



Evolving Views on Cognition in Animal Vocal Communication: Contributions from Scream Research

Jay W. Schwartz¹*, Jonathan W. M. Engelberg¹, and Harold Gouzoules¹

¹ Department of Psychology, Emory University

*Corresponding author (Email: schwartzj@wou.edu)

Citation – Schwartz, J. W., Engelberg, J. W. M., & Gouzoules, H. (2020). Evolving views on cognition in animal vocal communication: Contributions from scream research. *Animal Behavior and Cognition*, 7(2), 192-213. doi: <https://doi.org/10.26451/abc.07.02.12.2020>

Abstract – The last four decades have seen major advances in the study of the cognitive bases of animal vocal communication. The conceptual delineation of senders and receivers has led to a focus on the cognitive processes involved in call production and usage, and those concerning call perception. We review selected relevant literature and discuss how recent thought—specifically, the recognition of a distinction between acoustic variation within versus between call types—expands focus from two to four key questions, specifically, the identification of factors that influence: (1) the usage of different call types, (2) the particular acoustic structure of a vocalization, (3) receivers’ responses to different call types, and (4) receivers’ responses to acoustic variation within a call type. We present the findings of a case study examining the relationship between emotional arousal and the acoustics of scream classes of rhesus macaques (*Macaca mulatta*) and discuss how the findings relate to broader issues in animal vocal communication and cognition. Screams are phylogenetically widespread, generally highly emotional, and in some taxa, convey complex information about social interactions, making them a promising subject for future research addressing contemporary questions in the field.

Keywords – Vocal communication, Emotion, Arousal, External reference, Primate, Screams

Robert Seyfarth and Dorothy Cheney’s early postdoctoral collaborations with Peter Marler, beginning with their seminal articles (Seyfarth et al., 1980a, b) on the alarm calls of vervet monkeys (*Chlorocebus pygerythrus*)—showing that these calls elicit appropriate responses from receivers even in the absence of predators—had an essentially transcendent influence on the study of animal, and particularly nonhuman primate, vocal communication. The period leading up to their work together was striking in that monkey and ape vocal communication was thought to have only a limited, and basically ancillary, role in the complex social systems that were being discovered and documented in the burgeoning field and laboratory research of the time. For example, in an influential review of the topic, Lancaster (1968) wrote that:

... vocalizations do not carry the major burden of meaning in most social interactions, but function instead either to call visual attention to the signaler or to emphasize or enhance the effect of visual or tactile signals. In other words, a blind monkey would be greatly handicapped in his social interactions whereas a deaf one would probably be able to function almost normally (p. 442).

Views would change considerably after 1980.

Marler's former student, Thomas Struhsaker, had earlier conducted field research on vervet monkeys in the Amboseli Reserve in Kenya and had published the first extensive documentation of the vocal repertoire of the species (Struhsaker, 1967), including a description of the acoustically distinct alarm calls used in response to different kinds of predators. Struhsaker had set the stage for the later research on vervet alarm calls, but the initial response to the finding was rather tepid. Again, Lancaster (1968) cautioned that the:

differentiation of high-intensity alarm calls to communicate some information about the environment is a specialization that should not be thought of as pointing toward the kind of major revolution in information content suggested as a requisite of human language (p. 446).

There would be no follow-up investigations of vervet monkey alarm calls for almost a decade.

It is not a coincidence that interest in vervet alarms would be rekindled in the Marler lab. His colleague at The Rockefeller University, Donald Griffin, renowned for his pioneering studies of bat echolocation, would open the door to a whole new world of research possibilities with the publication of his 1976 book, *The Question of Animal Awareness*, which he revised and expanded in 1981. With the original slim volume, Griffin anticipated and encouraged research in what is now the rapidly expanding field of Cognitive Ethology, the study of the role of cognitive processes in natural animal behavior. Griffin (1976) saw research on communication as a "window on the minds of animals" (p. 87), and this would be the guiding principle that would come to underpin the remarkable research program Seyfarth et al. launched with their 1980 articles. What do vocalizations mean and to what extent do they reveal "How Monkeys See the World" (Cheney & Seyfarth, 1990)? An explosion of interest in monkey and ape vocal communication followed, summarized nicely in Cheney and Seyfarth (2008).

Our primary objective in this review is to describe a reframing of the issues that we believe is underway within the field of animal communication. We focus primarily on the literature on nonhuman primate vocal communication, but the principles extend more broadly. We begin by delineating key cognitive terms and concepts, reviewing in particular the significance of the distinction between senders and receivers (Seyfarth & Cheney, 2003b) for the role of cognition in nonhuman primate vocal communication. We explore how this distinction has organized the field, and review some of the last 40 years' most significant findings relating to the cognitive processes underlying the behavior of senders and receivers, some of which concern primate screams. We then address the distinction of acoustic variation within versus between call types, and discuss recent directions regarding the functions of, and cognitive processes involved in, the production and perception of each source of variation. These recent directions, we argue, amount to a reorganization of the central questions that frame empirical and theoretical research in the field. We suggest that screams—a phylogenetically widespread call type surrounded by complex cognitive dynamics—comprise a particularly promising model for research into some of the new questions that are emerging. To illustrate this, we present the results of a case study investigating the role of emotional arousal in determining acoustic variation of rhesus macaque (*Macaca mulatta*) screams.

Terminology

The history and sometimes contentious debate surrounding cognitive terms in the animal communication literature is extensive and, indeed, daunting to summarize (Cheney & Seyfarth, 1990, 2008). For the purposes of this paper—the primary objective of which is to outline an emerging framework for organizing research questions, rather than to provide a taxonomy of cognitive processes, *per se*—we present a brief overview of relevant terms and concepts.

That monkeys respond to conspecific predator alarm calls even in the absence of predators (Seyfarth et al., 1980a, b) suggests that the calls convey *information* about the type of threat. "Information" in this sense means that a signal disambiguates (or reduces the uncertainty about) a particular event or condition of the world for a receiver, potentially influencing its subsequent behavior (Seyfarth et al., 2010). This perspective entails the assumption that receivers possess cognitive capacities that allow for processing this information but makes no specific claim as to the precise nature of these capacities. Some monkey vocalizations, such as anti-predator alarm calls, have been proposed to be

functionally referential in that they fulfill two criteria: the calls occur with substantial reliability and exclusivity in response to specific events or stimuli (or *referents*; an exception being deceptive usage), and they elicit appropriate responses even in the absence of those referents (Macedonia & Evans, 1993; see also Gouzoules, 2005; Scarantino, 2013; Wheeler & Fischer, 2012). Like “information,” the term “functional reference” is agnostic regarding the specific cognitive processes involved in an instance of communication—a fact that has been the basis for criticism regarding the epistemological utility of the concept (Wheeler & Fischer, 2012). However, functionally referential vocalizations are generally assumed to be associated with *mental representations*, “images, rules, propositions, or symbols that guide an organism's ... behavior” (Cheney & Seyfarth, 1990, p. 15). Thus, organisms make decisions about behavior, in ways assisted by information about things, events, and experiences that is represented (somehow) in the brain. A more precise definition that includes brain mechanisms is lacking, although as Cheney and Seyfarth (2008, p. 240) noted, this deficiency does not mean the construct cannot be useful.

Within the domain of communication, mental representations have been distinguished from lower-level perceptual processing in which “recipients of a signal might simply be attending to the physical features of an alarm call ... and it is this percept alone, rather than some intervening mental representation of the referent, that guides subsequent behavior” (Zuberbühler et al., 1999, p. 2). In addition, mental representations about external referents have sometimes been contrasted with *emotional* (or *affective*) processes. These generally comprise temporary, multicomponent internal states characterized by levels of arousal (ranging from sleepy to alert) and valence (ranging from very negative to very positive), possibly but not necessarily entailing a subjective “feeling” component (Adolphs & Amler, 2018; Anderson & Adolphs, 2014; de Waal, 2011; Mendl et al., 2010; Paul & Mendl, 2018). Emotional states can occur in tandem with, evoke, or be elicited by mental representations. A related distinction is that between *external* and *internal reference*, with the former describing communication in which information conveyed relates to external events or stimuli, while, in the latter, information relates to the internal state of the sender; as we discuss in greater depth further below, these two kinds of reference can potentially coexist in a single calling system. Mental representations that are associated with vocalizations are assumed to allow for discriminations that relate to membership in *perceptual* (shared perceptual features), *functional* (shared relevance or consequence beyond perceptual attributes), and *relational / abstract* (e.g., instantiating a specific relationship such as kinship or dominance) categories (Shettleworth, 2010, p. 191).

Senders and Receivers

Communication is an intrinsically social event. Minimally, communication has taken place when a sender emits a signal that is detected by a receiver whose behavior might consequently change. In a pair of pivotal papers, Seyfarth and Cheney delineated a critical distinction between senders and receivers (Seyfarth & Cheney, 2003a, b). Responding to what they perceived as the perpetuation of a logically false dichotomy between emotional and externally referential signaling, their central argument was that vocalizations can be emotional on the part of senders and nonetheless convey external information, as the latter depends on receiver cognition—specifically, whether and how receivers conjure a mental representation of the external referent solely from hearing a vocalization. To illustrate this point, Seyfarth and Cheney drew on a point made by Premack (1972): a signal that (hypothetically) occurs entirely as a result of the emotional state of a vocalizer, and does not derive from mental representations of the external event that generated those emotions, can be used by receivers to obtain information about the presence of that external event, as long as the signal is reliably and exclusively correlated with the external stimulus. Thus, in principle, the production of vervet monkey alarms could be motivated by “three flavors of anxiety or arousal: leopard anxiety, eagle anxiety, and snake anxiety” (Dennett, 1983, p. 346) and this would not be incompatible with the view of these vocalizations as externally referential. As Seyfarth and Cheney note:

The affective and referential properties of signals are ...logically distinct, at least in animal communication, because the former depends on mechanisms of call production in the signaler,

whereas the latter depends on the listener's ability to extract information from events in its environment (Seyfarth & Cheney, 2003b, p. 154).

