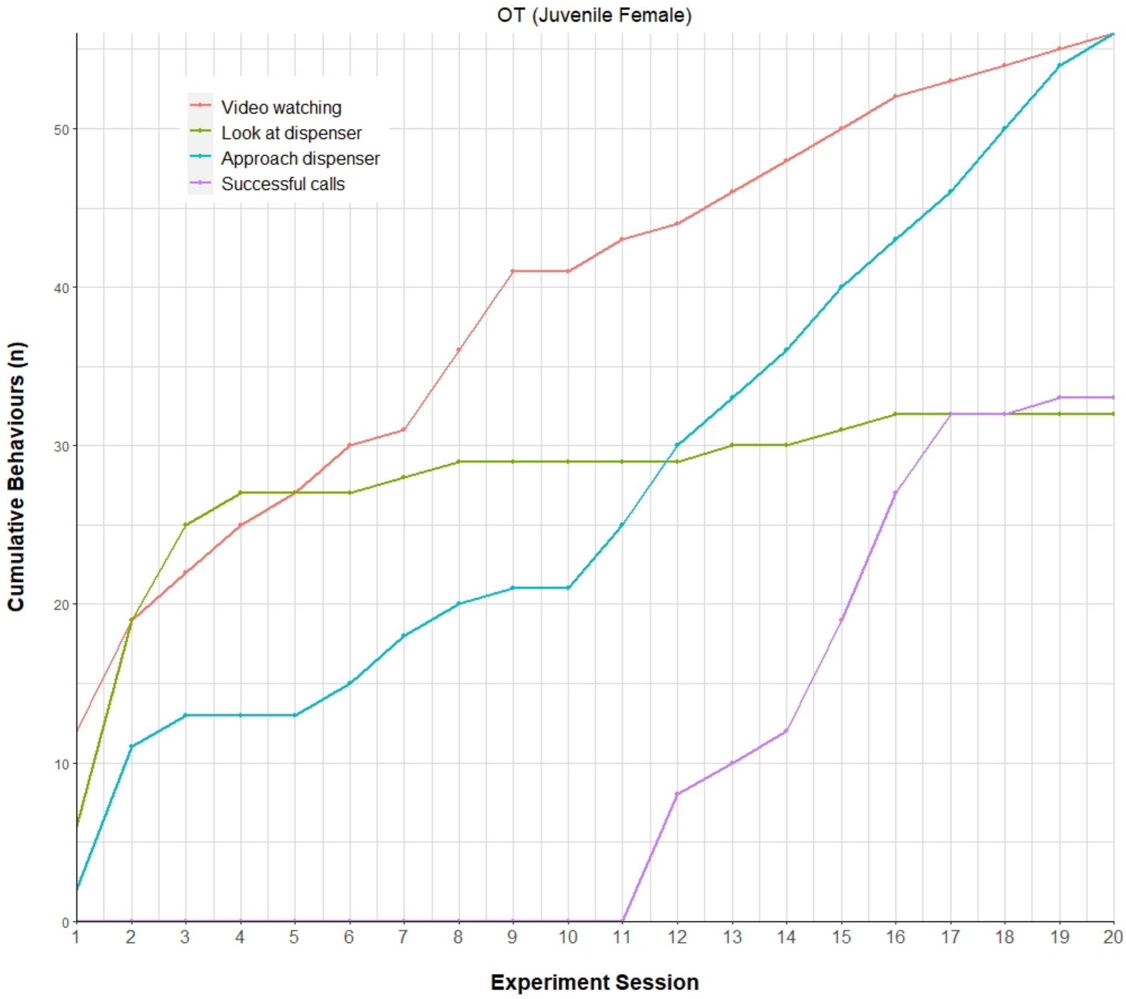

Figure 3.6 : Cumulative behaviours by subject OT in Experiment 2: Social learning of audio-visual demonstrations

3.6 Discussion

Human language is the product of human cognition, partially shared with animals, while speech appears to be uniquely human. No other species has the capacity to fluently combine phonemes into morphemes into phrases and sentences to not only interact socially but also to refer to real and imagined events [1]. This ultra-flexible sound production system is possible because of decoupling of vocal behaviour from direct biological function, insofar as humans are able to not only produce sounds that are linked to specific psychological states, such as distress is linked to crying, but additionally have the ability to produce sounds voluntarily [5,7]. Unhinged from biological function, human speech sounds are highly susceptible to social learning, which ultimately enables humans to establish countless signal-referent relations and connect with their minds.

This study attempted to explore the possible evolutionary roots of functionally decoupled and psychologically flexible calls in the primate lineage. To this end, we investigated if wild vervet monkeys can socially learn to produce affectively neutral move-grunt calls in an entirely novel context, i.e., to imperatively operate a food dispenser in their favour. To test this empirically, we provided wild vervet monkeys with opportunities to socially learn how to operate a food-dispenser, i.e., by producing spatially oriented move-grunts in order to activate the mechanism and receive a high-value food reward.

In a first experiment, we only provided a fragmentary glance of the event's causal structure, i.e., by playing another group member's move-grunts which was immediately followed by the release of food from the dispenser. Even though the event contained the key components of the causal chain, we did not find any evidence of learning in subjects. We concluded that a mere stimulus-response contingency was insufficient for subjects to adapt their own behaviour and successfully complete the task. In the second experiment, we featured the entire event, i.e. by showing a video clip of a monkey performing the entire event. Here we found that one monkey learned the contingency and spontaneously started to produce move-grunts in this novel condition, probably induced by watching the demonstration video.

In the first experiment, we played move-grunts from a hidden speaker followed by an immediate food reward from the dispenser. We expected that subjects would learn the association between move-grunt and food rewards and eventually start to produce move-grunt themselves to obtain the rewards. However, none of the subjects produced any move-grunt when stationed at the dispenser, suggesting that observing the stimulus-response contingencies was insufficient to learn a new call usage. Nevertheless, in three subjects that participated well, we observed behavioural pattern that indicated that some learning had taken place. On the first training day, monkeys were generally startled by the playback calls and looked at the speaker or the dispenser. On the following three days, however, the three subjects rapidly looked at the speaker first and then at the food-dispenser or the other way around. On the last three training days, subjects only looked at the dispenser and no longer at the speaker, indicating that the subjects learned the association between playback and subsequent food reward. Rapid auditory learning of call-event pairing has been found before in different primate and other mammal species [44–47].

In the second experiment, the subjects were able to watch a demonstration video of a conspecific successfully retrieving food by producing move grunts. The setup in this experiment was identical to the food dispenser nearby with which monkeys could interact freely. Out of 39 subjects who watched the video, one 4-year-old juvenile female (OT) started to produce move-grunt calls when stationed at the dispenser, a total of 33 times from session 12 onwards. Interestingly, she watched the video substantially more often than any other subject who participated in the experiment, suggesting that OT learned to produce a call in an arbitrary context by watching another individual doing the same, evidence for primary level usage learning [48].

An argument against this interpretation is that OT produced the first move-grunt for reasons unrelated to the video experience (e.g. by observing another individual moving at the same time) but then became conditioned due to the provisioning, learning from the video. However, this explanation is unlikely due to the fact that standard operant conditioning paradigms usually require enormous numbers of trials for subjects to become conditioned [22,23,25]. Furthermore, the natural rate of move-grunt producing is extremely low (1 move-grunt per 20 hours of observation per group), which makes it highly unlikely that OT produced her move-grunts due to a naturally adequate context. Finally, we did not observe any group movement during or after any successful move-grunts produced by OT. Also, OT did not move away with any group member immediately after any successful call. Thus, we conclude that our interpretation that OT learned a novel usage of move-grunts is much more plausible than alternative explanations. It is possible; however, that consistent reinforcement after move-grunt production strengthened the establishment of this new behaviour, especially in the beginning. In conclusion, this is the first experimental study of a socially learned novel call use in wild primates.

This finding is important because while earlier studies show that monkey vocalisations can be conditioned in controlled laboratory conditions [22,49–51], the operant conditioning paradigm used in such studies is not pertinent to the natural conditions in which primates live. We show that an alternative mechanism of observational social learning is likely to play an essential part in developing communicative competence [52]. Historically, social learning in primates has mainly been investigated in the acquisition of manual skills usually related to food acquisition [53,54]. But primates constantly monitor others during social interactions [55,56], suggesting that social learning could also play a role in becoming competent communicators. Future fine-grained studies on ontogeny and the interplay between social cognition and communication are required to test this possibility [57]. Social learning theory may have to be broadened to include all types of behaviour and communication [58–60].

Although the sample size is not optimal for comparing the outcomes of the two experiments, it is important to speculate why the demonstration video was probably more effective as a social learning stimulus than the playback of the call? Although the playbacks reliably predicted the availability of food, subjects were unable to see the causal connection, i.e., a conspecific uttering the call to trigger the production of food. Thus, the playback condition functions similarly to the 'ghost control' in manual tasks, with an invisible model, which may be cognitively more demanding for subjects [61]. The importance of an animate model with an agency in social learning is well studied in primates; however, it is generally in conjunction with manual action tasks [62]. Our experiments, probably for the first time, highlights the importance of witnessing a model with an agency to learn the communicative act.

Another unexpected observation in this study was OT's 'failed' (i.e., unvoiced) attempts to produce move-grunts, akin to the 'silent' vocalisations in the conditioning experiments with macaques [49,50]. During the 23 months of fieldwork on the same groups, we never observed anything similar in adults in any context (AD, personal obs). Previous attempts suggest it is difficult to train monkeys to vocalise on command [49,63]. In line with this, we noticed that OT also required significant physical efforts to elicit a successful call. Furthermore, we also observed that OT produced a few 'wrong' calls that resembled eagle alarm calls [64], despite the fact that there were no predators around and that OT did not show any anti-predatory behaviour. We speculate that OT was trying to produce a move-grunt but erroneously produced an alarm call.

These observations are also in line with recent findings on the neurology of primate call production. Coudé et al. identified vocalisation-selective neurons in the ventral premotor cortex, which fired only during the production of conditioned calls and remained inactive during silent vocalisations and spontaneous calls. Authors suggested that neural machinery required for voluntary vocalisation is not as fully formed as neural mechanisms involved in the spontaneous call in primates [49]. Notably, the volitional vocalisations also require greater efforts in motor preparation compared to volitional manual actions in primates [65]. In line with this, recent evidence suggests the correct response rate is much lower in vocal task compared to manual action task, probably due to separate neural networks for volitional vocalisations and manual actions in the ventrolateral prefrontal cortex [66]. This brain area is a homologue of the human Broca's area known to play a crucial role in human speech. The duality in neural networks might also explain the higher success of social learning for manual tasks than our results [37,38].

Overall, our results and behavioural observations support the neurobiological evidence and suggest that non-human primates can probably socially learn the novel call usage from appropriate virtual stimuli. However, they have limited cognitive control over calls due to a lack of neural machinery, making it rather difficult to elicit a pre-existing call in an emotionally and functionally novel context. Contrary to our initial prediction, a possibility that move-grunts are not completely decoupled from the specific affective state could also explain the difficulty in its production in a novel context. In other words, we cannot discount the possibility that monkeys might have understood the task by watching the demonstration video; however, they were not equipped to perform it [67].

To our knowledge, this is the first study to use the social learning paradigm to test for functional flexibility of vocalisations in wild primates experimentally. We acknowledge that our sample size is small and lack rigorous statistical analyses. Results can be considered suggestive as only one subject seemed to learn the novel usage. However, it is comparable to other experiments conducted in captivity [22,24,68]. Furthermore, we have also provided comprehensive qualitative behavioural observations, which could be useful for future studies, and opened a new possible experimental paradigm that can be used on highly habituated and provisioned populations of different primate species. Future studies with an improved experimental design are required to confirm our results. In the future, researchers can use multiple remote-controlled food dispensers, which might reduce any effects arising from the experimenter's proximity and improve the participation of subjects. A close-up camera and audio recorder that can record facial and mouth movements would help record behaviours like silent vocalisations and calls systematically. Acoustic analyses of the elicited calls were irrelevant for this study as rewards were based on the experimenter's immediate classification of the call, and bad weather conditions did not yield good recordings. However, post-hoc acoustic analysis of the well-recorded volitional calls is critical to understanding the differences in calls elicited in natural and novel contexts [49].

Furthermore, we argue that experiments in captive settings that test for the cognitive control of vocalisation with neuronal recordings can adopt the social learning paradigm. It is ecologically more relevant than the previously used operant conditioning, and we speculate that it would provide substantial insights into the neural mechanism of volitional call production. Finally, similar experiments with anthropoid apes and prelinguistic children are necessary for the comparative analyses to understand the evolution of functionally flexible sounds in the primate lineage, offering fundamental insights into human language evolution.

3.7 Author contribution

AD conceived the study with help from KZ. AD conducted the experiments analyses and wrote the first draft, which KZ reviewed. AD and KZ wrote the final version of the manuscript.

3.8 Acknowledgements

AD would like to thank Maria Teresa Navarette and Mábia Biff Cera for suggestions on the experimental design of Experiment 1. Finally, we would like to thank all the volunteers, researchers, and managers at the Inkawu vervet project for their help during data collection.

3.9 References

1. Fitch WT. 2000 The evolution of speech: A comparative review. Trends Cogn. Sci. 4, 258–267. (doi:10.1016/S1364-6613(00)01494-7)
2. 1999 Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambrifge University Press.
3. Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, Bakeman R. 2013 Functional flexibility of infant vocalization and the emergence of language. Proc. Natl. Acad. Sci. U. S. A. 110, 6318–23. (doi:10.1073/pnas.1300337110)
4. Hauser MD, Chomsky N, Fitch WTS. 2002 The faculty of language: what is it, who has it, and how did it evolve? Science (80-.). 298, 1569–79. (doi:10.1126/science.298.5598.1569)
5. Jhang Y, Oller DK. 2017 Emergence of functional flexibility in infant vocalizations of the first 3 months. Front. Psychol. 8. (doi:10.3389/fpsyg.2017.00300)
6. Kersken V, Zuberbühler K, Gomez JC. 2017 Listeners can extract meaning from non-linguistic infant vocalisations cross-culturally. Sci. Rep. 7, 1–7. (doi:10.1038/srep41016)
7. Griebel U, Oller DK. 2008 Evolutionary forces favoring communication flexibility. Evol. Commun. Flex. Complexity, Creat. Adapt. Hum. Anim. Commun. , 9–40. (doi:10.1093/icb/icp087)
8. Steels L, Kaplan F. 2001 AIBO's first words: The social learning of language and meaning. Evol. Commun. 4, 3–32. (doi:10.1075/eoc.4.1.03ste)
9. Tomasello M. 2010 Origins of Human Communication. MIT press.
10. Owren MJ, Rendall D. 2001 Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. Evol. Anthropol. 10, 58–71. (doi:10.1002/evan.1014)
11. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. Neurosci. Biobehav. Rev. 82, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)
12. Snowdon C. 2018 Cognitive Components of Vocal Communication: A Case Study. Animals 8, 126. (doi:10.3390/ani8070126)
13. Laporte MNC, Zuberbühler K. 2011 The development of a greeting signal in wild chimpanzees. Dev. Sci. 14, 1220–1234. (doi:10.1111/j.1467-7687.2011.01069.x)
14. Cheney DL, Seyfarth RM. 2018 Flexible usage and social function in primate vocalizations. Proc. Natl. Acad. Sci. U. S. A. 115, 1974–1979. (doi:10.1073/pnas.1717572115)
15. Clay Z, Archbold J, Zuberbühler K. 2015 Functional flexibility in wild bonobo vocal behaviour. PeerJ 3, e1124. (doi:10.7717/peerj.1124)
16. Dezecache G, Zuberbühler K, Davila-ross M. 2020 Flexibility in wild infant chimpanzee vocal behavior. , 1–17. (doi:10.1093/jole/lzaa009)

17. Hihara S, Yamada H, Iriki A, Okanoya K. 2003 Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. Neurosci. Res. 45, 383–389. (doi:10.1016/S0168-0102(03)00011-7)

18. Deshpande A, Gupta S, Sinha A. 2018 Intentional communication between wild bonnet macaques and humans. Sci. Rep. 8. (doi:10.1038/s41598-018-22928-z)

19. Hsu MJ, Chen LM, Agoramoorthy G. 2005 The vocal repertoire of Formosan macaques, Macaca cyclopis: Acoustic structure and behavioral context. Zool. Stud. 44, 275–294.

20. Arnold K, Zuberbühler K. 2013 Female Putty-Nosed Monkeys Use Experimentally Altered Contextual Information to Disambiguate the Cause of Male Alarm Calls. PLoS One 8, 2–8. (doi:10.1371/journal.pone.0065660)

21. Fichtel C. 2020 Monkey Alarm Calling: It Ain't all Referential, or is It? Anim. Behav. Cogn. 7, 101–107. (doi:10.26451/abc.07.02.04.2020)

22. Sutton D, Larson C, Taylor EM, Lindeman RC. 1973 Vocalization in rhesus monkey: Conditionability. Brain Res. 52, 225–231.