Thus, Seyfarth and Cheney (2003a, b) delineated two theoretically independent research questions that each have received much attention in the ensuing years: (1) what are the roles of emotion and other cognitive processes in signal production by senders, and (2) what is the role of information about, and mental representation of, external objects and events in receiver behavior? Indeed, in the following decades, research on external reference largely focused on receivers rather than senders, while attention to senders has, until recent years, largely focused on the emotional underpinnings of vocalizations. This literature is briefly reviewed below.

Receiver Cognition

Alarm Calls, External Reference, and Affect-Induction

The cognitive bases of the behavioral responses of vervet monkeys to playbacks in the experiments by Seyfarth et al. (1980a, b)—subsequently shown in many other species—have been the focus of much discussion (e.g., Cheney & Seyfarth, 1990, 2008; Gouzoules, 2005; Gouzoules & Gouzoules, 2006; Gouzoules et al., 1985, 1998; Owren & Rendall, 2001; Scarantino & Clay, 2015; Seyfarth & Cheney, 2003a, b, 2015, 2017; Seyfarth et al., 2005; Townsend & Manser, 2013; Wheeler & Fischer, 2015; Zuberbühler, 2003; Zuberbühler et al., 1999). Do leopard alarm calls elicit in a vervet a mental representation of a large, spotted cat with sharp claws and fangs? Perhaps it is a vaguer representation of some large quadruped, or simply one of terrestrial danger more generally? Precise empirical discrimination among these possibilities is not possible through the traditional playback experiment paradigm, but some conclusions that apply to all of these possible scenarios appear reasonable. First, the call evokes in the receiver a mental representation of some object, agent, and/or event that is external to the sender. Second, this mental representation is categorical, in the sense that receivers classify the referent of the “leopard” alarm as distinct from the referents of other alarm call types (Zuberbühler et al., 1999). It is plausible that this mental representation involves a visual component (i.e., a mental image of something), since receivers respond to alarm calls as if they have seen a predator (for direct evidence of this in birds, see Suzuki, 2018).

In principle, functionally referential communication need not involve mental representations, and some have argued other cognitive hypotheses are more parsimonious (Owren & Rendall, 2001; Rendall et al., 2009). One such alternative explanation is that receivers respond to calls via stimulus-response reflexes resulting in direct activation of motor neurons (e.g., Lloyd, 1989). Such an account would, however, predict that responses to calls would be consistent and fixed, whereas monkeys respond to conspecific calls in a variable, context-dependent fashion (e.g., Price & Fischer, 2014; Price et al., 2015). Incidentally, this observation suggests that receivers somehow take context into account when responding (or not) to externally referential signals, sparking debate on whether responses to purportedly functionally referential signals differ meaningfully from responses to other stimuli (Scarantino & Clay, 2015; Wheeler & Fischer, 2012, 2015). This observation is not incompatible with hypotheses about mental representation, but it is interesting and indicates that external reference is not sufficient to explain at least some cases of receiver behavior, since receivers perceive and process not only referential calls but other aspects of the social and natural environments.

Another alternative account posited to explain instances of functional reference is that signals function by manipulating the receiver's emotional or other affective state (Owren & Rendall, 1997, 2001; Rendall et al., 2009). Under this “affect-induction model,” the acoustic features of calls are hypothesized to play a crucial role in that they directly affect the receiver's internal state, e.g., characteristics of the sounds that irritate or excite receivers (Owren & Rendall, 2001). Affect-induction likely plays a role in some instances of animal communication (discussed further below). In principle, affect-induction could function in tandem with evocation of mental representations of external stimuli, but advocates of this model contend that it can fully account for many instances of vocal communication (Owren & Rendall,

1997, 2001; for review, see Gouzoules, 2005). However, a set of habituation-dishabituation experiments conducted by Zuberbühler et al. (1999) with Diana monkeys (*Cercopithecus diana*) strongly suggest that mental representation is involved in receivers' responses to alarm calls in this species. Monkeys habituated to leopard alarm calls did not dishabituate to a leopard's growl, even though the alarm call and the growl are acoustically distant. However, they did dishabituate to an eagle alarm call. These results contradict the view that acoustic properties and induced affective responses are primary determinants of receivers' responses to calls, and instead suggest that predator-specific calls evoke cognitive representations of different kinds of threats (Gouzoules, 2005; Zuberbühler, 2003; Zuberbühler et al., 1999). The mainstream view in the field of animal communication is that, in most cases, receiver responses to functionally referential calls probably entail mental representation of external referents (Gouzoules, 2005; Gouzoules et al., 1985; Seyfarth & Cheney, 2003a, b; Wheeler & Fischer, 2012; Zuberbühler, 2003; though see Rendall et al., 2009).

Representation of Individuals, Kin Groups, and Dominance Interactions

Following on the groundbreaking research on vervet alarm calls, other types of vocalizations were found to convey externally referential information concerning social dimensions of monkey and ape behavior. These findings, reviewed briefly below, suggest that animal vocalizations have the capacity to evoke mental representations far more complex than simple one-to-one, call-to-object relationships. First, receivers derive information from vocalizations that permits recognition, and mental representation, of individual senders as well as, in some species, the kin groups to which individuals belong (Seyfarth & Cheney, 2015; Seyfarth et al., 2005), although there has been debate about the degree to which different call types provide sufficient information to evoke these representations (Engelberg et al., 2019; Fugate et al., 2008; Rendall et al., 1996, 1998). Chacma baboon (*Papio ursinus*) societies, like those of several other catarrhine monkeys, are structured by a matrilineal hierarchy system in which females are ranked within and between matrilineal groups. Rank challenges—aggressive behavior by a lower-ranking individual and/or submission behavior by a higher-ranking individual—can occur between or within matrilineal groups; these sometimes lead to dominance rank reversals. Between-matriline rank challenges should be of greater interest to group members as they represent greater potential upheavals of the social structure. Bergman et al. (2003) played sequences of female dominance grunts and “submission” screams simulating interactions congruent with the hierarchy (dominant grunt of a higher-ranking individual, submissive scream of a lower-ranking individual) and incongruent rank challenges (the reverse sequence) within and between matrilineal groups. Listeners looked longer in response to incongruous, between-matriline sequences than to other combinations, even when the rank distance between the two vocalizers in the within-matriline sequence was great. To show this pattern, baboons must not only recognize groupmates from their vocalizations alone, but also hold some representation of the boundaries between, and hierarchical structure of, the matrilineal groups to which groupmates belong, and form expectations based on these representations.

Like baboons, rhesus macaques live in social groups with matrilineal hierarchies. During agonistic interactions, attacked or threatened victims often scream; rather than merely signaling submission (more formally signaled using bared-teeth facial displays; de Waal & Luttrell, 1985), rhesus screams recruit aid, primarily from matrilineal kin (Bernstein & Ehardt, 1985). Gouzoules et al. (1984) provided evidence that rhesus screams fall into several acoustically distinguishable classes, the usage of which correlates with the intensity of agonism received and the dominance rank of the opponent (see also Gouzoules, 2005; Gouzoules et al., 1985, 1998). Interventions by matrilineal kin may reduce the likelihood of harm to the victim and/or protect the rank of the matrilineal group against challenges from lower-ranking individuals (Kaplan et al., 1987). In playback experiments, mothers responded most strongly to the scream class most associated with potential risk of injury, but second-most strongly to the class associated with a threat from a lower-ranking opponent, even though the latter was rarely associated with physical danger (Gouzoules et al., 1984). Gouzoules et al. (1986) also found that the sex of juveniles and the degree of relatedness between the playback target and vocalizer influenced responses to these

screams. Playback of screams recorded from immature rhesus monkeys consistently elicited a response, not only from their mothers, but also from other adult female matrilineal relatives. Responses from these more distant relatives occurred in the absence of any cues from the immatures' mothers or any other contextual features that might have been originally available during an agonistic encounter. The pattern of female response to playback of immatures' screams closely resembled that of female aid to juvenile relatives described by Kaplan (1978) and suggested that scream vocalizations alone can function as a principal mechanism in recruitment. Moreover, the patterning of female response to different scream types provides additional support to the contention that screams convey detailed information about external events.

These findings suggest that screams evoke in receivers a mental representation of specific aspects of external events; furthermore, they indicate that receivers respond to instances of agonism associated with a threat to the matriline's place in the hierarchy. The social event that a rhesus macaque or baboon represents when responding to scream classes or call sequences may include the identity of the sender, kin and rank relationships, and different degrees of aggression. Do monkeys hold dichotomous mental categories for "higher-ranking" versus "lower-ranking," or do they mentally represent these as differences of degree? In support of the latter view, scream class usage is more consistent when the rank difference between the sender and the opponent is relatively great (Gouzoules et al., 1998). Specifically, usage of different scream classes might be underlain by a kind of fuzzy logic in which senders represent rank differences (and perhaps agonistic intensity) not as discrete categories separated by hard boundaries, but rather as graded, soft categories or perhaps continua (Gouzoules et al., 1998).

Vocalizations of many species may also convey information to receivers about the motivational state of the sender, i.e., internal reference (e.g., Gouzoules et al., 1985; Morton, 1977). However, most receivers probably do not attribute knowledge or emotional states, per se, to senders—doing so is highly advanced and probably exceeds the cognitive capabilities of most taxa (Cheney & Seyfarth, 2008; Seyfarth & Cheney, 2003a, b). Receivers instead use calls to simulate or anticipate a sender's likely subsequent behavior. Regardless, it is clear that calls that are not externally referential can nonetheless evoke complex mental representations in receivers. This has raised questions as to whether the cognitive bases of external reference are fundamentally different from those involved in receivers' responses to other kinds of calls and even to stimuli unrelated to communication (Wheeler & Fischer, 2012), a matter that remains open to empirical investigation.

In summary, research since 1980 has identified several possible cognitive capacities underlying receiver responses to calls, including categorical and hierarchical mental representation of external phenomena and of individuals and kin groups, induced emotional and other affective states, and stimulus-response reflexes. These are generally not mutually exclusive, and in many cases, mental representation and affect-induction are likely to play a role simultaneously.

Sender Cognition

Emotion

Dating back to Darwin (1872), there has been recognition that emotion plays a role in vocal production in animals (e.g., Briefer, 2012; Filippi, 2016; Fischer, 2017; Fischer & Price, 2016; Owren et al., 2011; Schamberg et al., 2018; Wheeler & Fischer, 2012). A view endorsed by some is that of animal vocalizations as largely involuntary outputs of emotional and other affective states (e.g., Owren et al., 2011; Panksepp, 2011). Indeed, evidence from brain stimulation as well as lesion studies suggests that vocalizations in some species are subcortically controlled, suggesting they might be closely tied to affect (Fichtel et al., 2001; Jürgens, 2009; reviewed in Owren et al., 2011). Humans, too, frequently emit involuntary, emotional vocalizations (sometimes referred to as "raw" affect bursts; Hawk et al., 2009; Schröder, 2003). Seyfarth and Cheney (2003a, b) noted that externally referential vocalizations in nonhuman primates could *in principle* be purely emotionally and involuntarily produced, but argued against this view, observing that human speech is often underlain by an intention to communicate and,

simultaneously, an emotional state, with the latter manifesting in predictable acoustic changes to the voice (e.g., Scherer, 2003). They also noted the wide array of factors that influence the usage of calls, including the composition of the audience (“audience effects,” Fichtel & Manser, 2010). For example, chimpanzees appear to modify their recruitment screams based not only on the intensity of aggression, as is the case for rhesus macaques (Gouzoules et al., 1984), but also on the rank composition of the audience: chimps exaggerate screams when there is a listener present who outranks the aggressor (Slocombe & Zuberbühler, 2007). Of course, differences in a sender’s emotional state in the presence, versus in the absence, of a particular audience could possibly account for these effects (Owren et al., 2011), but as one considers evidence for more extended context-dependency in call usage, the idea that this is underpinned by subtle differences in sender affect becomes less parsimonious, and thus invoking the involvement of other cognitive processes has been deemed more reasonable (Cheney & Seyfarth, 2018; Seyfarth & Cheney, 2017, 2018). Such efforts are discussed further below.

The predominant approach to emotion in the field of animal vocal production, adapted from the dimensional model (Mendl et al., 2010; Russell, 2003), recognizes the existence of at least two dimensions that characterize emotional states: arousal (level of activation, ranging from placid to alert) and valence (ranging from highly negative to highly positive) (Briefer, 2012; for reviews of emotion concepts in the field of vocal communication, see: Altenmüller et al., 2013; Juslin, 2013). While affect and emotion appear unable to wholly account for senders’ vocal production, research across taxa, especially in the last decade, has revealed consistent relationships between emotional states and acoustic properties of vocalizations (Altenmüller et al., 2013; Briefer, 2012; Zimmerman et al., 2013). Emotional arousal appears to generally result in increases in fundamental frequency (F0; the oscillation rate of the vocal folds, commonly perceived as pitch) (e.g., Briefer, Maigrot et al., 2015; Briefer, Tettamanti et al., 2015; Filippi et al., 2017; Rendall, 2003; Scherer, 2003; Szipl et al., 2017; for a possible exception, see Ordóñez-Gómez et al., 2019), as well as in variability of F0 (e.g., Bayart et al., 1990; Lingle et al., 2012; Rendall, 2003; Sugiura, 2007; Yamaguchi et al., 2010) and in noisiness (e.g., Liao et al., 2018; Puppe et al., 2005; Siebert et al., 2011; Stoeger et al., 2011; though see Blumstein & Chi, 2012; Linhart et al., 2015; Szipl et al., 2017). Emotional valence is associated with vocal duration in several species, with shorter vocalizations associated with more positive emotions and contexts (Briefer, 2012; Briefer, Maigrot et al., 2015; Briefer, Vizier et al., 2019; Fichtel et al., 2001; Friel et al., 2019; Taylor et al., 2009). These changes in the voice likely stem from the physiological effects of emotional states on the vocal apparatus, e.g., arousal tensing the laryngeal muscles thereby increasing the oscillation rate of the vocal folds. These cross-species trends are thought to represent evolutionary homology and are attributed to phylogenetic conservation of the mechanisms of vocal production as well as physiological arousal (Briefer, 2012; Filippi et al., 2017; Zimmermann et al., 2013).

Intentionality

Recent contributions to the area of sender cognition have explored the potential role of intentionality (Berthet et al., 2018; Crockford et al., 2017; Deshpande et al., 2018; Fischer & Price, 2016; Schel et al., 2013; Scott-Phillips, 2015a, b; Sievers & Gruber, 2016; Townsend et al., 2017). Authors have adapted Dennett’s (e.g., 1983) taxonomy of orders of intentionality to delineate different levels of social cognition on the part of senders (Fischer & Price, 2016; Schel et al., 2013; Townsend et al., 2017). Within this framework, “zero-order intentionality” entails no mental understanding of the outcome or function of one’s behavior; the view of a vocalization as purely emotional expression or otherwise involuntary would fall within this category. In “first-order intentionality,” senders use signals with some understanding that doing so brings about a response in receivers. Senders at this level do not necessarily understand that their signals change receivers’ mental states (e.g., awareness of a predator); communication exhibiting first-order intentionality might thus be best described as behavior induction. Mental state attribution is, however, a defining feature of “second-order intentionality,” in which a sender understands the function of communication in terms of its effects on a receiver’s mental state. Finally, in “third-order intentionality” a sender understands the function of a signal in terms of its effects on the receiver’s

perception of the sender's mental state; this and higher levels of intentional communication are thought to be unique to humans (Fischer & Price, 2016; Scott-Phillips, 2015a).

The essential point here is that there exists a spectrum of possible cognitive processes underlying signal usage, ranging from involuntary and emotional/affective signals (zero-order intentionality) to full-blown, human-like mental state attribution. The question is where different cases of vocalization by animals lie along this spectrum. Chimpanzees use “alert hoos” to signal potential but not impending danger such as the presence of a venomous snake in their path (Goodall, 1986). Recent research has shown that, upon encountering a researcher-planted artificial but life-like snake, wild chimpanzees not only emit alert hoos but target them selectively toward ignorant receivers, monitor receivers' reactions, and continue calling until the “snake” has been noticed (Crockford et al., 2012; Schel et al., 2013). This finding has been interpreted as evidence of at least first-order intentionality (Crockford et al., 2017; Schel et al., 2013), consistent with earlier evidence that captive chimpanzees used vocalizations flexibly to manipulate a human's behavior (Hopkins et al., 2007; Leavens et al., 2004). Thus, a sender understands that the behavior of the receiver is contingent upon the occurrence of the vocalizations, and attends to the receiver to assess whether the vocalizations are having their expected effect. The observation that senders are sensitive to whether the receiver is already informed or not (i.e., has already seen the snake or heard another individual's alert hoos) is perhaps suggestive of second-order intentionality, that is, an understanding on the part of senders that the alert hoos function not only to adjust receivers' behavior but also to inform them of the present danger (Seyfarth & Cheney, 2012). Definitive evidence for this interpretation is lacking, but given experimental corroboration that chimpanzees understand that seeing leads to knowing (reviewed in Call & Tomasello, 2008), it is not out of the question.

Evidence of second-order intentional signaling is lacking for most other nonhuman species (Cheney & Seyfarth, 2008; Seyfarth & Cheney, 2003b), but first-order intentionality may be a viable working hypothesis for many instances of animal communication (Townsend et al., 2017). For example, rhesus macaques have shown some of the behaviors associated with first-order intentionality in their gestural communication toward human caretakers (Canteloup et al., 2015), and it is plausible that this might extend to call usage: screaming macaques have been reported to scan the area, perhaps monitoring the reactions of potential aiders (“show-looking;” de Waal et al., 1976). In summary, although emotion is clearly involved in many instances of vocal signal production, other cognitive processes appear to be important as well.

Acoustic Variation Within and Between Call Types

Alongside research into the more complex cognitive processes at play in senders, the field has seen recent advances in thinking about the patterning of acoustic variation within and between call types. Researchers have recognized that many species' call types—including vervet monkey alarms—are not totally discrete, stereotyped acoustic categories; they often exhibit within-type variation and grade into one another to some degree, with some vocalizations exhibiting intermediate acoustic structures (Anikin et al., 2018; Fischer et al., 2016; Price et al., 2015; Schwartz et al., 2019). Based on this acoustic variation and gradation, contemporary literature around primate vocal communication portrays call types as clusters of vocalizations within a multidimensional acoustic space, separated from one another by fuzzy rather than hard boundaries (Fischer et al., 2016; Price et al., 2015; Wadewitz et al., 2015). Although this might seem like *prima facie* evidence against categorical representation of call types, the fact that call types are acoustically graded does not suggest that they are not represented as mental categories by receivers and/or senders; the many playback experiments reviewed above suggest categorical representation by receivers, while the question of how senders mentally represent their own call types remains open to investigation.