23. Aitken PG, Wilson WA. 1979 Discriminative vocal conditioning in rhesus monkeys: Evidence for volitional control? Brain Lang. 8, 227–240. (doi:10.1016/0093-934X(79)90051-8)

24. Koda H, Oyakawa C, Kato A, Masataka N. 2007 Experimental evidence for the volitional control of vocal production in an immature gibbon. Behaviour 144, 681–692. (doi:10.1163/156853907781347817)

25. Hage SR, Gavrilov N, Nieder A. 2013 Cognitive Control of Distinct Vocalizations in Rhesus Monkeys. J. Cogn. Neurosci. 25, 1692–1701. (doi:10.1162/jocn_a_00428)

26. Mineka S, Cook M. 1993 Mechanisms involved in the observational conditioning of fear. J. Exp. Psychol. Gen.

27. van de Waal E, Borgeaud C, Whiten A. 2013 Potent Social Learning and Conformity Shape a Wild Primate's Foraging Decisions. Science (80-.). 340, 483–485. (doi:10.1126/science.1232769)

28. Kendal RL, Boogert NJ, Rendell L, Laland KN, Webster M, Jones PL. 2018 Social Learning Strategies: Bridge-Building between Fields. Trends Cogn. Sci. 22, 651–665. (doi:10.1016/j.tics.2018.04.003)

29. Canteloup C, Waal E Van De, Gonzalez IP, Sueur C. 2020 The effects of data collection and observation methods on uncertainty of social networks in wild primates. , 1–12. (doi:10.1002/ajp.23137)

30. Mertz J, Surreault A, van de Waal E, Botting J. 2019 Primates are living links to our past: The contribution of comparative studies with wild vervet monkeys to the field of social cognition. Cortex 118, 65–81. (doi:10.1016/j.cortex.2019.03.007)

31. Whiten A, van de Waal E. 2018 The pervasive role of social learning in primate lifetime development. Behav. Ecol. Sociobiol. 72. (doi:10.1007/s00265-018-2489-3)

32. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey (Cercopithecus aethiops) alarm calls: semantic communication in a free ranging primate. Anim. Behav. 28, 1070–1094.

33. Struhsaker TT. 1967 Ecology of vervet monkeys (Cercopithecus aethiops) in The Masai-Amboseli Game Reserve, Kenya. Ecology 48, 891–904. (doi:10.2307/1934531)

34. Seyfarth RM, Cheney DL. 1986 Vocal development in vervet monkeys. Anim. Behav. 34, 1640–1658. (doi:10.1016/S0003-3472(86)80252-4)

35. Seyfarth RM, Cheney DL. 2010 Some general features of vocal development in nonhuman primates. Soc. Influ. Vocal Dev. , 249–273. (doi:10.1017/cbo9780511758843.013)

36. Cheney DL, Seyfarth RM. 1982 How vervet monkeys perceive their grunts: Field playback experiments. Anim. Behav. 30, 739–751. (doi:10.1016/S0003-3472(82)80146-2)

37. Price E, Caldwell CA. 2007 Artificially generated cultural variation between two groups of captive monkeys, Colobus guereza kikuyuensis. Behav. Processes 74, 13–20. (doi:10.1016/j.beproc.2006.09.003)

38. Gunhold T, Whiten A, Bugnyar T. 2014 Video demonstrations seed alternative problem-solving techniques in wild common marmosets. Biol. Lett. 10. (doi:10.1098/rsbl.2014.0439)

39. Canteloup C, Hoppitt W, van de Waal E. 2020 Wild primates copy higher-ranked individuals in a social transmission experiment. Nat. Commun. 11, 459. (doi:10.1038/s41467-019-14209-8)

40. Center for Conservation Bioacoustics. 2014 Raven Pro: Interactive Sound Analysis Software.

41. S. EC, Pfungst O. 1908 Das Pferd des Herrn von Osten (Der kluge Hans), ein Beitrag zur experimentellen Tier- und Menschen-Psychologie. Am. J. Psychol. 19, 422. (doi:10.2307/1413208)

42. Altmann J. 1974 Observational study of behavior: sampling methods. Behavior 49, 227–267. (doi:10.1163/156853974X00534)

43. Team RDC. 2020 R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. Vienna, Austria. ISBN 3-900051-07-0.

44. Wegdell FL, Hammerschmidt K, Fischer J. 2019 Conserved alarm calls but rapid auditory learning in monkey responses to novel flying objects. Nat. Ecol. Evol. (doi:10.1038/s41559-019-0903-5)

45. Petkov CI, Jarvis ED. 2012 Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. Front. Evol. Neurosci. 4, 1–24. (doi:10.3389/fnevo.2012.00012)

46. Hauser MD. 1988 How infant vervet monkeys learn to recognize starling alarm calls: the role of experience. Behaviour 105, 187–201. (doi:10.1163/156853988X00016)

47. Zuberbühler K. 2000 Interspecies semantic communication in two forest primates. Proc. R. Soc. B Biol. Sci. 267, 713–718. (doi:10.1098/rspb.2000.1061)

48. Shapiro AD, Slater PJB, Janik VM. 2004 Call usage learning in gray seals (Halichoerus grypus). J. Comp. Psychol. 118, 447–454. (doi:10.1037/0735-7036.118.4.447)

49. Coudé G, Ferrari PF, Rodà F, Maranesi M, Borelli E, Veroni V, Monti F, Rozzi S, Fogassi L. 2011 Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. PLoS One 6, 1–10. (doi:10.1371/journal.pone.0026822)

50. Aitken PG. 1981 Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. Brain Lang. 13, 171–184. (doi:10.1016/0093-934X(81)90137-1)

51. Ackermann H, Hage SR, Ziegler W. 2014 Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. Behav. Brain Sci. 37, 529–546. (doi:10.1017/S0140525X13003099)

52. Laland KN. 2004 Social learning strategies. Learn. Behav. 32, 4–14.

53. Bono AEJ, Whiten A, van Schaik C, Krützen M, Eichenberger F, Schnider A, van de Waal E. 2018 Payoff- and Sex-Biased Social Learning Interact in a Wild Primate Population. Curr. Biol. 28, 2800-2805.e4. (doi:10.1016/j.cub.2018.06.015)

54. Van de Waal E, Bshary R, Whiten A, Waal E Van De. 2014 Wild vervet monkey infants acquire the food-processing variants of their mothers. Anim. Behav. 90, 41–45. (doi:10.1016/j.anbehav.2014.01.015)

55. Maciej P, Patzelt A, Ndao I, Hammerschmidt K, Fischer J. 2013 Social monitoring in a multilevel society: A playback study with male Guinea baboons. Behav. Ecol. Sociobiol. 67, 61–68. (doi:10.1007/s00265-012-1425-1)

56. Seyfarth RM, Cheney DL. 2015 Social cognition. Anim. Behav. 103, 191–202. (doi:10.1016/j.anbehav.2015.01.030)

57. Seyfarth RM, Cheney DL. 2014 The evolution of language from social cognition. Curr. Opin. Neurobiol. 28, 5–9. (doi:10.1016/j.conb.2014.04.003)

58. Heyes C, Pearce JM. 2015 Not-so-social learning strategies. (doi:10.1098/rspb.2014.1709)

59. Heyes C. 2016 Who Knows? Metacognitive Social Learning Strategies. Trends Cogn. Sci. 20, 204–213. (doi:10.1016/j.tics.2015.12.007)

60. Lind J, Ghirlanda S, Enquist M. 2018 Social learning through associative processes: A computational theory. bioRxiv (doi:10.1101/446906)

61. Hopper LM. 2010 'Ghost' experiments and the dissection of social learning in humans and animals. Biol. Rev. 85, 685–701. (doi:10.1111/j.1469-185X.2010.00120.x)

62. Hopper LM, Lambeth SP, Schapiro SJ, Whiten A. 2015 The importance of witnessed agency in chimpanzee social learning of tool use. Behav. Processes 112, 120–129. (doi:10.1016/j.beproc.2014.10.009)

63. Pierce JD. 1985 A review of attempts to condition operantly alloprimate vocalizations. Primates 26, 202–213. (doi:10.1007/BF02382019)

64. Price T, Wadewitz P, Cheney DL, Seyfarth RM, Hammerschmidt K, Fischer J. 2015 Vervets revisited: A quantitative analysis of alarm call structure and context specificity. Sci. Rep. 5, 1–11. (doi:10.1038/srep13220)

65. Koda H, Kunieda T, Nishimura T. 2018 From hand to mouth: Monkeys require greater effort in motor preparation for voluntary control of vocalization than for manual actions. R. Soc. Open Sci. 5. (doi:10.1098/rsos.180879)

66. Gavrilov N, Nieder A. 2021 Distinct neural networks for the volitional control of vocal and manual actions in the monkey homologue of Broca's area. Elife 10, 1–24.

67. Tecumseh Fitch W. 2020 Animal cognition and the evolution of human language: Why we cannot focus solely on communication. Philos. Trans. R. Soc. B Biol. Sci. 375. (doi:10.1098/rstb.2019.0046)

68. Pomberger T, Risueno-Segovia C, Gultekin YB, Dohmen D, Hage SR. 2019 Cognitive control of complex motor behavior in marmoset monkeys. Nat. Commun. 10, 1–8. (doi:10.1038/s41467-019-11714-8)

3.10 Chapter 3: Appendix

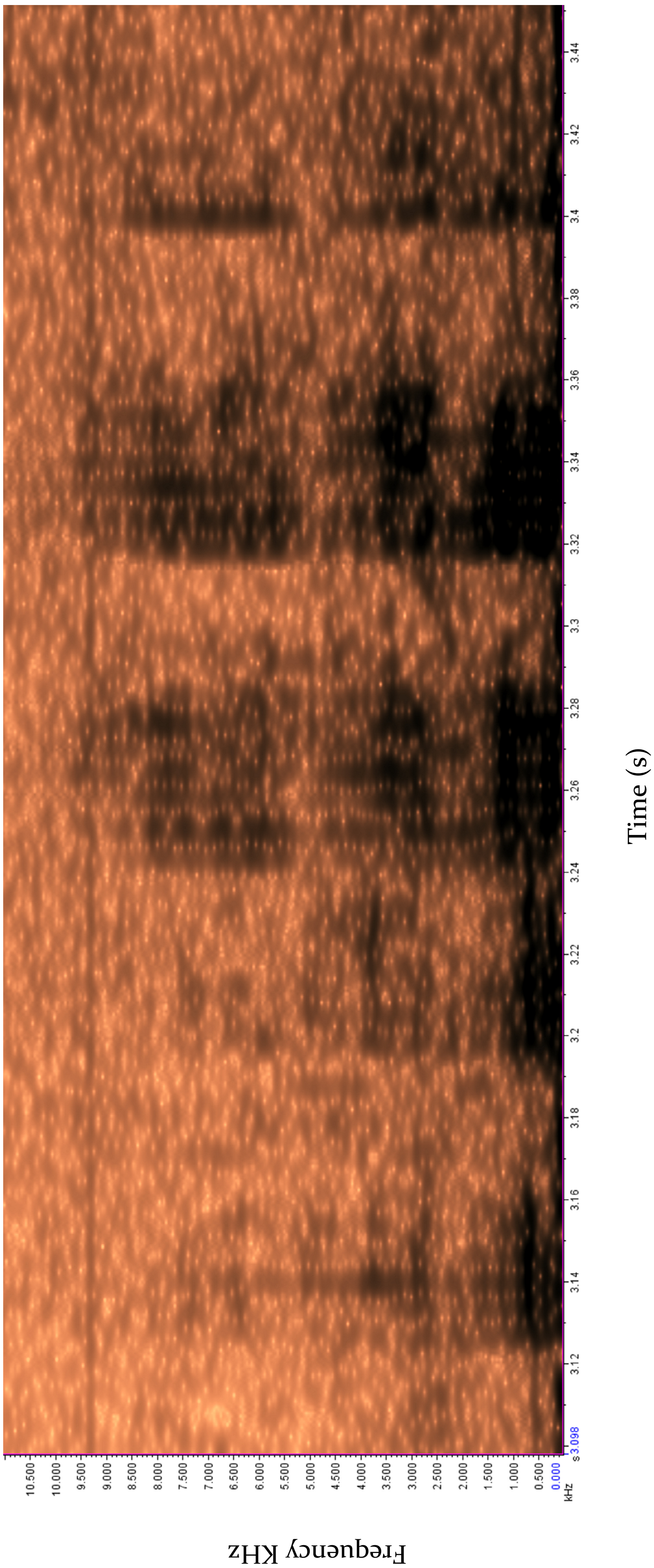

Figure A3.1: Spectrogram (Hamming window at 1024 DFT and 93.8% overlap) of the naturally recorded move-grunt call used for the Experiments

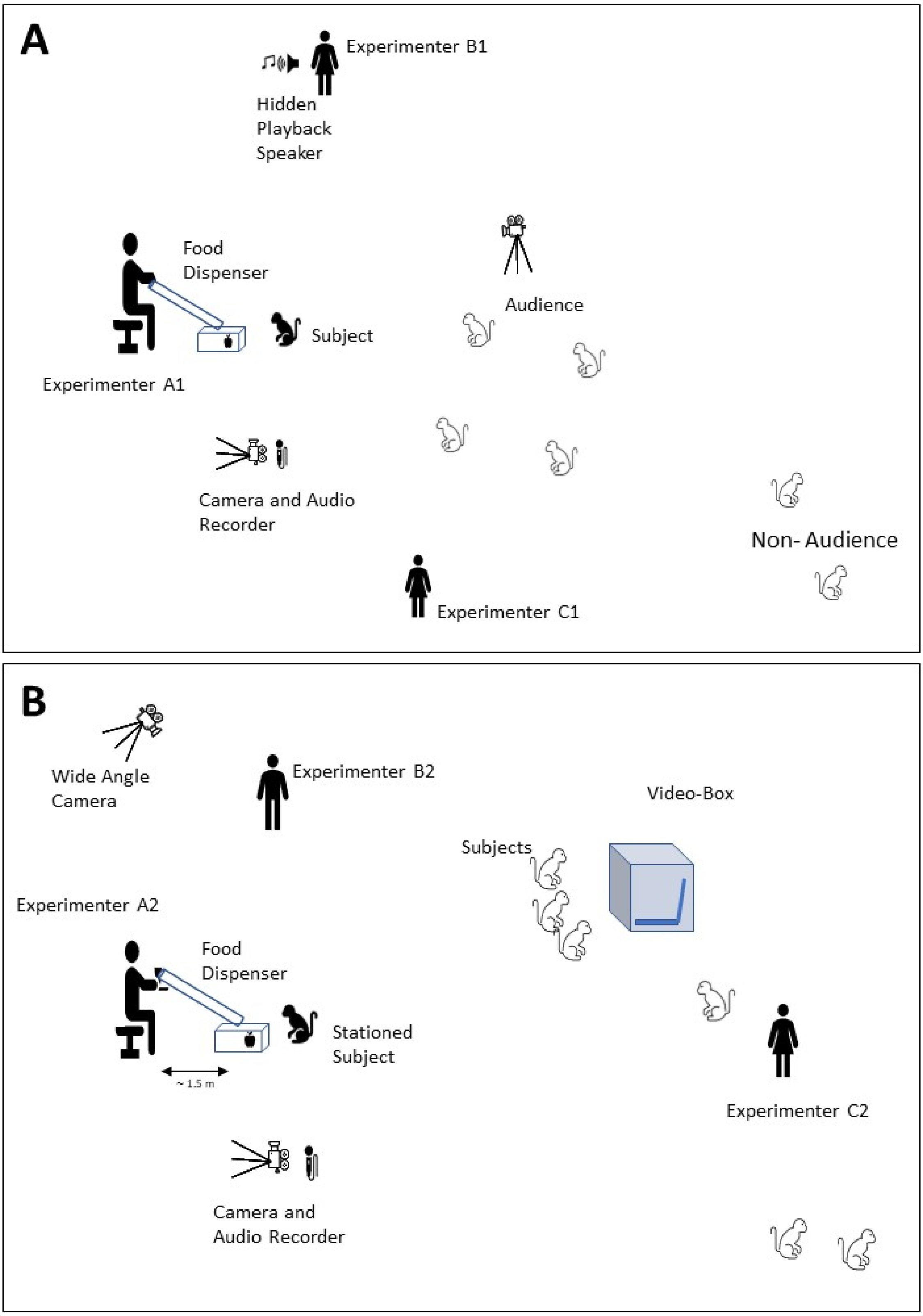

Figure A3.2: Heuristic illustrations (not to scale) of experimental setups (A) Experiment 1: Social learning of auditory demonstrations, (B) Experiment 2: Social learning of audio-visual demonstrations

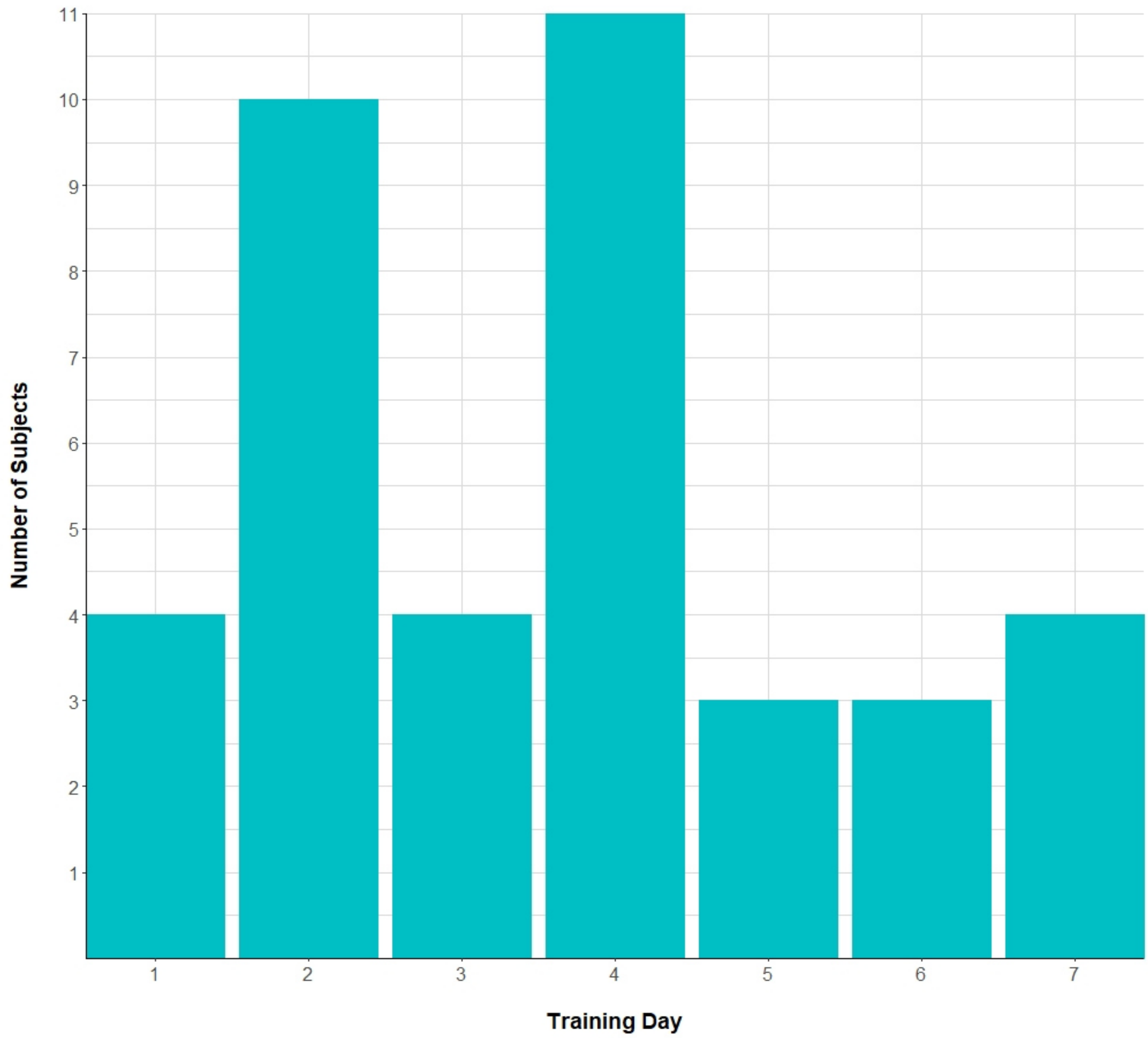

Figure A3.3: Number of subjects participated per training day for Experiment 1: Social learning of auditory demonstrations

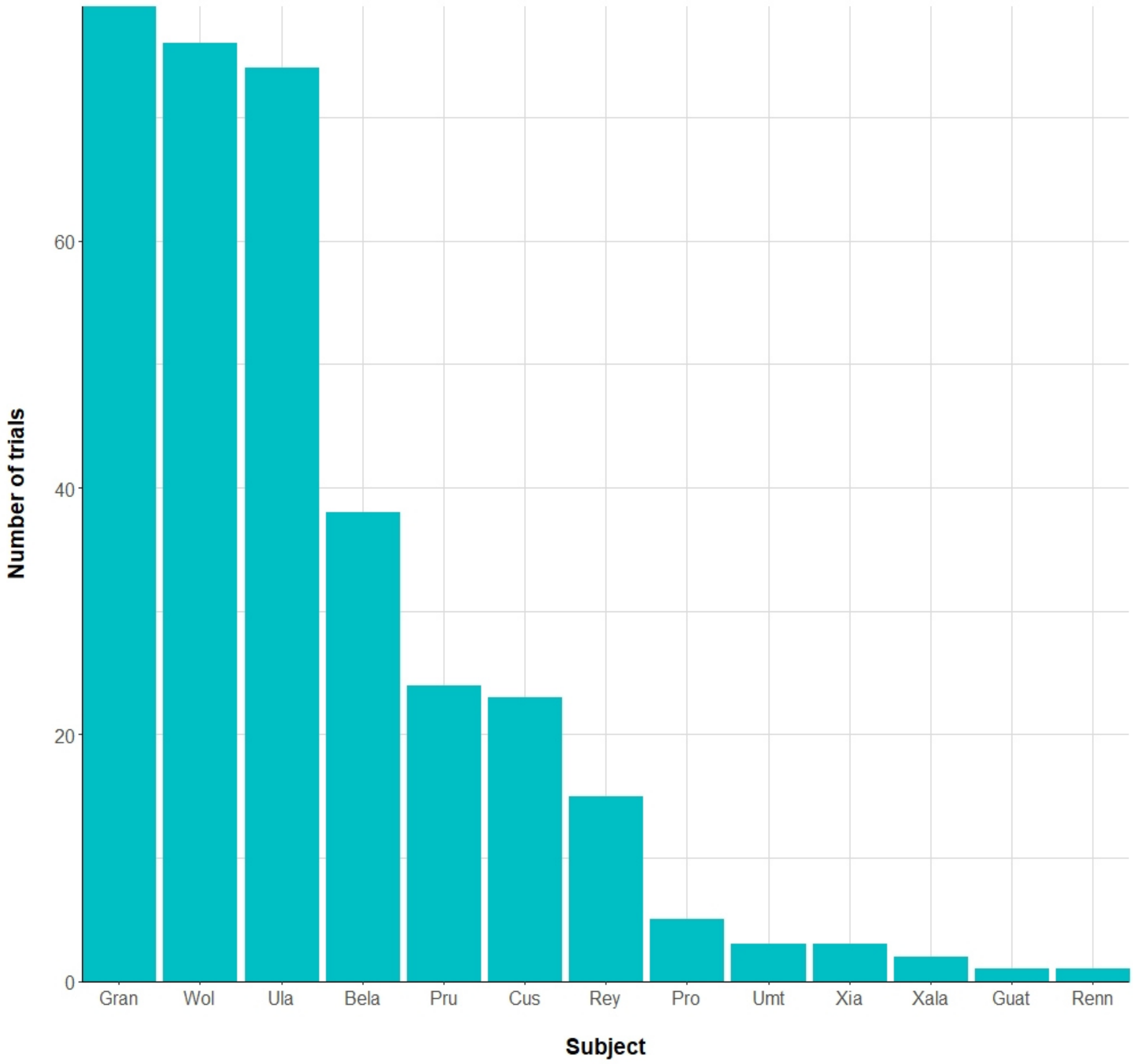

Figure A3.4: Subject wise participation in Experiment 1: Social learning of auditory demonstrations

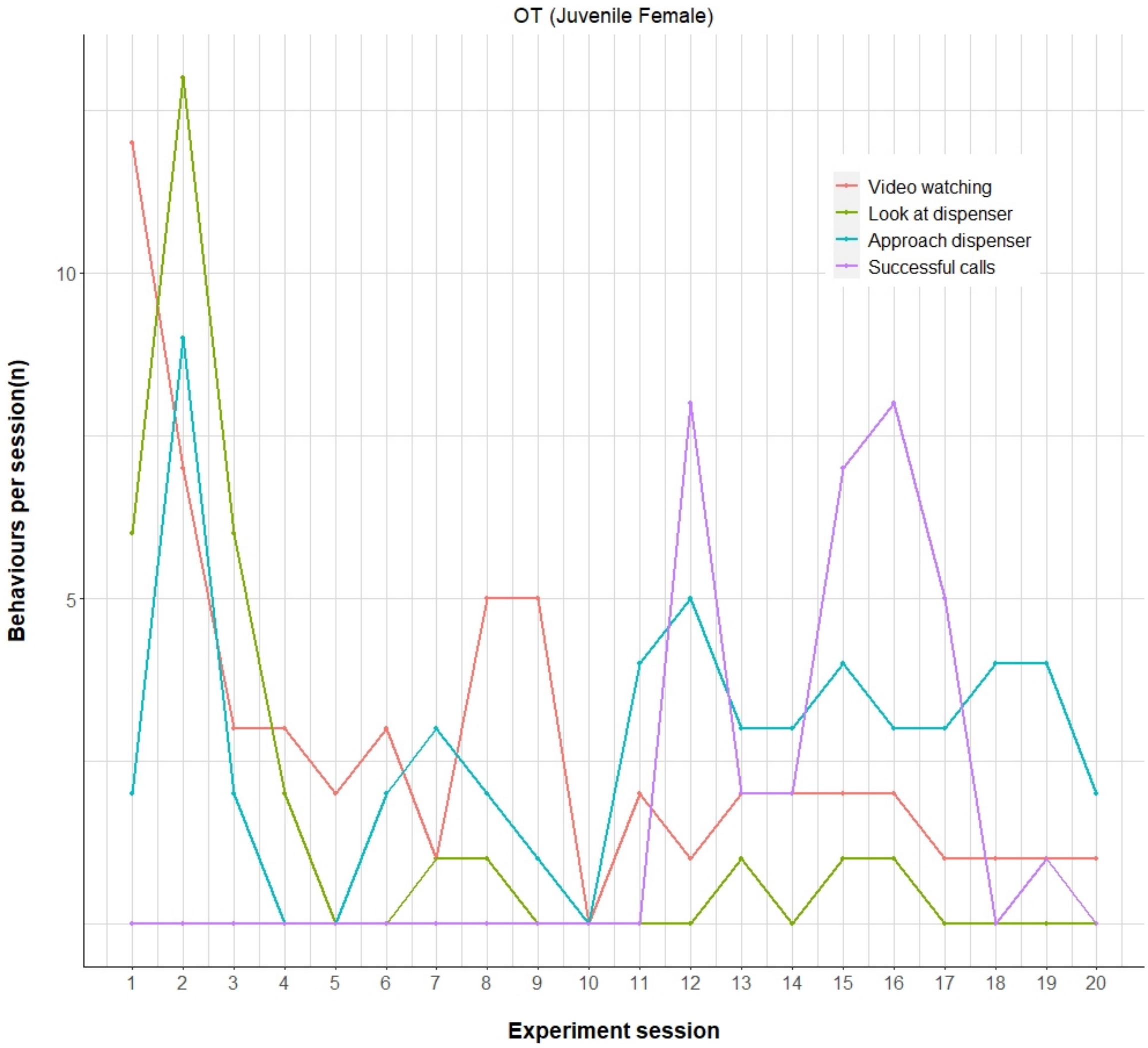

Figure A3.5: Focal behaviours by subject OT across sessions in Experiment 2: Social learning of audio-visual demonstrations

Chapter 4: Pragmatic inference in wild vervet monkeys

Adwait Deshpande[1,2,*] and Klaus Zuberbuhler[1,2,3]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchatel, Neuchatel, Switzerland, [2] Inkawu Vervet Project, Mawana Game Reserve, KwaZulu-Natal, South Africa, [3] School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom

4.1 Abstract

In human language, meaning resides in individual signals, syntactic combinations and contextual use, which has led to a distinction between lexical, compositional and pragmatic meaning. While there is evidence for the first two processes in animal communication, pragmatic meaning has not yet been investigated systematically. Male vervet monkeys produce the same alarm calls during encounters with terrestrial predators and neighbouring groups, suggesting that recipients must take the context into account to respond in an adaptive way. We empirically tested this hypothesis by playing male alarm calls during intergroup encounters or during neutral situations. We found that subjects responded in context-appropriate ways to the alarm calls, with significant differences in startle responses, vigilance behaviour and gazes towards the presumed caller. We concluded that non-human primates could derive meaning pragmatically, a capability that may be more widespread in animal communication.

4.2 Keywords

Pragmatics; Evolution of Language; Vocal communication; Non-human primates

4.3 Introduction

An important milestone in the Darwinian quest for evolutionary continuities has been the discovery that vervet monkeys (*Chlorocebus pygerythrus hilgerti*) produce acoustically distinct alarm calls to refer to different external events [1,2], which has led to the suggestion that animal alarm calls can function as semantic signals [3,4]. The evidence for referential or semantic signals is typically in the form of acoustically distinct calls that are produced in context-specific ways, such as monkey 'eagle' alarm calls given to particular species of eagles [5–10]. However, with the advanced acoustic analyses, the tight links between calls and their eliciting contexts often begin to dissolve. What initially looked like a predator-specific alarm call turns out to be a call that can also be given in other, often multiple contexts [11]. Three types of explanations have been offered to explain such phenomena. A first possibility is that callers sometimes make mistakes and make erroneous assumptions about reality, leading to inappropriate vocal behaviour. A second possibility is that animals occasionally deceive their audience, for example, by giving 'false' alarm calls or food calls [12,13]. A third possibility, and the topic of this study, is that animal communication is not limited to 'literal' meaning [14,15] (i.e., residing within single signals), but that meaning is pragmatically derived by taking context into account when producing and responding to signals [16].

In linguistic analyses, pragmatics refers to the study of meaning within context, in contrast to lexical or compositional semantics, where meaning resides within the units themselves [17]. For example, the sentence "What time is it?" can either be a simple request for information or an imperative, equivalent to "You are late!" [18]. In order to extract the intended speaker meaning, an interlocutor needs advanced cognitive abilities, such as mental state or common ground assessments, to decide whether the utterance was meant as a request or expression of disapproval [16]. The underlying psychological abilities are thought to be complex because they require the integration of multiple strands of information, something artificial conversational agents severely struggle with [19]. Arguably, however, most natural human communication is founded in pragmatics, whereas lexical and compositional semantics may act as its building blocks [20,21].

Despite considerable research efforts, non-human primates have not particularly excelled with demonstrating advanced meta-psychological abilities, such as mental state attribution, joint commitment or shared intentionality [22] (but see [23–26] for evidence in great apes). But how then has natural selection favoured the evolution of signals with ambiguous meanings as documented in primate communication? And how are recipients able to attribute meaning to them? One suggestion is that non-human primates may use much simpler forms of pragmatics that do not require deep psychological assessments but rely on simple accounting of ongoing context [27]. In such a simpler form, the receiver does not need to determine the caller's mental state but attributes meaning by taking into account contextual cues. For example, putty-nosed monkeys show less vigilance to terrestrial alarm calls if the calls are preceded by the sound of a falling tree compared to no preceding cues [28]. Similarly, Diana monkey responses to heterospecific alarm calls are based on prior information about probable cause [29,30]. Capuchin monkeys ignore alarm calls given during competitive feeding events but respond strongly to the same calls in other situations [31], and baboons respond to grunts depending on the context in which they were produced [32]. Taken together, these studies suggest primates can take into account simple social and environmental cues into account to infer the most probable cause of call production, i.e., causal inference becomes the basis for the attribution of meaning.

Here, we revisit the alarm call system of vervet monkeys, the classic example of 'lexical' semantics in animal communication, to investigate whether listeners can make inferences about the call-eliciting context rather than directly responding to the calls themselves. In a South African subspecies (*Chlorocebus pygerythrus pygerythrus*), males have been observed to produce 'leopard' alarm calls (i.e., 'alarm barks') to terrestrial predators but also during within and between-group aggression [11]. This raises the possibility that receivers are forced to infer the meaning of the male 'leopard' alarms based on the available contextual information. In line with this hypothesis is a recent study in vervet monkeys that has demonstrated flexible responses to alarm calls depending on group size and distance from the caller [33].

An alternative hypothesis is that males produce context-specific acoustic variants of alarm calls to terrestrial predators and within-species aggression, in which case recipients could directly attend to the signals to infer meaning, regardless of context. In line with this, recent acoustic analyses of vervet monkey alarm calls have revealed a graded system with overlap between different call-eliciting contexts, a pattern also found in Campbell's monkey alarm calls [34]. Research in barbary macaques and other species has shown that non-human primates, like humans, can perceive acoustically graded signals in categorical ways [35], adding plausibility to the acoustic variant hypothesis.

We empirically tested the two hypotheses against each other in a free-ranging South African population of vervet monkeys during natural intergroup encounters. Intergroup encounters are regular but highly variable social events that can range from peaceful co-foraging to episodes of severe physical aggression between the two groups [36]. Remarkably, during aggressive episodes, males from both groups sometimes produce alarms that acoustically resemble 'leopard alarms', along with other aggressive vocalisations, suggesting that males experience a similar state of arousal when dealing with hostile neighbours and terrestrial predators [11,37], despite the fact that the two situations require radically different behavioural responses. Participation in intergroup aggression usually varies between group members but is most common in high-ranked individuals and may also depend on their prior social interactions [36,38,39]. As a result, there are different subgroups of individuals who participate in frontline aggression during each encounter while others stay back and appear to remain uninterested. We took advantage of this fact to examine whether disengaged individuals use contextual information when hearing the male 'leopard' alarms of frontline males during intergroup encounters. We predicted that if non-participating individuals took context into account, they should respond differently to experimentally presented male predator alarms during ongoing intergroup encounters compared to control situations devoid of such contextual information.