The recently emerging consensus that most animal call types exhibit acoustic variation has raised the question of, on one hand, the factors involved whether an animal vocalizes and which call type is emitted versus, on the other, the factors determining the precise acoustic structure of that vocalization. This distinction—traditionally posed as “vocal usage” versus “vocal production”, respectively—is not

new (Seyfarth & Cheney, 2010). The novel contribution to this conversation from recent research on the patterning of acoustic variation within and between call types is the recognition that substantial within-type variation and between-type gradations are more ubiquitous across animals than previously thought.

Synthesizing recent advances in sender cognition and in acoustic patterning, some authors have proposed that, in general, the particular acoustic structure of a vocalization (i.e., within-type variation) reflects the sender's emotional arousal whereas call usage might reflect additional cognitive processes (Schamberg et al., 2018; Wadewitz et al., 2015). In support of this, a recent study with common marmosets (*Callithrix jacchus*) showed that within-type acoustic structure was associated with sender heart rate (a reliable measure of arousal), while call type was associated with the presence/absence and distance of a conspecific but not heart rate (Liao et al., 2018). Other studies have demonstrated a relationship between within-type acoustic structure and emotional valence in addition to or instead of arousal (Briefer, 2012; Briefer, Maigrot et al., 2015; Briefer, Vizier et al., 2019; Friel et al., 2019; Maigrot et al., 2018). Many recent studies on the vocal communication of emotion have at least implicitly, and sometimes explicitly, recognized the necessity of assessing the significance of variation within call types (e.g., Briefer, 2018; Briefer, Maigrot et al., 2015; Briefer, Vizier et al., 2019; Faragó et al., 2017; Friel et al., 2019; Linhart et al., 2015; Maigrot et al., 2018; Schwartz & Gouzoules, 2019). In addition to arousal, individual variation serving to convey the identity of the sender probably largely lies within call types, rather than or in addition to patterns of call usage (Rendall, 2003; Rendall et al., 1998).

Acoustic Variation Between and Within Call Types: Semi-Separate Channels?

The distinction between the factors determining call usage versus those shaping within-type acoustic structure represents an important step forward in the field of animal vocal communication and warrants revisiting an old question: the extent to which different kinds of information might be simultaneously available from animal vocalizations. In their 1980 paper on vervet alarms, Seyfarth et al. (1980b) suggested that emotional and external information might be simultaneously accessible from vocalizations, writing, "It ... seems appropriate to interpret arousal-related properties of alarm calls as ancillary to more specific call features, supplementing and enriching the meaning of calls rather than serving as a primary basis for meaning" (p. 1091). Twenty years later, research on meerkats (*Suricata suricatta*) provided direct support for this idea. Like vervet monkeys, meerkats produce distinct alarm calls in response to terrestrial predators, avian predators, and snakes (the snake alarm also being associated with other stimuli), and receivers respond appropriately (Manser, 2001; Manser et al., 2001). In addition to predator type, the alarms vary acoustically depending on the urgency of the threat (e.g., the proximity of the predator), with high-urgency alarms associated with a noisier acoustic structure and low-urgency alarms appearing more tonal (Manser, 2001); receivers respond not only to the predator type indicated by the call but also the urgency, with high-urgency calls eliciting more rapid responses and longer vigilance (Manser et al., 2001). The acoustic parameters important for distinguishing the predator types are different from those that correlate with urgency. Following the assumption that high-urgency situations evoke high-arousal emotional states in senders, these findings were interpreted as evidence that receivers could access information relating to the external predator type and the emotional state of the sender simultaneously, through separate acoustic channels (Manser et al., 2002; Seyfarth & Cheney, 2003a)—though the authors were careful to avoid specific claims about the cognitive processes involved. Later research from this group suggested that these two channels develop on different timescales, with the acoustic features associated with urgency ontogenetically preceding those that correlate with predator type (Hollén & Manser, 2007).

Although it was recognized early on that different acoustic parameters might convey emotional and external information, it was long unclear whether there might be a general pattern by which these two sets of acoustic parameters might be segregated. Recent thinking about acoustic variation within and between call types, and the differences in the cognitive mechanisms underlying call usage and vocal production, point to a speculative but intriguing possibility. Perhaps, like vervet monkeys (Price et al., 2015), meerkats possess acoustically variable, graded, fuzzy, but discriminable alarm call types, each

conveying information about a different category of external phenomena. The usage of these call types could be underlain by a variety of possible cognitive processes including, but not limited to, emotional state and first-order intentionality. Acoustic variation within each of these three call types, in contrast, probably correlates more directly with the emotional arousal of the sender (and therefore the urgency of the threat). If that is the case, perhaps acoustic variation between versus within call types might comprise semi-separate channels functioning to convey information about the predator type, and urgency, respectively.

This hypothesis could apply not only to meerkats but also to many of the species exhibiting externally referential call types. Even in species that have not shown evidence of external reference, it is plausible that acoustic variation between and within call types might convey functionally distinct significant categories of information to receivers. The general trend seems to be that within-type variation correlates with senders' emotional arousal and valence across species, and thus serves as a consistently available source of information about sender emotion and/or, potentially, any external phenomena that correlate reliably with the sender's emotional state. Within-type variation can, as discussed earlier, also convey the identity of the sender, insofar as the acoustic structure of a call varies by individual. The information available from the usage of different call types, in contrast, varies extensively depending on the species and the contexts in which calls occur, and can relate not only to external phenomena but also the probability of subsequent sender behavior (Seyfarth et al., 2010). Thus, the information conveyed by these channels may overlap, hence our designation of them as "semi-separate."

Evolution of Acoustic Variation Between and Within Call Types

The evidence reviewed above warrants a brief discussion of implications for general processes in vocal evolution, although such speculation, of course, requires caution. Specifically, it seems plausible that evolutionary forces might have different roles in shaping acoustic variation within versus between call types. On one hand, acoustic differences between call types seem responsive to selection pressures, or at least capable of rapid evolution, as demonstrated by differences in vocal repertoires across species (Dunn & Smaers, 2018; Fischer et al., 2016; Gouzoules & Gouzoules, 2000; Seyfarth & Cheney, 2018; though see Hammerschmidt & Fischer, 2019). For example, four closely related macaque species (*Macaca spp.*) exhibit clear differences in the acoustic qualities of screams used in the same agonistic context (Gouzoules & Gouzoules, 2000). On the other hand, some emotion-related acoustic variation within call types appears to be tied to the physiological and anatomical mechanisms of vocal production; for example, the correlation between emotional arousal and vocal F0 comes about as a natural consequence of the local effects on the laryngeal muscles of the generalized, body-wide arousal response. A hypothetical evolutionary change in the acoustic correlates of arousal within a call type might, in theory, require a change in either the generalized arousal response itself, or basic vocal anatomy—changes that could conceivably carry additional, negative fitness consequences. In addition, it is plausible that the relationships between emotion and acoustic variation within call types might be under stabilizing selection, due to benefits for senders associated with the function of conveying emotional information (Altenmüller et al., 2013). These pressures alone or in conjunction might have constrained emotion-related within-type acoustic variation over evolutionary time, explaining why the relationships between emotional arousal and within-type acoustic structure appear conserved across such a broad range of taxa (Briefer, 2012; Filippi et al., 2017; Zimmermann et al., 2013). Within-type acoustic variation might therefore comprise a potential source of "honest" information about the sender's emotional state (see Fitch & Hauser, 2006). This hypothesis generates testable predictions: if it is the case, then the acoustic correlates of emotions should be generally consistent not only across species, but also across contexts and call types within species.

From Two Questions to Four

To summarize, the field of animal vocal communication appears to be moving toward a view of acoustic variation between versus within call types as differing with respect to ultimate functions as well as cognitive mechanisms in both senders and receivers. We suggest that this distinction is analogous to an extension of Seyfarth's and Cheney's (2003b) efforts to delineate the boundary between senders and receivers as objects of empirical study. Building on the two broad questions laid out by Seyfarth and Cheney—(1) what are the roles of emotion and other cognitive processes in signal production by senders, and (2) what is the role of information about, and mental representation of, external objects and events in receiver behavior?—we see recent directions in the field as pointing to four questions centering on the factors that contribute to: (1) call usage by senders, (2) the particular within-type acoustic structure of a given call, (3) receivers' responses to different types of calls, and (4) receivers' responses to acoustic variation within call types (Figure 1). All four of these questions can and should be addressed at both the proximate and ultimate levels.