4.4 Methods

4.4.1 Study site and groups

The study was conducted between July- November 2019 at the Inkawu Vervet Project (IVP), located on the Mawana Game Reserve, Kwazulu Natal, South Africa (S 28° 00.327; E 031° 12.348). All human observers complete a standard training protocol and complete inter-observer reliability and individual identification test before participating in data collection. Researchers have access to several neighbouring groups of wild South African vervet monkeys that are well habituated to human presence and allow close observations by multiple observers. Data for this study were collected from one group (BD), consisting of 57 individuals (adult females: N=16; adult males: N= 12; juvenile females: N=15; juvenile males: N= 14).

4.4.2 Experimental procedure

We compared the effects of experimentally presenting 'leopard alarms' by a familiar adult male of the same group on subjects engaged in either an intergroup encounter (experimental trials) or a neutral situation (control trials), using a within-subject design (N=8 subjects; N=16 trials).

A pilot study conducted on the same population a few years earlier suggested that male 'leopard' alarms produced during intergroup encounters were acoustically similar to the alarms produced during terrestrial predator encounters [40]. To address the hypothesis that male alarms may consist of context-specific subtle acoustic variants, we recorded 'leopard' alarm calls of an adult male (age > six years) in both contexts, i.e., during a natural encounter with black-backed jackal (*Lupulella mesomelas*), a common terrestrial primate predator, six months prior to the beginning of the experiment, and during an aggressive interaction with a neighbouring group (AK), four months prior to the beginning of the experiment. Both playback calls were equally and randomly assigned across subjects. However, for each subject, the same playback call was used for experimental and control trials.

We edited the original recordings such that the playback stimuli were comparable in term of basic parameters (Figure 4.1). We removed background noise and padded the calls with two seconds of silence before and after the utterances using Raven Pro 1.5 [41]. Recordings were played back using a Motorola XPlay phone and an Anchor AN-MINI-RST-01 speaker. Call amplitudes were adjusted and kept constant across trials, such that they matched natural calls at a distance of 30-50m. The call providing male was present in the group for the entire duration of the experiment.

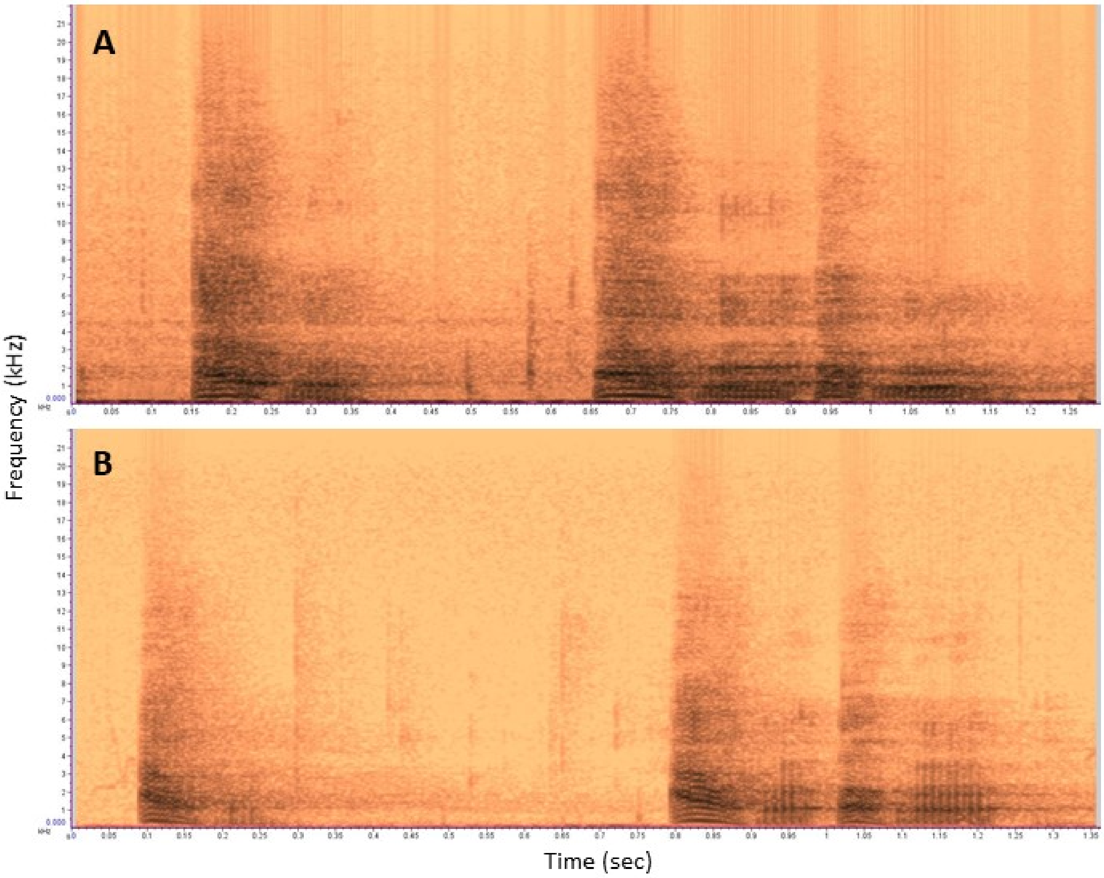

Figure 4.1 Spectrograms (Hamming window at 1024 DFT and 93.8% overlap) of the exemplars of leopard alarm calls of a male from BD group which were used as playback calls for the experiment (A) Exemplar of a leopard alarm call recorded during an intergroup encounter

4.4.3 Experimental trials

Since intergroup encounters occurred unpredictably, we conducted experimental trials opportunistically, provided the following conditions were met. First, we only initiated a trial after visible intergroup aggression. This was because male 'leopard' alarms were usually given during intergroup encounters that involved physical aggression between the two groups. The caller usually actively participated during intergroup aggressions and most likely remained on the frontline (see below). Second, previous studies [33] suggested that habitat type and the group spread affected responses to alarm calls, so we only conducted experiments in mixed habitat type (i.e., canopy cover >50% to <100%) and group spread (approximately 100 m, i.e., distance between subject and all other group members no more than 100m).

An intergroup encounter started when members of another group were within 50 m of any individual of the study group (BD) or if BD individuals were vocally reacting to another group with grunts, screams or chutter calls [42]. Once the encounter had started, we waited for at least 5min before carrying out a trial. We only conducted a trial if there were between-group aggressive interactions during this period. However, if any male from either group produced a 'leopard' alarm during this period, we aborted the trial.

Two observers were present during each experiment. The first observer positioned herself on the front line of the encounter, i.e., the imaginary line that separated the two groups, where most between-group interactions happen [36,38,39]. All individuals within 10m of the front line were considered participants in the encounter. Frontlines could shift during encounters, particularly when one group decided to retreat (Figure 4.2). The first observer continuously monitored the frontline and also operated the playback speaker.

The second observer selected and monitored the subject provided it was not participating in the encounter (i.e., located at least 30m from the front line, not showing any interest in the encounter but engaged in foraging, resting or grooming). The subject was continuously filmed using a JVC quadproof EverioR camcorder. Both observers communicated with each other via Motorola radios and, once all conditions were met, the first observer played a 'leopard' alarm from the speaker near the front line. The second observer video-recorded the subject's behaviours for the next 10min from the onset of the playback.

4.4.4 Control trials

During control trials, the same individuals were tested, but this time in the absence of intergroup encounter when most group members were either foraging, resting or grooming. We ensured that no intergroup encounter and no natural alarm calling event occurred for at least 2h before a trial. As with experimental trials, we only conducted control trials in mixed habitat and with a group spread of about 100m. The second observer located the subject and followed it for at least 10min. If no relevant event occurred during this time period (participation in intra-group aggression), the second observer informed the first observer to start the trial. Similar to experimental trials, the first observer then positioned herself between 30-50 m from the subject hidden behind a bush to broadcast the playback stimulus upon the instructions from the second observer on the radio. The second observer video recorded the subject continuously for 10min from the onset of the playback call.

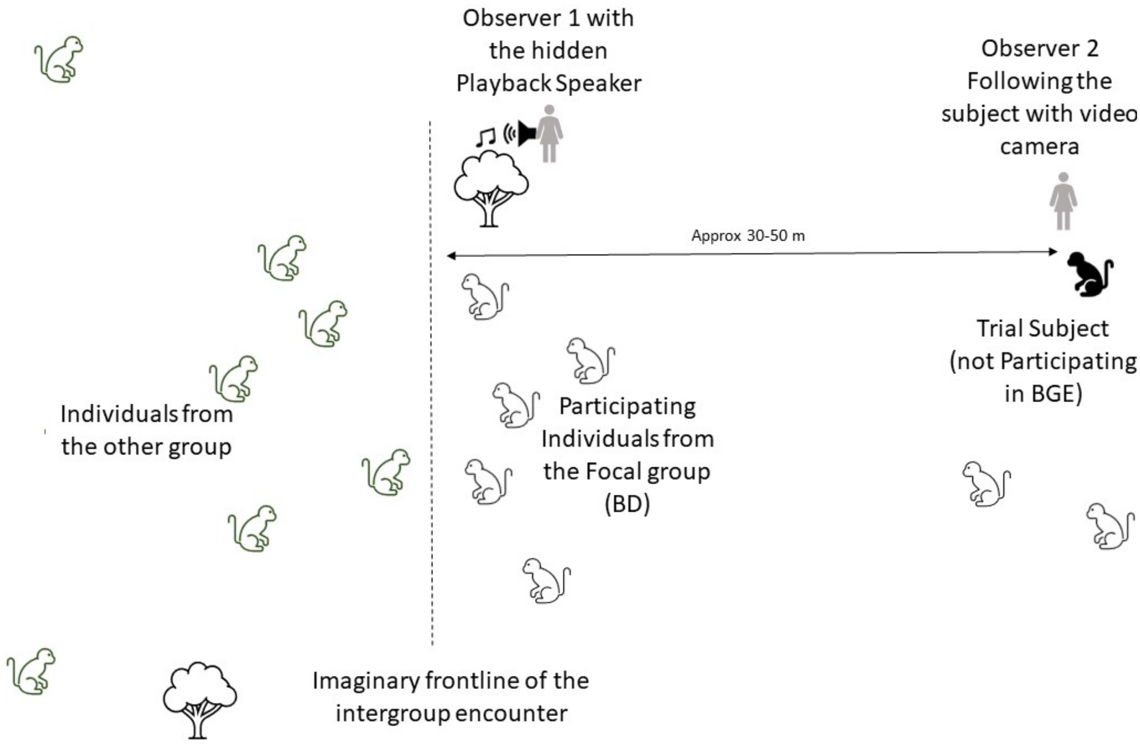

Figure 4.2: Heuristic illustrations (not to scale) showing an example of the experimental trials

4.4.5 Behavioural responses

For both experimental and control trials, the second observer continuously scored the overall group activity, the number of individuals present within 5m of the subject before each trial (audience size), and any other significant behaviours by the group members.

We coded key behaviours during the first minute following each playback trial using BORIS V. 7.10.2 coding software [43]. We were specifically interested in the following behaviours by the subject: (1) Looking towards the speaker (duration of looking in the direction of the speaker (s), as inferred from head orientation); (2) vigilance (duration of scanning environment with alert body posture (s)); (3) startle (sudden complete or upper body jerk movement, with a rapid head movement in any direction (yes/no)) [33]; (4) predator-specific responses (running into the nearest tree or bush (yes/no) [44,45].

To assess inter-observer reliability, 50% of trial videos were randomly selected and recoded by a second coder blind to the hypotheses and objectives but familiar with primate behaviour. We then compared the key behaviours described above between the two observers. As the speaker direction was not obvious in the video for the naïve coder, he was intimated about the general direction of the speaker at the start of the video. We calculated Pearson's correlations coefficients between two coders for the continuous variables (Looking towards the speaker: $r = 0.78$, $p = 0.02$; Vigilance: $r=0.88$, $p=0.003$) and Cohen's kappa was calculated for the startle responses ($k=0.91$). Thus, the inter-observer agreement was sufficiently high for all three key behaviours.

4.4.6 Statistical analyses

All statistical analyses were conducted in R v4.0.3 [46] using the lme4 [47] package for linear mixed effect models (LME). We log-transformed the response variables 'looking towards the speaker' and 'vigilance' to fit the assumptions of normal distribution. We created full models for both response variables by adding trial type (experimental vs control trials), recording context (intergroup vs predator) of playback 'leopard' alarm, audience size, and trial order as fixed effects. To control for repeated testing, we added subject identity as a random effect in the full models.

For model selection, we used an information-theoretic approach. We created two sets of candidate models corresponding to the two full models. Candidate models for each set were created by removing one fixed effect at a time from the full models, finally achieving null models without any fixed effect. We then compared the Akaike information criterion (AIC) values of all the models in the given set. We retained a final model from each set where AIC values increased by at least 2.0. In a post-hoc analysis, we checked Akaike's information criterion corrected for small sample sizes (AICc) values and AICc weights of the two selected final models compared to all other candidate models within the sets using the 'model.sel' function in MuMIn package [48]. The comparison revealed that both the final models had minimum AICc values and the highest AICc weights in the corresponding set (Appendix).

We could not use the same approach to assess the effect of treatments on the binary variables, startle and appropriate anti-predator reaction. Startle responses were never observed in the control trials and in only half of the experimental trials. Zero-inflation of startle responses combined with a small sample size resulted in non-converging LME. Hence, we tested the effect of trial type using the Barnard test (CSM approximate exact test) on the startle responses. Predator specific responses never occurred, neither in the experimental nor in the control trials.

4.5 Results

We conducted N=8 control trials and N=8 experimental trials with the BD group (N=5 with AK as an opposing group, N=3 trials with CR as an opposing group). Subjects looked toward speaker for significantly less time in the experimental compared to the control trials (LME: Estimates (log-transformed) = - 0.63 ±0.24, p= 0.008, ΔAIC= - 2.48; Figure 4.3). They also showed lower vigilance in experimental compared to control trials (LME: Estimates (log-transformed) = 0.50 ±0.18, p= 0.006, ΔAIC= - 2.15; Figure 4.4) and were less likely to show startle responses in experimental compared to control trials (Barnard Test = 0.015 p= 0.006). Additionally, during N=2 experimental trials, the adult male from the other group (AK) responded with 'leopard' alarms after hearing the playback of the calls.

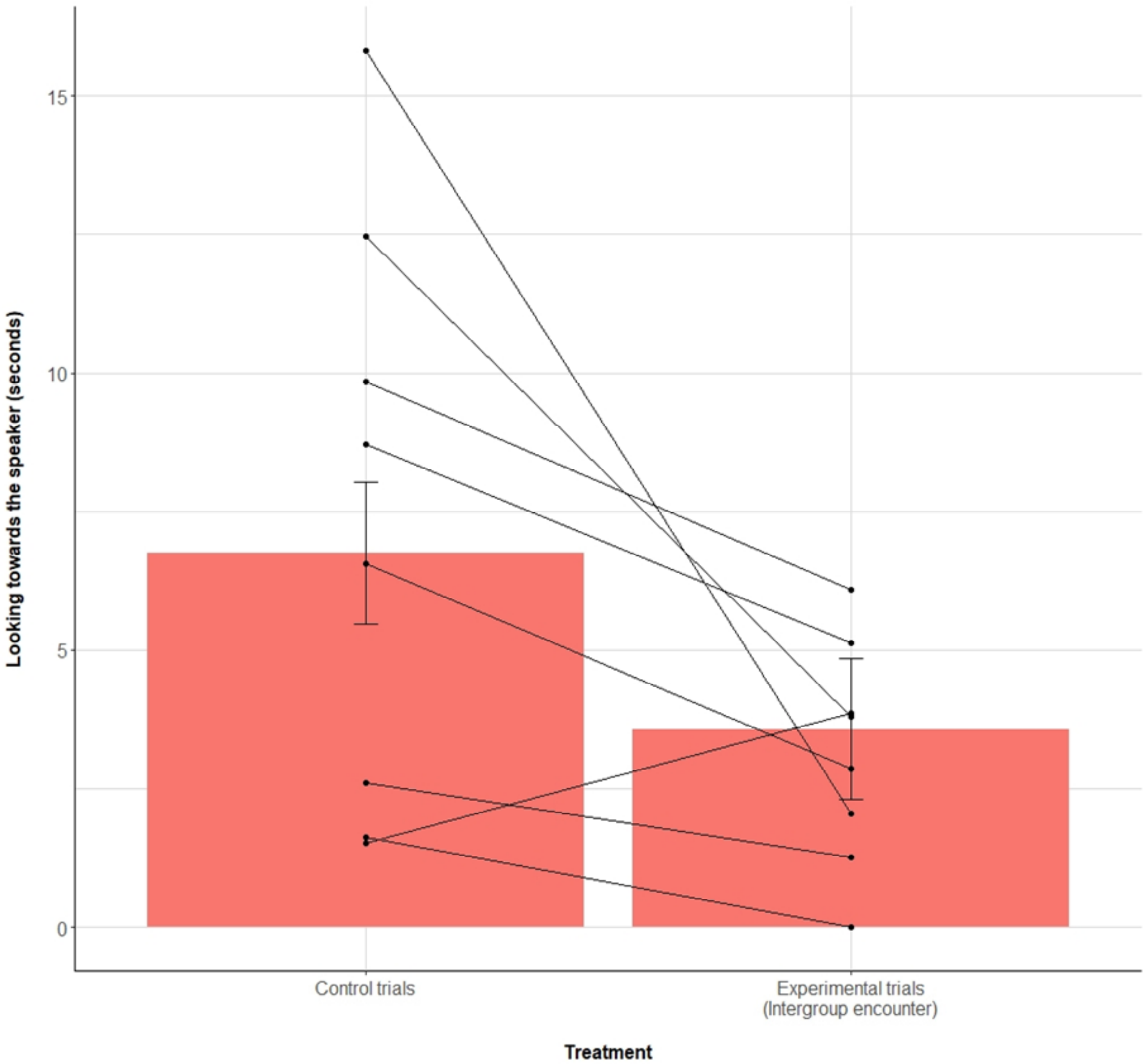

Figure 4.3: Duration of looking towards the speaker after leopard alarm playbacks in the control and experimental trials. Bars represent model estimates (±SE). Points connected by lines represent raw data points for each subject (n=8)

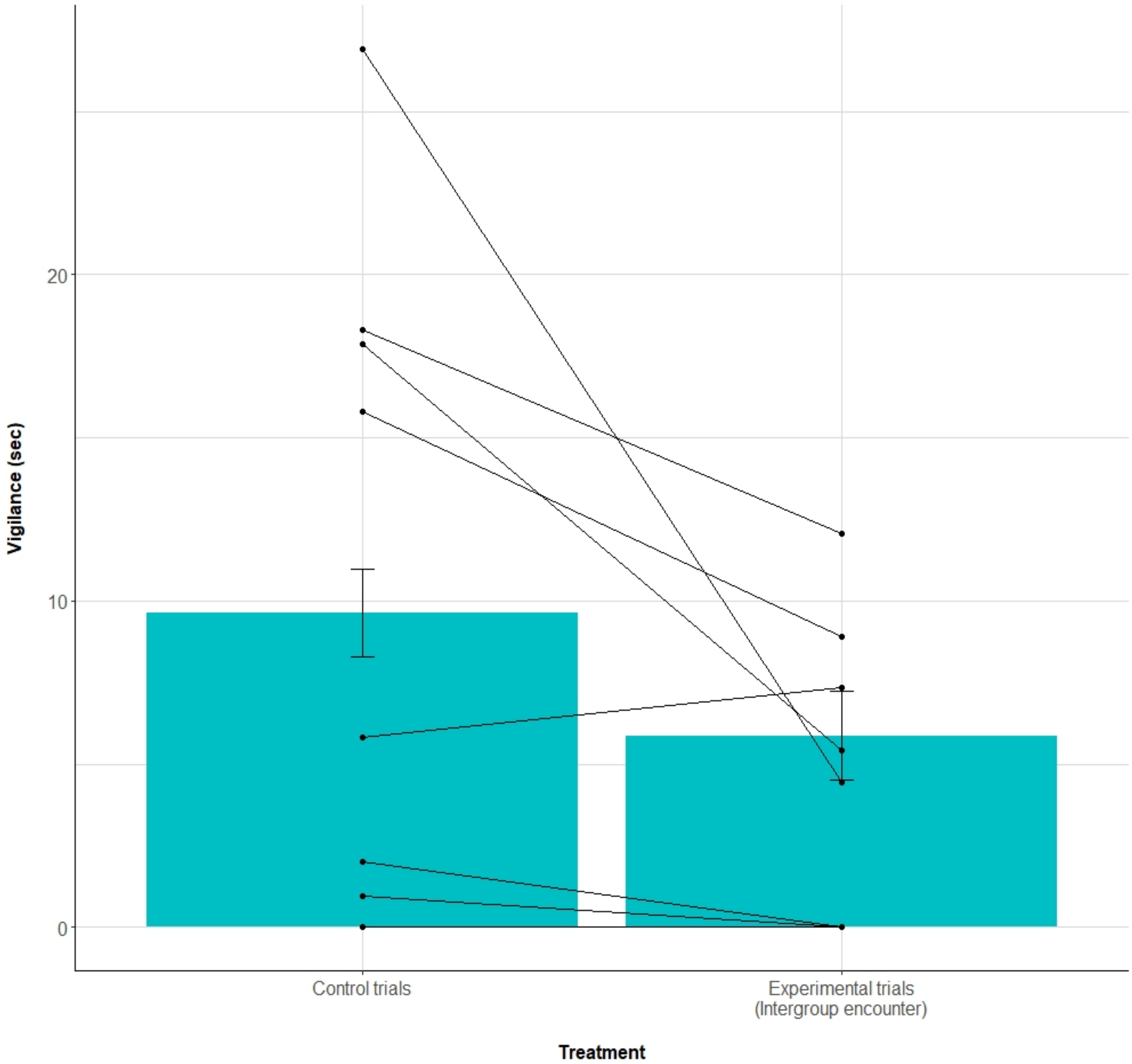

Figure 4.4: Vigilance duration after leopard alarm playbacks in the control and experimental trials. Bars represent model estimates (±SE). Points connected by lines represent raw data points for each subject (n=8)

4.6 Discussion

Most animal vocalisations are functionally specific, in the sense that the different call types within a species' repertoire are produced to relatively specific events, which effectively render them into referential signals that recipients can use to predict the associated events [49]. However, at least in primate communication, some call types are produced in a range of situations, which effectively prevents recipients from forming strong stimulus-response associations and responding directly to the call. In such instances, recipients are forced to use additional information, along with the actual vocal signal, to infer the nature of the associated event, i.e., the meaning of the call [27], similar to linguistic pragmatics.

In this study, we investigated the hypothesis that vervet monkeys relied on contextual information to infer the meaning of one of their calls, the male 'leopard' alarms, which are usually given to terrestrial predators, such as jackals, but also during non-predatory intergroup aggression. We were able to show experimentally that the recipient reacted differently to the same call exemplar, depending on the contexts in which it was produced. When 'leopard' alarms were produced during ongoing intergroup encounters, non-participating subjects were less interested in the calls, i.e., they looked at the speaker for less time, were less vigilant and never showed startle responses, compared to when the same calls were produced during foraging or resting. We concluded that if recipients already had an explanation for why a male 'leopard' alarm was produced (in this case due to the intergroup encounter), they were not specifically affected by the call, in contrast to neutral situations where the alarm call was inconsistent with the ongoing context and, in all likelihood, predicted that the male unexpectedly spotted a terrestrial predator. Naturally, it would have been desirable to add another condition to the experimental design, i.e., playing back male 'leopard' alarms in the presence of a terrestrial predator. Unfortunately, natural predator encounters are much too rare for any systematic study, although it might be possible to simulate predator presence with a suitable model using taxidermy. Also, we conducted all the control trials after the experimental trials because of the unpredictability of participation in intergroup encounters, identification of subjects was not possible before the experimental trial on any subject. It is not an optimal experimental design. Nevertheless, as we were expecting to heighten behavioural responses in the control trials, any effects of the habituation to the playback calls because of the previous experimental trials can be ruled out.

We selected the variable 'looking towards the speaker' because it was most likely an indication that the subject was looking for information about the reason for the call [2,44]. The fact that subjects looked for less time towards the speaker when trials were administered during intergroup interactions suggested that they already made assumptions regarding the cause of the call. During experimental trials, subjects already had time to learn about the intergroup context, either due to other monkeys' vocalisations typical for this context or due to having spotted members of the neighbouring group. Whatever source of information was available to them, the uncertainty about the cause of the (playback) call was low, allowing subjects to associate the call as a part of an ongoing intergroup interaction. In contrast, when the same calls were played during control trials, subjects had no relevant prior information available about the call's cause, which would explain their increased efforts in information seeking.

Another argument for additional information seeking in control trials could be a higher rate of false alarms in large groups, which could drive individuals to seek more information before responding, probably to avoid energetic costs [50]. Alternatively, additional information seeking in a neutral context can be used as a strategic response to counter the deceptive alarm calls [13,31]. However, systematic exploration of the presence and causes of deceptive alarm calls has not been conducted for south African vervet monkeys.

Vigilance was significantly less during experimental trials than during control trials. As the subjects, per design, did not participate in the intergroup encounter and stayed away from the frontline, a reasonable assumption for them was that the call was directed at a neighbouring individual rather than at a predator or at themselves. In control trials, however, no further information was available, which increased the uncertainty about the cause of the call and might have resulted in increased vigilance. Here, it is important to mention a relevant natural observation of aggressive encounter between two other groups in this population made outside this study. On 8th February 2018, the NH and AK groups had an aggressive encounter and the adult males and other individuals from both groups produced 'leopard' alarms on the frontline. However, during the encounter, a jackal (*Lupulella mesomelas*) was spotted by the observer (AD), stalking a non-participating individual (AD, personal obs.). Several frontline individuals also saw the jackal and produced alarm calls to it, but the non-participating individuals did not react with anti-predator behaviour or vigilance to the alarm calls (as predicted for this context and as documented in the experiment). In this case, the jackal's hunting attempt was unsuccessful as, at the last moment, the targeted monkey saw the jackal and escaped onto a tree. However, the observation highlights a rare situation in which listeners are likely to draw wrong conclusions, which could be costly. It also highlights the fact that 'leopard' alarm calls do not contain subtle acoustic cues that separate intergroup from predation contexts. Ambiguous contexts, like the one described here, are probably very rare, such that pragmatic inference remains an adaptive strategy.

Finally, startle responses were completely absent during experimental trials but were observed in half of the control trials. This result is in line with the interpretation that monkeys assumed the presence of a predator in the absence of a disambiguating intergroup context. In other words, the most plausible interpretation is that subjects, devoid of any contextual information, were more affected by an unanticipated male 'leopard' alarm than in the intergroup context where such calls are common. The fact that only some subjects responded this way might be linked to individual differences in age, sex, or social position. As we could not conduct proper statistical analyses for this variable, our results remain suggestive.

It is essential to note that we never saw predator-specific escape responses, as sometimes seen in response to terrestrial predators. For terrestrial carnivores, the most adaptive anti-predator response is to climb rapidly into the nearest tree or bush [44]. Nevertheless, such rapid flight responses have been observed in the study groups, but they typically occur when individuals forage close to each other (AD unpublished data). Thus, these textbook-style antipredator responses may occur primarily under specific circumstances when social contagion overrides individual decision-making [51,52]. Similar observations about the general lack of appropriate predator-specific responses and its occurrence due to social contagion were made in a recent study on the different population of the same species [33]. These observations reiterate that responses to alarm calls are not stereotypical. Alternatively, the Amboseli population in Kenya, on which the original study was conducted, regularly encountered and predated by leopards compared to the population in this study (Robert Seyfarth personal communication). Such differences in predatory pressure among populations could also explain the differences in predator-specific responses to leopard alarm calls.

How general are the vervet monkeys' documented pragmatic abilities, as shown by this experiment? Recent experiments on closely related green monkeys (*Chlorocebus sabaeus*) with a similar alarm call system are relevant here. During these experiments, researchers primed subjects one hour before the playbacks, which did not have a lasting impact on their subsequent behaviour. In our study, calling context was an ongoing event that may represent a more natural case and explain the diverging results, despite the fact that contextual information could impact long-term responses, albeit weakly [53]. Another difference between our study and the green monkey experiment relates to the nature of the playback stimulus (male 'leopard' alarm barks vs female 'leopard' alarm chirps, respectively). It is possible that male and female calls evoke different responses or serve different functions [54]. For example, male guenon alarm calls have a documented dual function as both anti-predator and sexual competition signals, suggesting that female alarms may be more effective in disambiguating context than male alarm calls [55]. The recent acoustic analyses of vervet alarm calls of study and other populations suggest that although male leopard alarm calls elicited in different contexts are acoustically similar, the acoustic overlap is not complete [11,40]. A possibility that unidentified subtle acoustic differences provide cues about the context in which calls are elicited cannot be entirely ruled out.

Furthermore, the function of male 'leopard' alarm in South African vervet monkeys has not yet been studied systematically. However, the playback call during two experimental trials triggered 'leopard' alarms from a male of the opposing group, pointing towards a possibility that male 'leopard' alarms also have a dual function in communicating to both predators and conspecific competitors, as noted by previous studies [11,56]. One possible approach could be to use a similar experimental design and record the responses of participating members of intergroup encounters and examine if male 'leopard' alarms during these events affect the overall outcome (winning and losing) of an encounter. We also suspect that within-context responses to male 'leopard' alarms vary between the individuals. Future studies that investigate inter-individual differences in the responses to alarm calls are necessary. More importantly, comparable experiments are needed to examine if the pragmatic inference is ubiquitous to all call types in this species.

Whatever the function of 'leopard' calls might serve, the important point remains that male 'leopard' alarm in all the contexts reliably predicted the presence of potential threats (although of different types) to receivers. As argued by Scarantino, 'leopard' alarms may contain natural information about the presence of threat as smoke carries information about a fire, Grice's notion of 'natural meaning' [57,58]. And contextual information facilitates locating and specifying the source of the threat, allowing individuals to respond appropriately. However, this speculation remains open to be tested in future studies.

4.7 Conclusion

We provide empirical evidence that vervet monkey responses to male 'leopard' alarms depend on the contexts in which calls are perceived, suggesting a basic pragmatic ability in non-human primate cognition and communication. Thus, it is possible that with the evolution of more advanced cognition, simple contextual pragmatics seen in non-human primates gradually became more important and included psychological variables, such as speaker intention and common ground, to evolve into the sophisticated, pragmatic abilities currently seen in human communication [21].

4.8 Acknowledgements

AD would like to thank Luke Quarles for volunteering as a second coder and Radu Slobodeanu for the help with statistical analyses. We would like to thank all the volunteers, researchers, and managers at the Inkawu Vervet Project for their help during data collection.

4.9 Author contribution

AD conceived the study with help from KZ. AD conducted the experiments, analyses and wrote the first draft, which KZ reviewed. AD and KZ wrote the final version of the manuscript.

4.10 References

1. Darwin C. 1872 *The expression of the emotions in man and animals*. London: J.Murray. (doi:10.1038/036294c0)
2. Struhsaker TT. 1967 Auditory communication among vervet monkeys (Cercopithecus aethiops). In *Social communication among primates*, pp. 281–324.
3. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Anim. Behav.* **28**, 1070–1094. (doi:10.1016/S0003-3472(80)80097-2)
4. Macedonia JM, Evans CS. 1993 Essay on Contemporary Issues in Ethology: Variation among Mammalian Alarm Call Systems and the Problem of Meaning in Animal Signals. *Ethology* **93**, 177–197. (doi:10.1111/j.1439-0310.1993.tb00988.x)
5. Zuberbühler K. 2000 Referential labelling in Diana monkeys. *Anim. Behav.* **59**, 917–927. (doi:10.1006/anbe.1999.1317)
6. Pereira ME, Macedonia JM. 1991 Ringtailed lemur anti-predator calls denote predator class, not response urgency. *Anim. Behav.* **41**, 543–544. (doi:10.1016/S0003-3472(05)80861-9)
7. Zuberbühler K. 2001 Predator-specific alarm calls in Campbell's monkeys, Cercopithecus campbelli. *Behav. Ecol. Sociobiol.* **50**, 414–422. (doi:10.1007/s002650100383)
8. Manser MB. 2001 The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proc. R. Soc. B Biol. Sci.* **268**, 2315–2324. (doi:10.1098/rspb.2001.1773)
9. Kirchhof J, Hammerschmidt K. 2006 Functionally referential alarm calls in tamarins (Saguinus fuscicollis and Saguinus mystax) - Evidence from playback experiments. *Ethology* **112**, 346–354. (doi:10.1111/j.1439-0310.2006.01165.x)
10. Cunningham S, Magrath RD. 2017 Functionally referential alarm calls in noisy miners communicate about predator behaviour. *Anim. Behav.* **129**, 171–179. (doi:10.1016/j.anbehav.2017.05.021)
11. Price T, Wadewitz P, Cheney DL, Seyfarth RM, Hammerschmidt K, Fischer J. 2015 Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* **5**, 1–11. (doi:10.1038/srep13220)
12. Gyger M, Marler P. 1988 Food calling in the domestic fowl, Gallus gallus: the role of external referents and deception. *Anim. Behav.* **36**, 358–365. (doi:10.1016/S0003-3472(88)80006-X)
13. Wheeler BC. 2009 Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proc. R. Soc. B Biol. Sci.* **276**, 3013–3018. (doi:10.1098/rspb.2009.0544)
14. Rendall D, Owren MJ, Ryan MJ. 2009 What do animal signals mean? *Anim. Behav.* **78**, 233–240. (doi:10.1016/j.anbehav.2009.06.007)
15. Wheeler BC, Fischer J. 2012 Functionally referential signals: A promising paradigm whose time has passed. *Evol. Anthropol.* **21**, 195–205. (doi:10.1002/evan.21319)
16. Scott-Phillips TC. 2010 Animal communication: insights from linguistic pragmatics. *Anim. Behav.* **79**, e1–e4. (doi:10.1016/j.anbehav.2009.10.013)