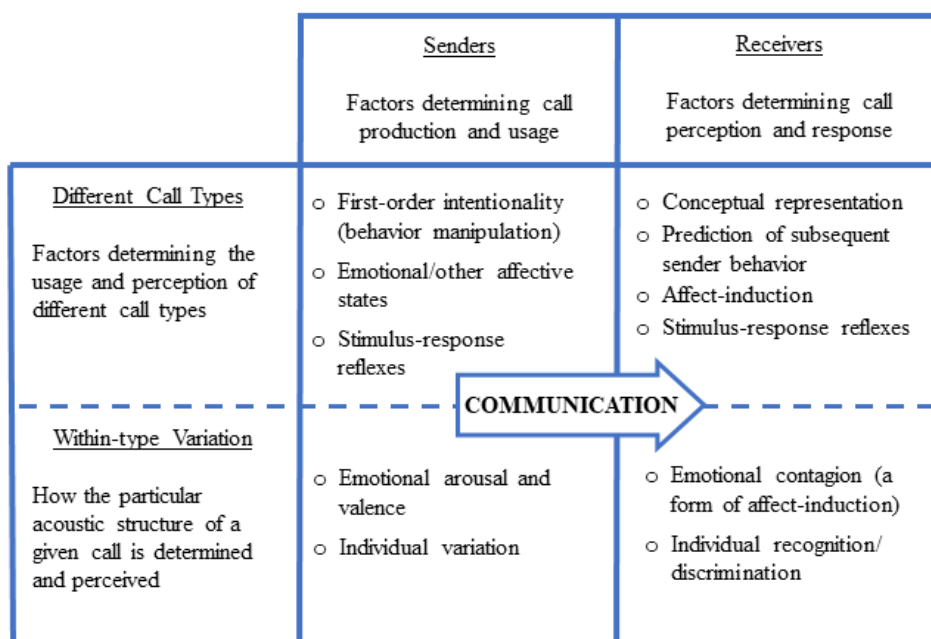


Figure 1. Organization of two distinctions leading to four broad questions. Each quadrant includes a brief, non-exhaustive list of proximate mechanisms supported by the literature. The processes listed are not mutually exclusive.

The evidence reviewed thus far collectively points to some of the ways that the cognitive processes involved in these four categories of behavior within instances of communication might differ (Figure 1). To summarize briefly, on the sender side, the long-held view of emotion as important for vocal usage and production has been borne out by much research, but it is no longer tenable to view all calling behavior as purely emotional or involuntary; first-order intentionality seems likely to play a role in many taxa as well. Often, emotion probably works in tandem with cognitive processes relating to first-order intentionality to influence calling behavior. On the other hand, within-type variation might generally relate more reliably to senders' emotional states, as well as inter-individual differences in vocal anatomy and production mechanisms. On the receiver side, as reviewed above, responses to different call types can be underlain by mental representation, affect-induction, stimulus-responses reflexes, or some combination of these. Receivers probably also represent the identity of the sender, conveyed through within-type acoustic variation (Seyfarth & Cheney, 2015). How receivers might respond to emotional information available from variation within call types is less clear. It is unlikely that most animals understand this information by attributing emotional states to senders (Cheney & Seyfarth, 2008; Seyfarth & Cheney,

2003a, b), but that does not mean that the emotional nature of within-type acoustic variation is irrelevant from the standpoint of receiver cognition. Emotional contagion, the process by which one individual's emotional state spreads to others through automatic perception-action processes (Preston & de Waal, 2002), could allow the emotional arousal of vocalizers to directly affect the emotional arousal of receivers, causing them to respond more urgently (Briefer, 2018). Thus, the affect-induction model (Owren & Rendall, 1997, 2001) might be especially well suited to explain receiver responses to within-type acoustic variation. A non-mutually exclusive possibility is that receivers might use within-type acoustic variation to predict the probable subsequent behavior of the sender, or to mentally represent additional external phenomena of emotional salience to senders, e.g., the proximity of a threat (Manser et al., 2001, 2002).

This recent shift in the framing of animal vocal communication research not only advances thinking about the findings of past research but also holds the potential to generate novel questions. Of the four questions delineated above, receivers' responses to acoustic variation within call types has received the least empirical attention. Few studies have examined whether receivers are even sensitive to emotion-related within-type variation, let alone how, except in humans (e.g., Schwartz & Gouzoules, 2019; Wood et al., 2017). The hypothesis that emotional contagion is involved in this process seems promising as a jumping-off point for future research (Briefer, 2018). Neuroscience research holds the potential to elucidate the brain networks that process emotional information from calls, and how these might differ from those that process externally referential components (Rauschecker, 2013). The other three questions also contain no shortage of interesting areas for future research. Additional topics worthy of further study include the *interactions between* cognitive mechanisms of call usage, perception of different calls, within-type acoustic structures, and perception of within-type variation, and the evolution of these different mechanisms. For example, how do the potential differences in cognitive bases of call usage versus within-type acoustic structure make available different kinds of information to receivers; how have these different kinds of information resulted in different selection pressures shaping the cognitive processes receivers use to perceive each kind of acoustic variation; and are these different kinds of acoustic variation under different evolutionary forces? These comprise promising future directions for the field.

Contributions from Scream Research: Rhesus Macaque Screams as a Case Study

Primate screams seem a promising call type for future research in the four areas outlined above, especially regarding the roles of emotion and other cognitive processes in call production and perception. Screams almost invariably occur in high-arousal contexts and can thus be viewed as highly emotional on the part of senders, yet chimpanzee screams show audience effects that are difficult to explain through sender emotion alone (Slocombe & Zuberbühler, 2007). On the receiver side, screams are inherently emotionally evocative, with human listeners showing emotional neural activation in response to screams and other sounds that share acoustic properties with them (Arnal et al., 2015; Belin & Zatorre, 2015). At the same time, the screams of rhesus macaques, chacma baboons, and chimpanzees appear to evoke mental representations of the identity of the vocalizer, the rank of the opponent, and/or the degree of agonistic intensity experienced (Bergman et al., 2003; Gouzoules, 2005; Gouzoules et al., 1984, 1985; Slocombe et al., 2010), and human listeners perceive acoustic variation among con- and heterospecific screams as containing cues to identity (Engelberg & Gouzoules, 2018; Fugate et al., 2008) and emotional arousal (Schwartz & Gouzoules, 2019). Research across taxa addressing how emotion and other cognitive processes map onto the usage and perception of screams, versus acoustic variation among screams, has the potential to contribute directly to the recent directions reviewed in this paper.

The vocal repertoire of rhesus macaques contains multiple scream classes, with acoustic variation both among and within these classes, and therefore seems like an especially promising model in which to explore these issues. Below, we take the opportunity to further flesh out the organizational framework described above by applying it to rhesus macaque screams, sharing the results of a preliminary case study

beginning to examine how emotional arousal relates to acoustic variation within versus between scream classes in this species.

As reviewed above, rhesus macaque screams function to recruit aid during agonistic interactions, and fall into acoustically distinct but graded scream classes, the usage of which correlates with the rank of the opponent and the intensity of aggression received; receivers respond differently depending on their relationship to the sender and on the scream classes used, suggesting the screams evoke a mental representation of aspects of the interaction playing out (Gouzoules et al., 1984, 1985). Most research into rhesus macaque screams has focused on the usage and perception of, and acoustic variation among, different scream classes, whereas the cognitive processes underlying production and perception of acoustic variation within scream classes has received relatively little attention. Although Gouzoules et al. (1984) challenged the idea that arousal could account entirely for differences among scream classes, they noted the existence of acoustic variation within scream classes (see also Le Prell et al., 2002); following the recent directions laid out within the present paper, it is plausible that arousal plays a greater role in that variation than it does in scream class usage. The question of how this variation relates to individual identity has received attention (Fugate et al., 2008; Rendall et al., 1998), but the relationship between emotional states and acoustic variation within scream classes has remained open.

To investigate this, we (Schwartz, J. W., & Gouzoules, H.) recorded and analyzed the acoustic properties of rhesus macaque screams emitted during naturally occurring social interactions. Detailed methods are available online in the [Supplementary Materials](#). We attributed varying levels of emotional arousal to vocalizers depending on the degree of agonistic intensity received, ranging in intensity from no visible threat, to threats with and without physical contact (e.g., a lunge forward with or without a grab, respectively), to high-intensity chase and attack interactions. We recorded multiple screams of a given scream class from bouts associated with each of these different contexts; often multiple scream classes occurred within a single bout, likely relating to dynamic changes in the nature of an interaction as it played out. We then compared the emotional arousal of a sender, operationally defined according to the intensity of the preceding threat, to the acoustic properties of screams, while accounting for scream class. Under the hypothesis that within-class acoustic variation correlates more strongly with emotional arousal than between-class acoustic variation or scream class usage, we predicted that several acoustic correlates of emotional arousal in mammals—including parameters related to F0, energy distribution, tonality, and scream duration (e.g., Briefer, 2012)—would correlate with agonistic intensity, irrespective of variation between scream classes.

Detailed results are available online in the [Supplementary Materials](#). Based on numbers of screams recorded, we limited analysis to two distinct scream classes: tonal screams ($N = 178$; clearly identifiable F0 contour, and a duration lasting 0.25–1 s) and pulsed screams ($N = 184$; < 0.25 s in duration, ranging in tonality from noisy and cough-like to tonal and chirp-like). Note that tonal and pulsed screams are just two of the scream classes exhibited by rhesus macaques, and might not be representative of screams as a whole in this species, especially since they elicit weaker responses than noisy and arched screams (corresponding to severe aggression and rank challenge, respectively; Gouzoules et al., 1984, 1986). We found a significant main effect of the intensity of the preceding agonism on mean F0 and start F0 of screams irrespective of scream class (Figure 2). Additionally, minimum F0 and the frequency value of the upper limit of the first quartile of energy (dfa25) both correlated with agonism intensity among pulsed screams, showing higher values in high-arousal attack and chase contexts, though not among tonal screams (Figure 2). Insofar as agonistic intensity correlates with the arousal of the sender, these acoustic changes are probably attributable to increased tension in the vocal folds resulting in greater overall F0, and match broader trends across mammals (Briefer, 2012). Thus, although arousal appears unable to account for the acoustic differences among scream classes (Gouzoules et al., 1984), it does correlate significantly with acoustic variation within tonal and pulsed scream classes, in a fashion consistent with other species and with known mechanisms of vocal emotion expression. Furthermore, counter to the notion that arousal underlies acoustic gradations between scream classes, the acoustic parameters that correlated with arousal in the present study (those related to F0) are not those that were originally described as distinguishing the scream classes from one another (Gouzoules et al., 1984); temporal

characteristics and those related to spectral noise would better fit that criterion, yet these did not exhibit significant arousal-dependent variation within our sample.

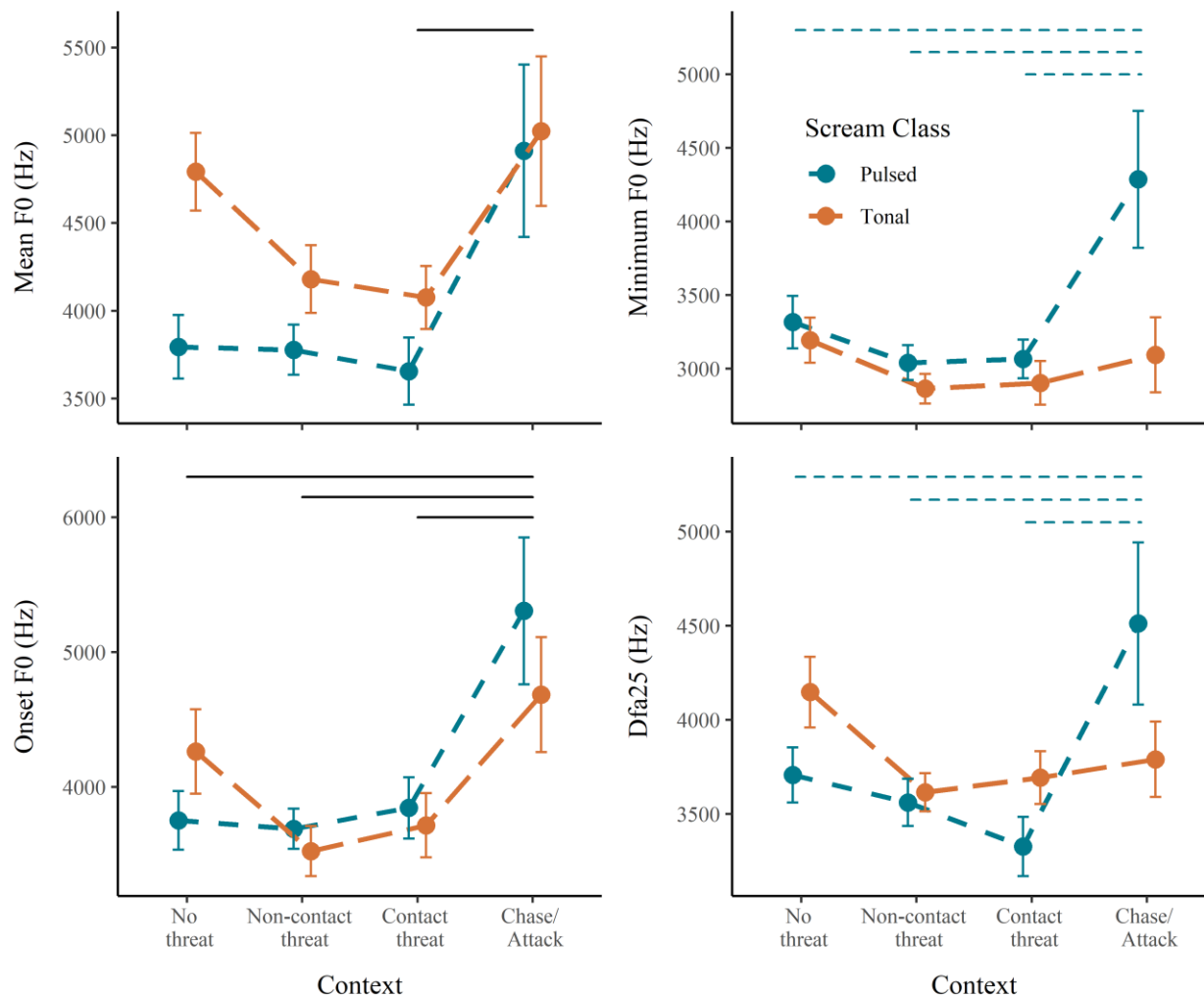


Figure 2. Acoustic variation within scream classes, and relationship to context, with lowest arousal on the left and highest arousal on the right. Horizontal lines denote significant pairwise differences for all screams regardless of class (solid black line) or for pulsed screams only (dotted line).

It is important to reiterate two limitations of this preliminary investigation. First, our conclusions are limited to tonal and pulsed screams, which were originally reported as less significant for receivers (Gouzoules et al., 1984, 1985). Second, it might be overly simplistic to reduce an entire agonistic interaction to a single threat and a single associated arousal level. Agonistic interactions are dynamic, potentially escalating or de-escalating in perceived intensity (and arousal) over the course of a bout of screaming; we accounted for this when escalations were obvious (when a low-intensity threat was followed by a higher-intensity threat; see [online Supplementary Materials](#)), but some changes in senders' arousal states over the course of an interaction probably remain unaccounted for.

Having stated these limitations, the findings of this preliminary case study permit two tentative conclusions with respect to broader issues outlined in this review regarding the cognitive mechanisms of call usage and vocal production, and call function. First, the cognitive processes determining the particular acoustic structure of a tonal or pulsed scream (i.e., within-class variation) appear to differ from those determining scream class usage and between-class acoustic differences, with emotional arousal clearly playing a role in the former but less clearly in the latter. It remains unclear what additional

cognitive mechanisms might be involved in the usage of different scream classes. Previous writing has suggested that senders hold mental representations of the rank differences between themselves and opponents (H. Gouzoules et al., 1998). In addition, we have observed a tendency among screaming rhesus macaques to scan the area, perhaps monitoring the reactions of potential aiders (Schwartz, J. W., & Gouzoules, H., personal observation; for a similar observation in crab-eating macaques, *Macaca fascicularis*, see “show-looking” in de Waal et al., 1976), which if true, would be considered evidence supporting first-order intentionality (though, on its own, insufficient to conclude such, Schel et al., 2013).

Second, the acoustic parameters showing arousal-dependent variation while accounting for scream class (primarily related to F0) seem to be different from those that were originally reported to differentiate scream classes (e.g., duration, tonality, Gouzoules et al., 1984), consistent with the hypothesis that acoustic variation between and within scream classes function as semi-separate channels for different types of information, processed by receivers through different cognitive processes. More direct evidence of this hypothesis is yet lacking, but some tentative speculations are possible. As discussed above, different scream classes convey information about the rank difference between the vocalizer and the opponent and about the intensity of aggression involved; cognitive mechanisms involved in responses to different scream classes probably involve mental representation of kin- and rank-relationships and the details of an incident of aggression. Simultaneously, the present results suggest F0 variation within tonal and pulsed scream classes might convey further information about the intensity of aggression, the internal state of the vocalizer, and/or the sender’s likely subsequent behavior. Given the cross-species emotional significance of vocal F0 (Briefer, 2012) and the emotional salience of F0 for human listeners (Filippi et al., 2017; Kelly et al., 2017; Sauter et al., 2010), including in scream perception (Schwartz & Gouzoules, 2019), it seems plausible that rhesus macaque listeners might respond to F0 variation within these scream classes via emotional contagion (Briefer, 2018), although this has not been tested. Note that F0 is not always perceivable in rhesus macaque screams, especially the class of noisy screams exhibited in the most intense contexts (Gouzoules et al., 1984). Furthermore, the distinction between these two possible categories of information is not strict, since the usage of scream classes depends in part on the same variable we used as a proxy for arousal, (the intensity of agonism received by the vocalizer, Gouzoules et al., 1984), but it might still be significant in terms of biological function and/or cognitive mechanisms. Future research into rhesus macaques’ responses to different scream classes versus to within-class variants, and the role of arousal in the acoustic properties of other scream classes (noisy and arched screams), has the potential to shed light on these questions.

Conclusions

The 40 years since the initial vervet playback studies by Seyfarth et al. (1980a, b) have seen major advances in scientific understanding of animal communication. A critical framework for these advances has been the delineation of senders and receivers and the cognitive processes underlying the behavior of each (Seyfarth & Cheney, 2003a, b). Recently, the field has also begun to distinguish between- and within-type acoustic variation, effectively dividing two broad questions—what factors determine (1) call usage and production by senders, and (2) receivers’ responses to calls?—into four: what factors determine (1) call usage, (2) within-type acoustic structure, (3) receivers’ responses to different call types, and (4) receivers’ responses to within-type acoustic variation? Screams, which are phylogenetically widespread, emotional, and in some cases convey complex information, hold particular promise for addressing contemporary questions in the field of animal communication. Ever-present in this field is the view that animal communication may provide a “window on the minds of animals” (Griffin, 1976), which served as a guiding principle for the influential initial work of Seyfarth et al. (1980a, b), and much research since. This principle will doubtless remain a cornerstone as the field continues to delineate and address novel questions.

Acknowledgments

We thank Drs. Mar M. Sanchez and Mark E. Wilson for their help and oversight in data collection. The rhesus macaque scream research presented here was supported by the National Institutes of Health under award numbers 1R01HD077623 & P51OD011132, and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1343012.