17. Carnap VR. 1958 BEOBACHTUNGSSPRACHE UND THEORETISCHE SPRACHE. *dialectica* **12**, 236–248. (doi:10.1111/j.1746-8361.1958.tb01461.x)

18. Wharton T. 2009 *Pragmatics and non-verbal communication*. Cambridge University Press.

19. Baratgin J, Dubois-Sage M, Jacquet B, Stilgenbauer JL, Jamet F. 2020 Pragmatics in the False-Belief Task: Let the Robot Ask the Question! *Front. Psychol.* **11**, 1–17. (doi:10.3389/fpsyg.2020.593807)

20. Sperber D, Wilson D. 1995 *Relevance: Communication and Cognition*. Oxford: Blackwell Publishing Ltd. (doi:10.1111/j.1468-0017.1989.tb00246.x)

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

22. Scott-Phillips TC. 2016 Meaning in great ape communication: summarising the debate. *Anim. Cogn.* **19**, 233–238. (doi:10.1007/s10071-015-0936-3)

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

24. Krupenye C, Kano F, Hirata S, Call J, Tomasello M. 2016 Great apes anticipate that other individuals will act according to false beliefs. *Science (80-.).* **354**, 110–114. (doi:10.1126/science.aaf8110)

25. Crockford C, Wittig RM, Mundry R, Zuberbühler K. 2012 Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* **22**, 142–146. (doi:10.1016/j.cub.2011.11.053)

26. Heesen R, Bangerter A, Zuberbühler K, Rossano F, Iglesias K, Guéry JP, Genty E. 2020 Bonobos engage in joint commitment. *Sci. Adv.* **6**. (doi:10.1126/SCIADV.ABD1306)

27. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)

28. Arnold K, Zuberbühler K. 2013 Female Putty-Nosed Monkeys Use Experimentally Altered Contextual Information to Disambiguate the Cause of Male Alarm Calls. *PLoS One* **8**, 2–8. (doi:10.1371/journal.pone.0065660)

29. Zuberbühler K. 2000 Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* **76**, 195–207. (doi:10.1016/S0010-0277(00)00079-2)

30. Zuberbühler K. 2000 Causal knowledge of predators' behaviour in wild Diana monkeys. *Anim. Behav.* **59**, 209–220. (doi:10.1006/anbe.1999.1296)

31. Wheeler BC, Hammerschmidt K. 2013 Proximate factors underpinning receiver responses to deceptive false alarm calls in wild tufted capuchin monkeys: Is it counterdeception? *Am. J. Primatol.* **75**, 715–725. (doi:10.1002/ajp.22097)

32. Rendall D, Seyfarth RM, Cheney DL, Owren MJ. 1999 The meaning and function of grunt variants in baboons. *Anim. Behav.* **57**, 583–592. (doi:10.1006/anbe.1998.1031)

33. Ducheminsky N, Henzi SP, Barrett L. 2014 Responses of vervet monkeys in large troops to terrestrial and aerial predator alarm calls. *Behav. Ecol.* **25**, 1474–1484. (doi:10.1093/beheco/aru151)

34. Keenan S, Lemasson A, Zuberbühler K. 2013 Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. *Anim. Behav.* **85**, 109–118. (doi:10.1016/j.anbehav.2012.10.014)

35. Hammerschmidt K, Fischer J. 1998 The vocal repertoire of Barbary macaques: A quantitative analysis of a graded signal system. *Ethology* **104**, 203–216. (doi:10.1111/j.1439-0310.1998.tb00063.x)

36. Willems EP, Arseneau TJM, Schleuning X, van Schaik CP. 2015 Communal range defence in primates as a public goods dilemma. *Philos. Trans. R. Soc. B Biol. Sci.* **370**. (doi:10.1098/rstb.2015.0003)

37. Zuberbühler K, Jenny D, Bshary R. 1999 The predator deterrence function of primate alarm calls. *Ethology* **105**, 477–490. (doi:10.1046/j.1439-0310.1999.00396.x)

38. Arseneau-Robar TJM, Taucher AL, Müller E, Van Schaik CP, Bshary R, Willems EP. 2016 Female monkeys use both the carrot and the stick to promote male participation in intergroup fights. *Proc. R. Soc. London B Biol. Sci.* **283**. (doi:10.1098/rspb.2016.1817)

39. Arseneau-Robar TJM, Müller E, Taucher AL, Van Schaik CP, Bshary R, Willems EP. 2018 Male monkeys use punishment and coercion to de-escalate costly intergroup fights. *Proc. R. Soc. B Biol. Sci.* **285**. (doi:10.1098/rspb.2017.2323)

40. Besson E. 2017 Vervet monkeys alarm calls: context specific or not? University of Neuchatel. See https://www.unine.ch/files/live/sites/compcog/files/Thesis/170828_Msc_Thesis_BESSON_Emmanuelle.pdf.

41. Center for Conservation Bioacoustics. 2014 Raven Pro: Interactive Sound Analysis Software.

42. Cheney DL, Seyfarth RM. 1982 How vervet monkeys perceive their grunts: Field playback experiments. *Anim. Behav.* **30**, 739–751. (doi:10.1016/S0003-3472(82)80146-2)

43. Friard O, Gamba M. 2016 BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol.* **7**, 1325–1330. (doi:10.1111/2041-210X.12584)

44. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey (Cercopithecus aethiops) alarm calls: semantic communication in a free ranging primate. *Anim. Behav.* **28**, 1070–1094.

45. Seyfarth RM, Cheney DL, Marler P. 1980 Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801–803. (doi:10.1126/science.7433999)

46. Team RDC. 2020 R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna, Austria. ISBN 3-900051-07-0*.

47. Bates D, Maechler M, Bolker B, Walker S. 2015 Package lme4. *J. Stat. Softw.* **67**, 1–91. (doi:http://lme4.r-forge.r-project.org)

48. Barton K. 2020 MuMIn: Multi-model inference.

49. Smith J, Harper D. 2003 *Animal signals*. Oxford University Press.

50. Beauchamp G, Ruxton GD. 2007 False alarms and the evolution of antipredator vigilance. *Anim. Behav.* **74**, 1199–1206. (doi:10.1016/j.anbehav.2007.02.014)

51. Hoppitt W, BlackBurn L, Laland KN. 2007 Response facilitation in the domestic fowl. *Anim. Behav.* **73**, 229–238. (doi:10.1016/j.anbehav.2006.05.013)

52. Armstrong EA. 1951 The nature and function of animal mimesis. *Bull. Anim. Behav.* **9**, 46–48.

53. Price T, Fischer J. 2014 Meaning attribution in the West African green monkey: Influence of call type and context. *Anim. Cogn.* **17**, 277–286. (doi:10.1007/s10071-013-0660-9)

54. Mehon FG, Stephan C. 2021 Female putty-nosed monkeys (Cercopithecus nictitans) vocally recruit males for predator defence.

55. Zuberbühler K. 2005 Effects of Natural and Sexual Selection on the Evolution of Guenon Loud Calls. In *The Guenons: Diversity and Adaptation in African Monkeys*, pp. 289–306. Kluwer Academic Publishers. (doi:10.1007/0-306-48417-x_20)

56. Cheney DL, Seyfarth RM. 1992 *How monkeys see the world: Inside the mind of another species*. The University of Chicago Press.

57. Scarantino A. 2010 Animal communication between information and influence. *Anim. Behav.* **79**, 1–5. (doi:10.1016/j.anbehav.2010.03.005)

58. Grice H. 1957 Meaning. *Philos. Rev.*

4.11 Chapter 4: Appendix

Table A4.1: A table showing the set of candidate models for the response variable of looking towards the speaker (log-transformed) (LLS). Treatment= Type of treatment; Playback= Type of playback call; Audience.5m= number of group members within 5m of a subject during the playback; Order= Order of the trial. The identity of the subject is added as a random effect in each model. The final model is highlighted in bold where AIC values increased by more than 2.0

Model	Degrees of freedom	Akaike information criterion (AIC)	Δ AIC
LLS ~ Treatment +Playback + Audience.5m. + Order + (1|Subject)	7	46.5	0.00
LLS ~ Treatment +Playback + Order +(1|Subject)	6	43.14	-3.36
LLS ~ Treatment + Playback + (1|Subject)	5	39.25	-3.89
LLS ~ Treatment + (1|Subject)	**4**	**39.37**	**0.12**
LLS ~ 1 + (1|Subject)	3	41.85	+2.48

Table A4.2: A table showing the set of candidate models for the response variable of vigilance(log-transformed) (LVIG). Treatment= Type of treatment; Playback= Type of playback call; Audience.5m= number of group members within 5m of a subject during the playback; Order= Order of the trial. The identity of the subject is added as a random effect in each model. The final model is highlighted in bold where AIC values increased by more than 2.0

Model	Degrees of freedom	Akaike information criterion (AIC)	Δ AIC
LVIG ~ Treatment +Playback + Audience.5m. + Order + (1|Subject)	7	49.29	0.00
LVIG ~ Treatment +Playback + Order +(1|Subject)	6	44.46	-4.84
LVIG ~ Treatment + Playback + (1|Subject)	5	39.8	-4.66
LVIG ~ Treatment + (1|Subject)	**4**	**39.62**	**-0.18**
LVIG ~ 1 + (1|Subject)	3	41.77	+2.15

Chapter 5: General discussion

5.1 Summary of findings

Chapter 2 focuses on the role of social learning in the vocal communication of non-human primates. Human language is mainly a socially learned behaviour enabled by a specialised cognitive architecture. Very few studies have tested if social learning is also a relevant mechanism through which non-human primates acquire competence in vocal communication. In this experiment, I provided monkeys with novel visual stimuli (unfamiliar animal models) and simultaneously broadcasted alarm calls. The subject showed increased vigilance towards the model when presented with the same model again without playback of alarm calls. These results can be considered as one-trial social learning of 'meaning', i.e., call-context associations. Such rapid learning is akin to the fast mapping seen in human children [1] and also documented in a dog [2]. Thus, these results provide experimental evidence that conspecific alarm calls facilitate social learning in wild primates and confirm observations that vervet monkeys socially acquire some aspects of anti-predatory behaviour [3]. Furthermore, some of the juvenile subjects also produced alarm calls similar to the calls they heard during first exposure, suggesting call comprehension leads to call usage [4]. These conclusions were based on a small sample size, and the study would have benefitted from additional control treatments. However, this experiment is one of the few that investigated social learning in the vocal communication of wild non-human primates, a promising approach to determine the flexibility in vocal communication systems. If individuals can socially learn to modify their behaviour, we can infer that the behaviour in question is not entirely innate and reflexive. Future studies should extend the social learning paradigm to assess the degree of flexibility in other call types. Another interesting finding from this experiment was the interaction between core and socially acquired knowledge [5,6]. Although individuals associated the alarm calls with the novel predator models, they generally reacted less to the less animate model, indicating how core knowledge about animacy combines with knowledge acquired through social learning [7,8].

Chapter 3 focuses on the functional flexibility of calls. Functional flexibility is the ability to produce the same sound in different contexts, irrespective of emotional valence [9]. It has been argued that functional flexibility is a fundamental infrastructural property of human speech, as it allows the production of phonemes irrespective of a speaker's emotions. In non-human primates, contextually flexible calls appear to be the exception, although they have been documented in some species [10,11]. However, it can be argued that different contexts might evoke the same emotional state in callers, resulting in the production of the same call across contexts. One way to rule out this possibility is to examine if non-human primates can produce an existing call in an entirely novel context. To address this, I designed an experimental setup in which wild vervet monkeys could socially learn to produce move-grunts to obtain food reward from a manual food dispenser. In the first condition, I provided the learning stimulus in the auditory modality only, i.e., by playing back move-grunts as a predictor of a food reward. No subject was able to learn the task in this experiment, i.e., to produce move grunts themselves to generate food. Here, observations suggested that subjects merely associated the playback with the food rewards and simply waited near the dispenser in apparent anticipation that the event would take place. In the second condition, the learning stimulus was in the form of a complete sound/video demonstration of conspecific successfully obtaining a food reward after producing a move-grunt. In this setup, out of 39 subjects, one juvenile female acquired the new skill, in all likelihood, through watching the demonstration video. Remarkably, she also failed to produce fully formed move-grunts, which we discuss in relation to recent neurobiological findings on cognitive control on vocalisations. Overall, our results and behavioural observations are in line with the previous studies and suggest that non-human primates can socially learn novel call usage, provided they obtain appropriate social input. However, the basic theory remains unchanged, insofar as non-human primates only have very limited control over call production due to a lack of an adequate underlying neural machinery, making it near impossible for them to use calls in functionally different contexts. Nevertheless, the study opens a new experimental paradigm that could be used to investigate other species. More importantly, future neurological studies in captive settings may want to consider social learning paradigms similar to this experiment when training subjects to produce vocalisations on command. I have also suggested potential improvements in the experimental setup that can be implemented by future studies.

Chapter 4 focuses not on how the meaning of call emerges from past social experiences but on how individuals use ongoing social context to interpret call meaning. In human language, the 'intended' meaning of utterances can be independent, or even contradictory, to the 'literal' meaning, as, for example, in sarcasm. However, humans can rapidly and reliably distinguish literal from intended (or pragmatic) meaning using sophisticated psychological abilities. Some researchers even argue that most human communication is based on pragmatics rather than semantics [12]. Interestingly, pragmatic interpretation of utterances is one of the hardest problems for automated speech recognition systems, suggesting that it builds on a psychological apparatus that requires more than just rich and diverse language input. A review of animal communication literature suggests that non-human primates also use an arguably simpler form of pragmatics to decipher the meaning of ambiguous calls [13,14]. However, this conclusion is still very preliminary as it rests on a small number of studies. Here, I tested the hypothesis that non-human primates can make pragmatic inferences to decipher the meaning of other individuals' alarm calls. In the experiment, I waited for natural inter-group encounters to occur before conducting a playback experiment and compared the monkeys' responses to the same stimulus in control contexts. I found that subjects reacted strongly in the control contexts and showed vigilance when hearing recordings of alarm calls, whereas this was not the case if alarm calls were presented during intergroup encounters. Overall, the results confirmed the use of contextual information to decipher ambiguous calls in non-human primates and suggested the possibility that, with the evolution of more advanced cognitive abilities, the simple contextual pragmatics already seen in non-human primates could turn into more sophisticated forms, as currently seen in human communication.

In conclusion, the overarching theme of the three studies concerns the amount of flexibility available to non-human primates in vocal communication, using wild vervet monkeys as a model system. Overall, the results provide a contribution to a growing body of evidence for vocal flexibility and weaken the argument for a more rigid affect-conditioning model in which callers are mere agents of their affective states [15]. The findings also contribute to questions about human language evolution by providing evidence of gradation in three language capacities: fast-mapping, functional flexibility and pragmatics. However, all three experiments also face limitations due to the nature of the experimental designs. In all experiments, subjects were chosen opportunistically and were not equally distributed across sex and age categories, which limits the interpretations. Also, due to complex experimental procedures, it was not possible to increase the sample size, which sometimes resulted in suboptimal statistical analyses. Despite these limitations, all the experiments may have the potential to serve as foundations for future studies that could replicate and improve upon these experimental designs. Finally, the results of the three experiments also highlight the need for a more nuanced view of non-human primate communication beyond simple dichotomies between hard-wired and cognitively complex. In the next section, I attempt to provide such a more nuanced view by proposing a classification scheme for different non-human primate vocalisations that can account for the growing body of often contradictory data interpretation between behaviourist and cognitivist viewpoints.

5.2 A new classification approach for primate vocalisations

The behaviourist view is represented generally by the affect-conditioning model [16]. It suggests in non-human primates, vocalisations are directly linked to the internal states of the callers and function to influence the behaviour of the recipient. Consequently, calls either evoke automatic responses in recipients, or responses are subject to Pavlovian conditioning. Overall, the system is hard-wired [15]. In a recent review, Fischer and Price [14] propose a similar view; however, they attribute more cognitive control over the receiver's responses. On the other extreme, cognitivist propose that non-human primates have cognitive control over the call production and reception and may even possess mental representations probably similar to what humans experience [17–19].

Interestingly, even though behaviourists and cognitivists have opposing viewpoints towards non-human primate vocalisations, one commonality is that both generally attempt to explain call repertoires of non-human primate species through one single mechanism or process [14,16,20]. However, none of the viewpoints can explain all the variation seen in the data, and there is no convincing reason to defend such a generalised view for the entire call repertoire. It is entirely possible that different call types evolved in response to very different selective pressures, resulting in different underlying mechanisms and developmental trajectories. For example, in humans, seemingly hard-wired vocalisations, such as crying or laughing, appear to have different underlying mechanisms than proto-phones [21,22]. Similarly, it has been shown that non-human primate gestures develop at different rates during ontogeny [23]. In line with this argument, a recent review also asks a pertinent question, 'must all signals be evolved?' and argues for a novel classification scheme for non-human primate signals (although this review mainly focuses on gestures) [24].

Evidence from neurobiological studies hints towards the possibility that different call types in primates may have different underlying mechanisms, for instance, distinct neural networks responsible for spontaneous and volitional calls [25,26]. Furthermore, different subregions of periaqueductal grey of non-human primate brains are responsible for eliciting specific call types, and lesions in these areas result in selective mutism [27,28]. (see [29] for a review of the neurobiology of different types of vocalisations).