References

- Adolphs, R., & Andler, D. (2018). Investigating emotions as functional states distinct from feelings. *Emotion Review*, *10*, 191–201. <http://dx.doi.org/10.1177/1754073918765662>
- Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, *157*, 187–200. <http://dx.doi.org/10.1016/j.cell.2014.03.003>
- Altenmüller, E., Schmidt, S., & Zimmermann, E. (2013). A cross-taxa concept of emotion in acoustic communication: An ethological perspective. In E. Altenmüller, S. Schmidt & E. Zimmermann (Eds.), *The evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man* (339–363). Oxford University Press.
- Anikin, A., Bååth, R., & Persson, T. (2018). Human non-linguistic vocal repertoire: Call types and their meaning. *Journal of Nonverbal Behavior*, *49*, 53–80. <http://dx.doi.org/10.1007/s10919-017-0267-y>
- Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A. L., & Poeppel, D. (2015). Human screams occupy a privileged niche in the communication soundscape. *Current Biology*, *25*(12), 2051–2056. <http://dx.doi.org/10.1016/j.cub.2015.06.043>
- Bayart, F., Hayashi, K. T., Faull, K. F., Barchas, J. D., & Levine, S. (1990). Influence of maternal proximity on behavioral and physiological responses to separation in infant rhesus monkeys (*Macaca mulatta*). *Behavioral Neuroscience*, *104*, 98–107. <http://dx.doi.org/10.1037/0735-7044.104.1.98>
- Belin, P., & Zatorre, R. J. (2015). Neurobiology: Sounding the alarm. *Current Biology*, *25*, R805–R806. <http://dx.doi.org/10.1016/j.cub.2015.07.027>
- Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, *302*, 1234–1236. <http://dx.doi.org/10.1126/science.1087513>
- Bernstein, I. S., & Ehardt, C. L. (1985). Agonistic aiding: Kinship, rank, age, and sex influences. *American Journal of Primatology*, *8*, 37–52. <http://doi.org/10.1002/ajp.1350080105>
- Berthet, M., Neumann, C., Mesbahi, G., Cäsar, C., & Zuberbühler, K. (2018). Contextual encoding in titi monkey alarm call sequences. *Behavioral Ecology and Sociobiology*, *72*, 1–11. <http://dx.doi.org/10.1007/s00265-017-2424-z>
- Blumstein, D. T., & Chi, Y. Y. (2012). Scared and less noisy: Glucocorticoids are associated with alarm call entropy. *Biology Letters*, *8*, 189–192. <http://dx.doi.org/10.1098/rsbl.2011.0832>
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *7*, 1–20. <http://dx.doi.org/10.1098/rsbl.2011.0832>
- Briefer, E. F. (2018). Vocal contagion of emotions in non-human animals. *Proceedings of the Royal Society B: Biological Sciences*, *285*, 20172783. <http://dx.doi.org/10.1098/rspb.2017.2783>
- Briefer, E. F., Maigrot, A. L., Mandel, R., Freymond, S. B., Bachmann, I., & Hillmann, E. (2015). Segregation of information about emotional arousal and valence in horse whinnies. *Scientific Reports*, *4*, 1–11. <http://dx.doi.org/10.1038/srep09989>
- Briefer, E. F., Tettamanti, F., & McElligott, A. G. (2015). Emotions in goats: Mapping physiological, behavioural and vocal profiles. *Animal Behaviour*, *99*, 131–143. <http://dx.doi.org/10.1016/j.ANBEHAV.2014.11.002>
- Briefer, E. F., Vizier, E., Gyax, L., & Hillmann, E. (2019). Expression of emotional valence in pig closed-mouth grunts: Involvement of both source- and filter-related parameters. *The Journal of the Acoustical Society of America*, *145*, 2895–2908. <http://dx.doi.org/10.1121/1.5100612>
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*, 187–192. <http://dx.doi.org/10.1016/j.tics.2008.02.010>
- Canteloup, C., Bovet, D., & Meunier, H. (2015). Intentional gestural communication and discrimination of human attentional states in rhesus macaques (*Macaca mulatta*). *Animal Cognition*, *18*, 875–883. <http://dx.doi.org/10.1007/s10071-015-0856-2>
- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world*. University of Chicago Press.
- Cheney, D. L., & Seyfarth, R. M. (2008). *Baboon metaphysics: The evolution of a social mind*. University of

- Chicago Press.
- Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *115*, 1974–1979. <http://dx.doi.org/10.1073/pnas.1717572115>
- Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, *22*, 142–146. <http://dx.doi.org/10.1016/J.CUB.2011.11.053>
- Crockford, C., Wittig, R. M., & Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances*, *3*, e1701742. <http://dx.doi.org/10.1126/sciadv.1701742>
- Darwin, C. (1872). The expression of the emotions in man and animals. John Murray. <http://dx.doi.org/10.1037/h0076058>
- de Waal, F. B. M. (2011). What is an animal emotion? *Annals of the New York Academy of Sciences*, *1224*, 191–206. <http://dx.doi.org/10.1111/j.1749-6632.2010.05912.x>
- de Waal, F. B. M., & Luttrell, L. M. (1985). The formal hierarchy of rhesus macaques: An investigation of the bared-teeth display. *American Journal of Primatology*, *9*, 73–85. <http://dx.doi.org/10.1002/ajp.1350090202>
- de Waal, F. B., van Hooff, J. A. R. A. M., & Netto, W. J. (1976). An ethological analysis of types of agonistic interaction in a captive group of Java-monkeys (*Macaca fascicularis*). *Primates*, *17*, 257–290. <http://doi.org/10.1007/BF02382785>
- Dennett, D. C. (1983). Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behavioral and Brain Sciences*, *6*, 343. <http://dx.doi.org/10.1017/S0140525X00016393>
- Deshpande, A., Gupta, S., & Sinha, A. (2018). Intentional communication between wild bonnet macaques and humans. *Scientific Reports*, *8*, 5147 <http://dx.doi.org/10.1038/s41598-018-22928-z>
- Dunn, J. C., & Smaers, J. B. (2018). Neural correlates of vocal repertoire in primates. *Frontiers in Neuroscience*, *12*, 534. <http://dx.doi.org/10.3389/fnins.2018.00534>
- Engelberg, J. W. M., & Gouzoules, H. (2018). The credibility of acted screams: Implications for emotional communication research. *Quarterly Journal of Experimental Psychology*, *72*, 1889–1902. <http://dx.doi.org/10.1177/1747021818816307>
- Engelberg, J. W. M., Schwartz, J. W., & Gouzoules, H. (2019). Do human screams permit individual recognition? *PeerJ*, *7*, e7087. <http://dx.doi.org/10.7717/peerj.7087>
- Faragó, T., Takács, N., Miklósi, Á., & Pongrácz, P. (2017). Dog growls express various contextual and affective content for human listeners. *Royal Society Open Science*, *4*, 170134. <http://dx.doi.org/10.1098/rsos.170134>
- Fichtel, C., Hammerschmidt, K., & Jürgens, U. (2001). On the vocal expression of emotion. A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behavior*, *138*, 97–116. <http://dx.doi.org/10.1163/15685390151067094>
- Fichtel, C., & Manser, M. (2010). Vocal communication in social groups. In P. Kappeler (Ed.), *Animal behaviour: Evolution and mechanisms* (pp. 29–54). Springer Berlin. <http://dx.doi.org/10.1007/978-3-642-02624-9>
- Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: A comparative approach to the emergence of language. *Frontiers in Psychology*, *7*, 1393. <http://dx.doi.org/10.3389/fpsyg.2016.01393>
- Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Pašukonis, A., Hoeschele, M., Ocklenburg, O., de Boer, B., Sturdy, C. B., Newen, A., & Güntürkün, O. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859), 20170990. <http://dx.doi.org/10.1098/rspb.2017.0990>
- Fischer, J. (2017). Primate vocal production and the riddle of language evolution. *Psychonomic Bulletin and Review*, *24*, 72–78. <http://dx.doi.org/10.3758/s13423-016-1076-8>
- Fischer, J., & Price, T. (2016). Meaning, intention, and inference in primate vocal communication. *Neuroscience and Biobehavioral Reviews*, *82*, 22–31. <http://dx.doi.org/10.1016/j.neubiorev.2016.10.014>
- Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2016). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour*, *134*, 229–237. <http://dx.doi.org/10.1016/j.anbehav.2016.06.012>
- Fitch, W. T., & Hauser, M. D. (2006). Unpacking “honesty”: Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (pp. 65–137). Springer. http://dx.doi.org/10.1007/0-387-22762-8_3
- Friel, M., Kunc, H. P., Griffin, K., Asher, L., & Collins, L. M. (2019). Positive and negative contexts predict duration of pig vocalisations. *Scientific Reports*, *9*, 2062. <http://dx.doi.org/10.1038/s41598-019-38514-w>
- Fugate, J. M. B., Gouzoules, H., & Nygaard, L. C. (2008). Recognition of rhesus macaque (*Macaca mulatta*) noisy screams: Evidence from conspecifics and human listeners. *American Journal of Primatology*, *70*, 594–604. <http://dx.doi.org/10.1002/ajp.20533>