Based on these indications, i.e., that different calls might have different underlying mechanisms, I suggest a new approach for the classification of the call repertoires that is based on differences in the control primates have in the production and reception of their vocalisations. I suggest three different call categories and provide examples to illustrate the concepts. Finally, I make testable predictions about how the call repertoire of different species are distributed between the three categories.

5.2.1 Category 1: Calls with minimal cognitive control

This category includes probably the most basic vocalisations found in mammals, requiring minimal cognitive control over usage and comprehension. The calls in this category are evolutionarily conserved. The production and reception mechanisms of the innate vocalisations can be best described by the affect-conditioning model [16]. These calls are coupled with the internal states of the caller, and the structure of the call evokes unconditioned responses from the receiver [15]. Individual identity can be encoded in this class of vocalisations, which might help the audience or bystanders to respond appropriately depending on their relationship with the caller [16]. Innate calls do not undergo any changes in terms of cognitive control over their onset during ontogeny. However, some motor control over the offset of the calls is achieved during ontogeny, enabling individuals to prolong the calls. Similarly, receivers also do not undergo any changes in their unconditioned responses during the ontogeny [15].

A prime example of such calls is screams, squeaks and distress calls found in most primate species, including humans [30,31]. These are high pitch and noisy vocalisations, often with a chaotic, nonlinear acoustic structure that probably functions to irritate or repel receivers [32,33]. Often, screams are used by victims to dissuade an attacker during aggression [31]. Simultaneously, these calls are hard to ignore and are produced more by young animals looking for social support [34–36]. Correlations between different attributes of callers with the acoustic structure of the screams have been documented in several primate species [31]. For example, in vervet monkeys, screams reflect the caller's identity, social role during interaction (aggressor/victim), and the intensity of interaction [37]. Some control over scream production is documented in chimpanzees [38,39] and humans, where screams can evoke specific emotional neural activity [40]. And finally, screams are evolutionarily conserved vocalisations as it has been shown that heterospecific screams evoke similar responses in different primate species [41–43].

5.2.2 Category 2: Calls with limited cognitive control

The second category includes calls that show limited cognitive control in both production and reception. For this category, the internal states of the callers might be linked to different call types. However, coupling with the affective state is weaker than in the previous category, allowing callers to inhibit production under unfavourable circumstances. These calls may develop with little environmental control from birth. However, usage is fine-tuned during the ontogeny through social learning and individual learning [4,44]. These calls are usually broadcasted and are not directed to a single receiver. Similarly, throughout ontogeny, receivers initially show inflexible responses to the calls but then gradually learn to make pragmatic inferences and form novel call-context associations with experience. Overall, this category fits Fischer and Price's characterisation of non-human primate vocalisations [14]

The most fitting example for this category is the non-human primate alarm calls. The alarm calls are loosely coupled with a predator class; however, the usage gets fine-tuned to specific predator species during ontogeny as per Cheney and Seyfarth's description development of alarm calls in vervet monkeys [3,45]. The limited control over call usage can also explain the combinations of alarm calls used by different simian species [46,47]. Also, the flexible call usage based on audience composition and predator types has been documented in several non-human primate species [48,49]. Finally, the pragmatic inference in alarm calls is evident in different primate species suggesting that alarm calls can be classified into this category [50–52].

5.2.3 Category 3: Calls with significant cognitive control

The production and reception of the calls in this category are relatively free from specific biological functions and involve a significant cognitive component. The influence of the affective state on the call production is thus correspondingly minimal, and calls are functionally flexible [9,10]. Callers have volitional control over both usage and comprehension, which require learning through different mutually non-exclusive processes like ontogenetic ritualisation, social learning or individual associative learning [53–55].

Non-human primate social calls, such as grunts and other greeting calls, would probably come under this category. Social calls generally facilitate social interactions between individuals [11,56]. Usage of the social call is strategic and is generally based on interactional history with the receivers [57], and behavioural patterns of greeting calls differ based on the species' social structure [58]. Thus, comprehension and usage of the social calls develop parallel with social cognition and knowledge. In line with this, preliminary reports suggest that correct usage of different grunts develops at a different speed in vervet monkeys [3]. Wild chimpanzees produce grunt types irrespective of emotional valences, at least during infancy, hinting towards the functional flexibility of these grunts [59]. Overall, usage and reception patterns of non-human primate social calls fit the criteria of this category.

The ideas presented through this classification scheme are not novel. Cheney and Seyfarth suggested to focus on social calls to look for potential precursors of language [13,60]. Gruber and Grandjean also suggested a multidimensional classification for non-human primate vocalisations [61]. Rather, I have only collated the different views together and briefly put forward the idea of classification based on differences in underlying cognitive control, which makes testable predictions about how the call repertoire of different non-human primate species is distributed across the three categories. Comparisons of call repertoires of different species using this framework could help us understand the potential social and ecological attributes of a species that would have shaped their call repertoire.

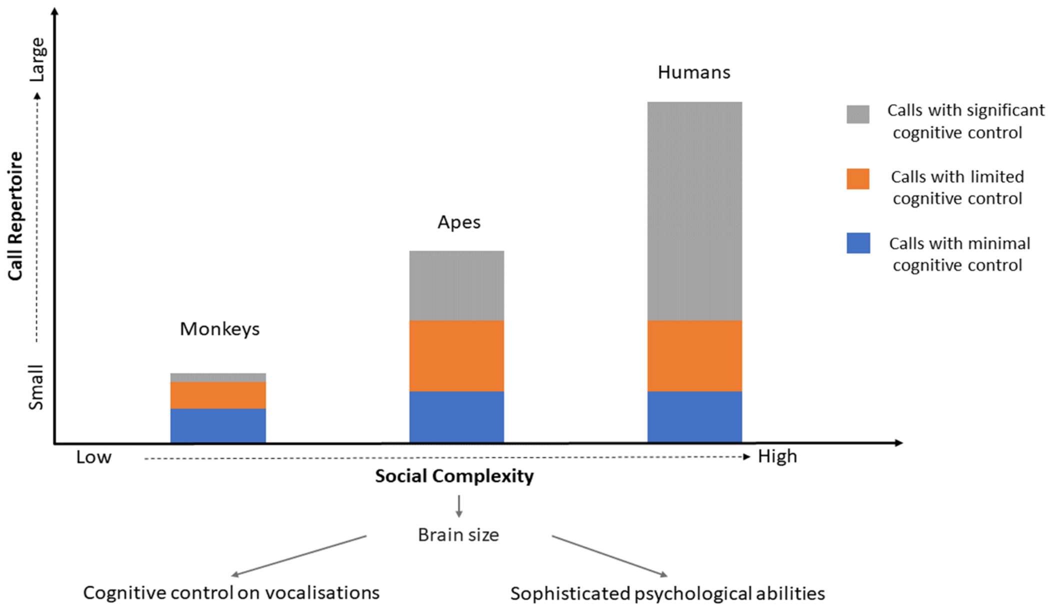

Figure 5.1: Hypothetical distributions of call repertoires of different primates based proposed classification of the calls. The call repertoire increases with the social complexity. Furthermore, social complexity also correlates with brain size. The increased brain size

For example, the classification predicts social calls would fall under the category of significant cognitive control, given the highly strategic usage of such calls. It is then reasonable to predict that species living in socially complex societies would have a greater number of calls in the third category. A similar view is also presented by Schamberg et al. in their framework. Indeed, McComb and Semple found strong relationships among the size of primate groups and the size of vocal repertoires [62], and there is now growing evidence that suggests parallels between social and vocal complexity [63,64] (but see [65] for a counter-argument). Furthermore, as suggested by the social brain hypothesis, both neocortex and relative brain size increase with the social complexity [66–68]. The increase in relative brain size and neocortex and the rapid evolution of the larynx in primates [69] could also result in higher cognitive control on vocalisations in socially complex species. With this logic, humans might have evolved an extensive call repertoire with a greater number of calls under significant cognitive control (Figure 5.1) [70,71]. This hypothetical scenario demonstrates how the classification of call types based on underlying cognitive control could help to trace the evolution of primate vocal communication systems. Future studies may find it helpful to catalogue call repertoires of different non-human primate species based on differences in presumed cognitive control Many of these predictions can be tested in the near future using advanced technologies, which I briefly review in the next section.

Future directions

In this final section, I briefly review possible future research directions on primate vocal behaviour, particularly how emerging technologies could revolutionise our understanding of primate and human communication.

One major inadequacy of many vocal communication studies is the coarse acoustic analyses. Traditionally classification of calls is done by choosing arbitrary acoustic parameters and extracting them manually. Although the traditional technique works for specific research questions and provides a broad classification of calls, it fails to detect the nuanced differences between call types and prone to human errors. Advances in machine learning (ML) algorithms can overcome such deficits [72]. Recent studies have shown that ML-based fine-grained call classification could highlight previously unseen differences in vocalisations [59,73,74]. In the future, ML-based acoustic analyses can support the playback experiments that can test the effects of subtle acoustic differences on the behaviour of receivers. However, ML algorithms are generally data-hungry [72] and require large datasets as training data for neural networks to accurately classify the species' repertoire. Only trained species-specific networks can be used for the classification of smaller datasets generally obtained through field experiments. Creating such large datasets of recorded natural vocalisations for any species requires serious time and efforts on the field. Furthermore, sorting and cleaning large acoustic datasets is an equally tedious job. However, the advanced lightweight audio recorders attached to individuals are becoming increasingly affordable and can significantly reduce manual efforts by continuously recording calls emitted by a free-ranging individual for days [75–78]. Finally, detecting, cutting, sorting, and feature extraction of call units from lengthy recordings can be fully or partially automated using ML techniques [79–84].

Additionally, along with microphones availability of cost-effective attachable dataloggers fitted with accelerometers and other sensors can simultaneously detect important physiological parameters during vocal communication events [85,86]. Such data from multiple channels can tremendously increase our understanding of the coupling between internal states and calls. One of the limitations of the field experiments (including the experiments in this thesis) is coarse behavioural measurements either in the field or through the coding of videos and the dependence on inter-observer reliability. In the future, multisensory calibrated bio-loggers can increase the accuracy of behavioural measurements in field experiments by many folds [87]. Furthermore, novel deep learning algorithms can also detect and track multiple individuals from videos and automatically code different body postures [88–90]. Such algorithms can provide information on how individuals react to the vocalisations at a high temporal resolution of almost 30 Hz [91]. Together these technologies can provide precise measurements of behaviours in the field as well as in captivity.

Another area in which technology has advanced in recent years is that of 3D animation techniques aided by deep learning algorithms. It is possible to create a customised, realistic-looking conspecific demonstrator using this technology for social learning experiments, such as the video demonstration experiment in Chapter 3 [92]. Using realistic avatars as models for social stimuli might prove very important for understanding cognitive phenomena, such as audience effects, turn-taking and dynamics of vocalisations used during social interactions [93,94]. However, the most significant advancements in our understanding would be enabled by the upcoming state-of-the-art brain-computer interfaces (BCIs), implants that can be thought of as 'mind readers' [95]. These devices can map neuronal activity at a higher resolution in real-time. In future, such implants would make it possible to directly compare neuronal activity during vocal behaviours of freely moving humans and non-human primates.

The BCI and ML technologies described above have a bright future in animal behaviour studies. However, these methods are currently not yet fully developed and generalised to be directly deployed for studying different species and behaviours. For example, BCIs could be a promising paradigm in interpreting the neuronal basis of behaviours. However, its application and implementations are facing serious ethical concerns [96,97]. Understanding the implications of BCI on the welfare of both animal and human models should be our priority, along with developing a robust framework of ethical principles for BCI research [98].

One of the limitations of ML techniques is that the exact basis of classification or clustering of data used by the ML algorithm is not clear in many of the cases. A poor understanding of the basis then can create significant constraints on the reproducibility of the outcomes [72]. Also, it can lead to outcomes that are based on the hidden biases in the raw data [99,100]. Finally, although ML algorithms can classify simple postures in animals, detecting and classifying complex social interactions between individuals using ML techniques remains challenging and needs considerable research and development. Thus, careful and detailed field-based observations of animals by researchers remain at the core of animal communication and cognition research.

5.3 Conclusion

Results presented here have implications for current theory and future research on communication and cognition in non-human primates. All three facets—social learning, functional flexibility, pragmatic inference— were tested with field experiments and, put together, ascribe more flexibility than previously suggested to the vocal behaviour of a small-brained, non-human primate [101]. Furthermore, the sheer presence of these facets indicates that the corresponding albeit primitive cognitive infrastructure, which is also essential for human language, already existed in the common ancestor of humans and Old World monkeys [102,103], in line with the general hypothesis that human language evolved on the homologous path within the primate phylogenetic tree. However, the lack of evidence for other key cognitive features of the language, such as hierarchical organisation of thought and recursion in non-human species, indicates the need for further empirical investigations into the origins of specific cognitive properties [70,104]. Finally, the experiments in this thesis reiterate the importance of comparative research on animal communication to gain significant insights into the roots and evolution of human language [70]. In the future, detailed natural observations and experimental field research may be able to increasingly benefit from advanced technologies and data processing, whereas theoretical advancement is likely to benefit from a non-dichotomous view when mapping the evolutionary journey of communication systems from non-humans to humans. With such an integrative approach, science may soon tell us if Malcolm de Chazal was right when he said, "monkey sees a monkey when (s)he looks into the mirror".

5.4 References

1. Markson L, Bloom P. 1997 Evidence against a dedicated system for word learning in children. *Nature* **385**, 813–815. (doi:10.1038/385813a0)

2. Kaminski J, Call J, Fischer J. 2004 Word learning in a domestic dog: Evidence for 'fast mapping'. *Science (80-.).* **304**, 1682–1683. (doi:10.1126/science.1097859)

3. Seyfarth RM, Cheney DL. 1986 Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640–1658. (doi:10.1016/S0003-3472(86)80252-4)

4. Janik VM, Slater PB. 2000 The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11. (doi:10.1006/anbe.2000.1410)

5. Spelke E. 1994 Initial knowledge: six suggestions. *Cognition* **50**, 431–445. (doi:10.1016/0010-0277(94)90039-6)

6. Spelke ES, Lee SA. 2012 Core systems of geometry in animal minds. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2784–2793. (doi:10.1098/rstb.2012.0210)

7. Tsutsumi S, Ushitani T, Tomonaga M, Fujita K. 2012 Infant monkeys' concept of animacy: The role of eyes and fluffiness. *Primates* **53**, 113–119. (doi:10.1007/s10329-011-0289-8)

8. Shettleworth SJ. 2012 Modularity, comparative cognition and human uniqueness. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2794–2802. (doi:10.1098/rstb.2012.0211)

9. Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, Bakeman R. 2013 Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6318–23. (doi:10.1073/pnas.1300337110)

10. Clay Z, Archbold J, Zuberbühler K. 2015 Functional flexibility in wild bonobo vocal behaviour. *PeerJ* **3**, e1124. (doi:10.7717/peerj.1124)

11. Faraut L, Siviter H, Pesco FD, Fischer J. 2019 How life in a tolerant society affects the usage of grunts: evidence from female and male Guinea baboons. *Anim. Behav.* **153**, 83–93. (doi:10.1016/j.anbehav.2019.05.003)

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

13. Seyfarth RM, Cheney DL. 2016 Precursors to language: Social cognition and pragmatic inference in primates. *Psychon. Bull. Rev.* **24**, 1–6. (doi:10.3758/s13423-016-1059-9)

14. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)

15. Owren MJ, Rendall D. 2001 Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol. Anthropol.* **10**, 58–71. (doi:10.1002/evan.1014)

16. Owren MJ, Rendall D. 1997 An Affect-Conditioning Model of Nonhuman Primate Vocal Signaling. pp. 299–346. (doi:10.1007/978-1-4899-1745-4_10)

17. Zuberbühler K. 2016 Social concepts and communication in nonhuman primates. In *Psychological Mechanisms in Animal Communication*, Springer.

18. Seyfarth RM, Cheney DL. 2016 The origin of meaning in animal signals. *Anim. Behav.* , 1–8. (doi:10.1016/j.anbehav.2016.05.020)

19. Snowdon C. 2018 Cognitive Components of Vocal Communication: A Case Study. *Animals* **8**, 126. (doi:10.3390/ani8070126)

20. Schamberg I, Wittig RM, Crockford C. 2018 Call type signals caller goal: a new take on ultimate and proximate influences in vocal production. *Biol. Rev.* **93**, 2071–2082. (doi:10.1111/brv.12437)

21. Oller DK, Griebel U, Warlaumont AS. 2016 Vocal Development as a Guide to Modeling the Evolution of Language. *Top. Cogn. Sci.* **8**, 382–392. (doi:10.1111/tops.12198)

22. Griebel U, Oller DK. 2008 Evolutionary forces favoring communication flexibility. *Evol. Commun. Flex. Complexity, Creat. Adapt. Hum. Anim. Commun.* , 9–40. (doi:10.1093/icb/icp087)

23. Bard KA, Dunbar S, Maguire-Herring V, Veira Y, Hayes KG, Mcdonald K. 2014 Gestures and social-emotional communicative development in chimpanzee infants. *Am. J. Primatol.* **76**, 14–29. (doi:10.1002/ajp.22189)

24. Fröhlich M, van Schaik CP. 2020 Must all signals be evolved? A proposal for a new classification of communicative acts. *Wiley Interdiscip. Rev. Cogn. Sci.* , 1–13. (doi:10.1002/wcs.1527)

25. Gavrilov N, Nieder A. 2021 Distinct neural networks for the volitional control of vocal and manual actions in the monkey homologue of Broca's area. *Elife* **10**, 1–24.

26. Hage SR. 2018 Dual neural network model of speech and language evolution: new insights on flexibility of vocal production systems and involvement of frontal cortex. *Curr. Opin. Behav. Sci.* **21**, 80–87. (doi:10.1016/j.cobeha.2018.02.010)

27. Carrive P. 1993 The periaqueductal gray and defensive behavior: Functional representation and neuronal organization. *Behav. Brain Res.* **58**, 27–47. (doi:10.1016/0166-4328(93)90088-8)

28. Jürgens U, Pratt R. 1979 Role of the periaqueductal grey in vocal expression of emotion. *Brain Res.* **167**, 367–378. (doi:10.1016/0006-8993(79)90830-8)

29. Nieder A, Mooney R. 2020 The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190054. (doi:10.1098/rstb.2019.0054)

30. Newman JD. 2007 Neural circuits underlying crying and cry responding in mammals. *Behav. Brain Res.* **182**, 155–165. (doi:10.1016/j.bbr.2007.02.011)

31. Schwartz JW, Engelberg JWM, Gouzoules H. 2020 Evolving views on cognition in animal vocal communication: Contributions from scream research. *Anim. Behav. Cogn.* **7**, 192–213. (doi:10.26451/abc.07.02.12.2020)

32. Blumstein DT, Récapet C. 2009 The sound of arousal: The addition of novel non-linearities increases responsiveness in marmot alarm calls. *Ethology* **115**, 1074–1081. (doi:10.1111/j.1439-0310.2009.01691.x)

33. Fitch WT, Neubauer J, Herzel H. 2002 Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav.* **63**, 407–418. (doi:10.1006/anbe.2001.1912)

34. Gouzoules H, Gouzoules S, Tomaszycki M. 1998 Agonistic screams and the classification of dominance relationships: Are monkeys fuzzy logicians? *Anim. Behav.* **55**, 51–60. (doi:10.1006/anbe.1997.0583)

35. Gouzoules H, Gouzoules S. 2000 Agonistic screams differ among four species of macaques: The significance of motivation-structural rules. *Anim. Behav.* **59**, 501–512. (doi:10.1006/anbe.1999.1318)

36. Hammerschmidt K, Fischer J. 1998 The vocal repertoire of Barbary macaques: A quantitative analysis of a graded signal system. *Ethology* **104**, 203–216. (doi:10.1111/j.1439-0310.1998.tb00063.x)

37. Mercier S, Déaux EC, van de Waal E, Bono AEJ, Zuberbühler K. 2019 Correlates of social role and conflict severity in wild vervet monkey agonistic screams. *PLoS One* **14**, e0214640. (doi:10.1371/journal.pone.0214640)

38. Slocombe KE, Zuberbühler K. 2007 Chimpanzees modify recruitment screams as a function of audience composition. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 17228–33. (doi:10.1073/pnas.0706741104)

39. Fedurek P, Slocombe KE, Zuberbühler K. 2015 Chimpanzees communicate to two different audiences during aggressive interactions. *Anim. Behav.* **110**, 21–28. (doi:10.1016/j.anbehav.2015.09.010)

40. Arnal LH, Flinker A, Kleinschmidt A, Giraud AL, Poeppel D. 2015 Human Screams Occupy a Privileged Niche in the Communication Soundscape. *Curr. Biol.* **25**, 2051–2056. (doi:10.1016/j.cub.2015.06.043)

41. Engelberg JWM, Schwartz JW, Gouzoules H. 2019 Do human screams permit individual recognition? *PeerJ* **2019**. (doi:10.7717/peerj.7087)

42. Fugate JMB, Gouzoules H, Nygaard LC. 2008 Recognition of rhesus macaque (Macaca mulatta) noisy screams: Evidence from conspecifics and human listeners. *Am. J. Primatol.* **70**, 594–604. (doi:10.1002/ajp.20533)

43. Owren MJ, Rendall D. 2003 Salience of Caller Identity in Rhesus Monkey (Macaca, mulatta) Coos and Screams: Perceptual Experiments with Human (Homo sapiens) Listeners. *J. Comp. Psychol.* **117**, 380–390. (doi:10.1037/0735-7036.117.4.380)

44. Hoppitt W, Laland KN. 2008 Chapter 3 Social Processes Influencing Learning in Animals: A Review of the Evidence. *Adv. Study Behav.* **38**, 105–165. (doi:10.1016/S0065-3454(08)00003-X)

45. Seyfarth RM, Cheney DL. 1980 The Ontogeny of Vervet Monkey Alarm Calling Behavior: A Preliminary Report. *Z. Tierpsychol.* **54**, 37–56. (doi:10.1111/j.1439-0310.1980.tb01062.x)

46. Arnold K, Zuberbühler K. 2008 Meaningful call combinations in a non-human primate. *Curr. Biol.* **18**, 202–203. (doi:10.1016/j.cub.2008.01.040)

47. Berthet M, Mesbahi G, Pajot A, Casar C, Neumann C, Zuberbühler K. 2018 Titi monkeys combine alarm calls to create probabilistic meaning. *Sci. Adv.* , 1–10.

48. Zuberbühler K. 2001 Predator-specific alarm calls in Campbell's monkeys, Cercopithecus campbelli. *Behav. Ecol. Sociobiol.* **50**, 414–422. (doi:10.1007/s002650100383)

49. Crockford C, Wittig RM, Mundry R, Zuberbühler K. 2012 Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* **22**, 142–146. (doi:10.1016/j.cub.2011.11.053)

50. Zuberbühler K. 2000 Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* **76**, 195–207. (doi:10.1016/S0010-0277(00)00079-2)

51. Wheeler BC, Hammerschmidt K. 2013 Proximate factors underpinning receiver responses to deceptive false alarm calls in wild tufted capuchin monkeys: Is it counterdeception? *Am. J. Primatol.* **75**, 715–725. (doi:10.1002/ajp.22097)

52. Arnold K, Zuberbühler K. 2013 Female Putty-Nosed Monkeys Use Experimentally Altered Contextual Information to Disambiguate the Cause of Male Alarm Calls. *PLoS One* **8**, 2–8. (doi:10.1371/journal.pone.0065660)

53. Halina M, Rossano F, Tomasello M. 2013 The ontogenetic ritualization of bonobo gestures. *Anim. Cogn.* **16**, 653–666. (doi:10.1007/s10071-013-0601-7)

54. Whiten A, van de Waal E. 2018 The pervasive role of social learning in primate lifetime development. *Behav. Ecol. Sociobiol.* **72**. (doi:10.1007/s00265-018-2489-3)

55. Enquist M, Lind J, Ghirlanda S. 2016 The power of associative learning and the ontogeny of optimal behaviour. *R. Soc. Open Sci.* **3**. (doi:10.1098/rsos.160734)

56. Cheney DL, Seyfarth RM, Silk JB. 1995 The role of grunts in reconciling opponents and facilitating interactions among adult female Baboons. *Anim. Behav.* **50**, 249–257. (doi:10.1006/anbe.1995.0237)

57. Silk JB, Seyfarth RM, Cheney DL. 2016 Strategic use of affiliative vocalizations by wild female baboons. *PLoS One* **11**, 1–10. (doi:10.1371/journal.pone.0163978)

58. Fedurek P, Neumann C, Bouquet Y, Mercier S, Magris M, Quintero F, Zuberbühler K. 2019 Behavioural patterns of vocal greeting production in four primate species. *R. Soc. Open Sci.* **6**, 182181. (doi:10.1098/rsos.182181)

59. Dezecache G, Zuberbühler K, Davila-ross M. 2020 Flexibility in wild infant chimpanzee vocal behavior. , 1–17. (doi:10.1093/jole/lzaa009)

60. Seyfarth RM, Cheney DL, Bergman TJ. 2005 Primate social cognition and the origins of language. *Trends Cogn. Sci.* **9**, 264–266. (doi:10.1016/j.tics.2005.04.001)

61. Gruber T, Grandjean D. 2017 A comparative neurological approach to emotional expressions in primate vocalizations. *Neurosci. Biobehav. Rev.* **73**, 182–190. (doi:10.1016/j.neubiorev.2016.12.004)

62. McComb K, Semple S. 2005 Coevolution of vocal communication and sociality in primates. *Biol. Lett.* **1**, 381–385. (doi:10.1098/rsbl.2005.0366)

63. Peckre L, Kappeler PM, Fichtel C. 2019 Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behav. Ecol. Sociobiol.* **73**. (doi:10.1007/s00265-018-2605-4)

64. Bouchet H, Blois-Heulin C, Lemasson A. 2013 Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Front. Psychol.* **4**, 1–15. (doi:10.3389/fpsyg.2013.00390)

65. Hammerschmidt K, Fischer J. 2019 Baboon vocal repertoires and the evolution of primate vocal diversity. *J. Hum. Evol.* **126**, 1–13. (doi:10.1016/j.jhevol.2018.10.010)

66. Shultz S, Dunbar RIM. 2007 The evolution of the social brain: Anthropoid primates contrast with other vertebrates. *Proc. R. Soc. B Biol. Sci.* **274**, 2429–2436. (doi:10.1098/rspb.2007.0693)

67. Dunbar RIM, Shultz S. 2007 Evolution in the social brain. *Science (80-.).* **317**, 1344–1347. (doi:10.1126/science.1145463)

68. Dunbar RIM. 2009 The social brain hypothesis and its implications for social evolution. *Ann. Hum. Biol.* **36**, 562–572. (doi:10.1080/03014460902960289)

69. Bowlingi DL *et al.* 2020 Rapid evolution of the primate larynx? *PLoS Biol.* **18**, 1–21. (doi:10.1371/JOURNAL.PBIO.3000764)

70. Hauser MD, Chomsky N, Fitch WTS. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science (80-.).* **298**, 1569–79. (doi:10.1126/science.298.5598.1569)

71. Fitch WT. 2000 The evolution of speech: A comparative review. *Trends Cogn. Sci.* **4**, 258–267. (doi:10.1016/S1364-6613(00)01494-7)

72. Valletta JJ *et al.* 2017 Applications of machine learning in animal behaviour studies. *Anim. Behav.* **124**, 203–220. (doi:10.1016/j.anbehav.2016.12.005)

73. Dezecache G, Zuberbühler K, Davila M, Christoph R. 2020 A machine learning approach to infant distress calls and maternal behaviour of wild chimpanzees. *Anim. Cogn.* (doi:10.1007/s10071-020-01437-5)

74. Fedurek P *et al.* 2016 Sequential information in a great ape utterance. *Sci. Rep.* **6**, 38226. (doi:10.1038/srep38226)

75. Greif S, Yovel Y. 2019 Using on-board sound recordings to infer behaviour of free-moving wild animals. *J. Exp. Biol.* **222**. (doi:10.1242/jeb.184689)

76. Couchoux C, Aubert M, Garant D, Réale D. 2015 Spying on small wildlife sounds using affordable collar-mounted miniature microphones: An innovative method to record individual daylong vocalisations in chipmunks. *Sci. Rep.* **5**, 1–8. (doi:10.1038/srep10118)

77. Gill LF, D'Amelio PB, Adreani NM, Sagunsky H, Gahr MC, ter Maat A. 2016 A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods Ecol. Evol.* **7**, 1349–1358. (doi:10.1111/2041-210X.12610)

78. Studd EK *et al.* 2019 Use of Acceleration and Acoustics to Classify Behavior, Generate Time Budgets, and Evaluate Responses to Moonlight in Free-Ranging Snowshoe Hares. *Front. Ecol. Evol.* **7**. (doi:10.3389/fevo.2019.00154)

79. Clink DJ, Klinck H. 2019 GIBBONR: An R package for the detection and classification of acoustic signals using machine learning.

80. Mac Aodha O *et al.* 2018 Bat detective—Deep learning tools for bat acoustic signal detection. *PLoS Comput. Biol.* **14**, 1–19. (doi:10.1371/journal.pcbi.1005995)

81. Mielke A, Zuberbühler K. 2013 A method for automated individual, species and call type recognition in free-ranging animals. *Anim. Behav.* **86**, 475–482. (doi:10.1016/j.anbehav.2013.04.017)

82. Stowell D, Stylianou Y, Wood M, Pamuła H, Glotin H. 2018 Automatic acoustic detection of birds through deep learning: the first Bird Audio Detection challenge. , 1–21.

83. Letters E, Ovaskainen O, Moliterno de Camargo U, Somervuo P, Letters E. 2018 Animal Sound Identifier (ASI): software for automated identification of vocal animals. *Ecol. Lett.* **21**, 1244–1254. (doi:10.1111/ele.13092)

84. Wellman M, Srour N, Hillis D. 1997 Acoustic Feature Extraction for a Neural Network Classifier.

85. Ropert-Coudert Y, Wilson RP. 2005 Trends and perspectives in animal-attached remote sensing. *Front. Ecol. Environ.* **3**, 437–444. (doi:10.1890/1540-9295(2005)003[0437:TAPIAR]2.0.CO;2)

86. Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. 2013 Observing the unwatchable through acceleration logging of animal behavior. *Anim. Biotelemetry* **1**, 1–16. (doi:10.1186/2050-3385-1-20)

87. Dahl CD, Wyss C, Zuberbühler K, Bachmann I. 123 ADSocial information in equine movement gestalts. *Anim. Cogn.* **0**, 0. (doi:10.1007/s10071-018-1193-z)

88. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. 2018 DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289. (doi:10.1038/s41593-018-0209-y)

89. Lin TY, Goyal P, Girshick R, He K, Dollar P. 2020 Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 318–327. (doi:10.1109/TPAMI.2018.2858826)

90. Rathore A, Sharma A, Sharma N, Guttal V. 2020 Multi-Object Tracking in Heterogeneous environments (MOTHe) for animal space-use studies.

91. Graving JM, Chae D, Naik H, Li L, Koger B, Costelloe BR, Couzin ID. 2019 Deepposekit, a software toolkit for fast and robust animal pose estimation using deep learning. *Elife* **8**, 1–42. (doi:10.7554/eLife.47994)

92. Siebert R, Taubert N, Spadacenta S, Dicke PW, Giese MA, Thier P. 2020 A Naturalistic Dynamic Monkey Head Avatar Elicits Species-Typical Reactions and Overcomes the Uncanny Valley. **7**, 1–17.

93. Zuberbühler K. 2008 Audience effects. *Curr. Biol.* **18**, 189–190. (doi:10.1016/j.cub.2007.12.041)

94. Pika S, Wilkinson R, Kendrick KH, Vernes SC. 2018 Taking turns: bridging the gap between human and animal communication. *Proc. R. Soc. B Biol. Sci.* **285**. (doi:10.1098/rspb.2018.0598)

95. Willett FR, Avansino DT, Hochberg LR, Henderson JM, Shenoy K V. 2021 High-performance brain-to-text communication via handwriting. *Nature* **593**, 249–254. (doi:10.1038/s41586-021-03506-2)

96. Hildt E. 2015 What will this do to me and my brain? Ethical issues in brain-to-brain interfacing. *Front. Syst. Neurosci.* **9**. (doi:10.3389/fnsys.2015.00017)

97. Kotchetkov IS, Hwang BY, Appelboom G, Kellner CP, Connolly ES. 2010 Brain-computer interfaces: Military, neurosurgical, and ethical perspective. *Neurosurg. Focus* **28**, 1–6. (doi:10.3171/2010.2.FOCUS1027)

98. Pham M, Goering S, Sample M, Huggins JE, Klein E. 2018 Asilomar survey: researcher perspectives on ethical principles and guidelines for BCI research. *Brain-Computer Interfaces* **5**, 97–111. (doi:10.1080/2326263X.2018.1530010)

99. Dreyfus HL, Dreyfus SE. 1992 What artificial experts can and cannot do. *AI Soc.* **6**, 18–26. (doi:10.1007/BF02472766)

100. Kim B, Kim H, Kim K, Kim S, Kim J. 2019 Learning not to learn: Training deep neural networks with biased data. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2019-June**, 9004–9012. (doi:10.1109/CVPR.2019.00922)

101. Hammerschmidt K, Fischer J. 2013 Constraints in Primate Vocal Production. In *Evolution of Communicative Flexibility*, pp. 92–119. Cambridge, MA, USA: The MIT Press. (doi:10.7551/mitpress/9780262151214.003.0005)

102. Cheney DL, Seyfarth RM. 1992 *How monkeys see the world: Inside the mind of another species*. The University of Chicago Press.

103. Cheney DL, Seyfarth RM. 2008 *Baboon metaphysics: the evolution of a social mind*. The University of Chicago Press.

104. Tecumseh Fitch W. 2020 Animal cognition and the evolution of human language: Why we cannot focus solely on communication. *Philos. Trans. R. Soc. B Biol. Sci.* **375**. (doi:10.1098/rstb.2019.0046)