- Goodall, J. (1986). *The chimpanzees of Gombe*. Cambridge University Press.
- Gouzoules, H. (2005). Monkeying around with symbolism: Are vocalizations simple symbols...or more like cymbals? In L. Namy (Ed.), *Symbol use and symbolic representation: Developmental and comparative perspectives*. (pp. 245–265). Lawrence Erlbaum Associates Publishers.
- Gouzoules, H., & Gouzoules, S. (2000). Agonistic screams differ among four species of macaques: The significance of motivation-structural rules. *Animal Behaviour*, *59*, 501–512. <http://dx.doi.org/10.1006/anbe.1999.1318>
- Gouzoules, H., & Gouzoules, S. (2006). The conundrum of communication. In C. Campbell, A. Fuentes, K. MacKinnon, M. Panger, & S. Bearder (Eds.), *Primates in perspective* (pp. 621–635). Oxford University Press.
- Gouzoules, H., Gouzoules, S., & Marler, P. (1985). External reference and affective signaling in mammalian vocal communication. In G. Zivin (Ed.), *The development of expressive behavior: Biology environment interactions* (pp. 77–101). Academic Press.
- Gouzoules, H., Gouzoules, S. and Marler, P. (1986). Vocal communication: A vehicle for the study of social relationships. In R. G. Rawlins & M. J. Kessler (Eds), *The Cayo Santiago macaques* (pp. 111-129). State University of New York Press.
- Gouzoules, H., Gouzoules, S., & Tomaszycki, M. (1998). Agonistic screams and the classification of dominance relationships: Are monkeys fuzzy logicians? *Animal Behaviour*, *55*, 51–60. <http://dx.doi.org/10.1006/anbe.1997.0583>
- Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: Representational signalling in the recruitment of agonistic aid. *Animal Behaviour*, *32*, 182–193. [http://dx.doi.org/10.1016/S0003-3472\(84\)80336-X](http://dx.doi.org/10.1016/S0003-3472(84)80336-X)
- Griffin, D. R. (1976). *The question of animal awareness: Evolutionary continuity of mental experience*. Rockefeller University Press.
- Hammerschmidt, K., & Fischer, J. (2019). Baboon vocal repertoires and the evolution of primate vocal diversity. *Journal of Human Evolution*, *126*, 1–13. <http://dx.doi.org/10.1016/j.jhevol.2018.10.010>
- Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). “Worth a thousand words”: Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, *9*, 293–305. <http://dx.doi.org/10.1037/a0015178>
- Hollén, L. I., & Manser, M. B. (2007). Motivation before meaning: Motivational information encoded in meerkat alarm calls develops earlier than referential information. *American Naturalist*, *169*, 758–767. <http://dx.doi.org/10.1086/516719>
- Hopkins, W. D., Tagliatalata, J. P., & Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour*, *73*, 281–286. <http://dx.doi.org/10.1016/j.anbehav.2006.08.004>
- Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice*, *23*, 1–10. <http://dx.doi.org/10.1016/j.jvoice.2007.07.005>
- Juslin, P. N. (2013). Vocal affect expression: Problems and promises. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *The evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man* (pp. 252-273). Oxford University Press.
- Kaplan, J. R. (1978). Fight interference and altruism in rhesus monkeys. *American Journal of Physical Anthropology*, *49*, 241-250.
- Kaplan, J. R., Chikazawa, D. K., & Manuck, S. B. (1987). Aspects of fight interference in free-ranging and compound-dwelling rhesus macaques (*Macaca mulatta*). *American Journal of Primatology*, *12*, 287–298. <http://dx.doi.org/10.1002/ajp.1350120306>
- Kelly, T., Reby, D., Levréro, F., Keenan, S., Gustafsson, E., Koutseff, A., & Mathevon, N. (2017). Adult human perception of distress in the cries of bonobo, chimpanzee, and human infants. *Biological Journal of the Linnean Society*, *120*, 919–930. <http://dx.doi.org/10.1093/biolinnean/blw016>
- Lancaster, J. (1968). Primate communications and the emergence of human language. In P. C. Jay (Ed.), *Primates, studies in adaptation and variability* (pp. 439–457). Holt, Rinehart, and Winston.
- Le Prell, C. G., Hauser, M. D., & Moody, D. B. (2002). Discrete or graded variation within rhesus monkey screams? Psychophysical experiments on classification. *Animal Behaviour*, *63*, 47–62. <http://dx.doi.org/10.1006/anbe.2001.1888>
- Leavens, D. A., Hostetter, A. B., Wesley, M. J., & Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Animal Behaviour*, *67*, 467–476. <http://dx.doi.org/10.1016/j.anbehav.2003.04.007>
- Liao, D. A., Zhang, Y. S., Cai, L. X., & Ghazanfar, A. A. (2018). Internal states and extrinsic factors both determine

- monkey vocal production. *Proceedings of the National Academy of Sciences*, *115*, 3978–3983. <http://dx.doi.org/10.1073/pnas.1722426115>
- Lingle, S., Wyman, M. T., Kotrba, R., Teichroeb, L. J., & Romanow, C. A. (2012). What makes a cry a cry? A review of infant distress vocalizations. *Current Zoology*, *58*, 698–726. <http://dx.doi.org/10.1093/czoolo/58.5.698>
- Linhart, P., Ratcliffe, V. F., Reby, D., & Špinka, M. (2015). Expression of emotional arousal in two different piglet call types. *PLoS ONE*, *10*, 1–13. <http://dx.doi.org/10.1371/journal.pone.0135414>
- Lloyd, L. B. (1989). *Simple minds*. MIT Press.
- Macedonia, J. M., & Evans, C. S. (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. <http://dx.doi.org/10.1111/j.1439-0310.1993.tb00988.x>
- Maigrot, A.-L., Hillmann, E., & Briefer, E. F. (2018). Encoding of emotional valence in wild boar (*Sus scrofa*) calls. *Animals*, *8*, 85. <http://dx.doi.org/10.3390/ani8060085>
- Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*, 2315–2324. <http://dx.doi.org/10.1098/rspb.2001.1773>
- Manser, M. B., Bell, M. B., & Fletcher, L. B. (2001). The information that receivers extract from alarm calls in suricates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*, 2485–2491. <http://dx.doi.org/10.1098/rspb.2001.1772>
- Manser, M. B., Seyfarth, R. M., & Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences*, *6*, 55–57. [http://dx.doi.org/10.1016/S1364-6613\(00\)01840-4](http://dx.doi.org/10.1016/S1364-6613(00)01840-4)
- Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 2895–2904. <http://dx.doi.org/10.1098/rspb.2010.0303>
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, *111*, 855–869. <http://dx.doi.org/10.1086/283219>
- Ordóñez-Gómez, J. D., Santillan-Doherty, A. M., & Hammerschmidt, K. (2019). Acoustic variation of spider monkey (*Ateles geoffroyi*) contact calls is related to caller isolation and affects listeners' responses. *PLOS ONE*, *14*, e0213914. <http://dx.doi.org/10.1371/journal.pone.0213914>
- Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, *73*, 530–544. <http://dx.doi.org/10.1002/ajp.20913>
- Owren, M. J., & Rendall, D. (1997). An affect-conditioning model of nonhuman primate vocal signaling. In D. H. Owings, M. D. Beecher, & N. S. Thompson (Eds.), *Perspectives in ethology* (Vol. 12) (pp. 299–346). Penum Press.
- Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, *10*, 58–71. <http://dx.doi.org/10.1002/evan.1014>
- Panksepp, J. (2011). The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience and Biobehavioral Reviews*, *35*, 1791–1804. <http://dx.doi.org/10.1016/j.neubiorev.2011.08.003>
- Paul, E. S., & Mendl, M. T. (2018). Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. *Applied Animal Behaviour Science*, *205*, 202–209. <http://dx.doi.org/10.1016/j.applanim.2018.01.008>
- Premack, D. (1972). Concordant preferences as a precondition for affective but not for symbolic communication (or how to do experimental anthropology). *Cognition*, *1*, 251–264. [http://dx.doi.org/10.1016/0010-0277\(72\)90021-2](http://dx.doi.org/10.1016/0010-0277(72)90021-2)
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 1–20. <http://dx.doi.org/10.1017/S0140525X02000018>
- Price, T., & Fischer, J. (2014). Meaning attribution in the West African green monkey: Influence of call type and context. *Animal Cognition*, *17*, 277–286. <http://dx.doi.org/10.1007/s10071-013-0660-9>
- Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., & Fischer, J. (2015). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Scientific Reports*, *5*, 1–11. <http://dx.doi.org/10.1038/srep13220>
- Puppe, B., Schön, P. C., Tuchscherer, A., & Manteuffel, G. (2005). Castration-induced vocalisation in domestic piglets, *Sus scrofa*: Complex and specific alterations of the vocal quality. *Applied Animal Behaviour*

- Science*, 95, 67–78. <http://dx.doi.org/10.1016/j.applanim.2005.05.001>
- Rauschecker, J. P. (2013). Brain networks for the encoding of emotions in communication sounds of human and nonhuman primates. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *The evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man* (pp. 49–60). Oxford University Press.
- Rendall, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *The Journal of the Acoustical Society of America*, 113, 3390–3402. <http://dx.doi.org/10.1121/1.1568942>
- Rendall, D., Owren, M. J., & Rodman, P. S. (1998). The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *The Journal of the Acoustical Society of America*, 103, 602–614. <http://dx.doi.org/10.1121/1.421104>
- Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, 78, 233–240. <http://dx.doi.org/10.1016/j.anbehav.2009.06.007>
- Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, 51, 1007–1015. <http://dx.doi.org/10.1006/anbe.1996.0103>
- Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *Quarterly Journal of Experimental Psychology*, 63, 2251–2272. <http://dx.doi.org/10.1080/17470211003721642>
- Scarantino, A. (2013). Rethinking functional reference. *Philosophy of Science*, 80, 1006–1018. <http://dx.doi.org/10.1086/673900>
- Scarantino, A., & Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour*, 100, e1–8. <http://dx.doi.org/10.1016/j.anbehav.2014.08.017>
- Schamberg, I., Wittig, R. M., & Crockford, C. (2018). Call type signals caller goal: A new take on ultimate and proximate influences in vocal production. *Biological Reviews*, 93, 2071–2082. <http://dx.doi.org/10.1111/brv.12437>
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE*, 8, e76674. <http://dx.doi.org/10.1371/journal.pone.0076674>
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256. [http://dx.doi.org/10.1016/S0167-6393\(02\)00084-5](http://dx.doi.org/10.1016/S0167-6393(02)00084-5)
- Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication*, 40, 99–116. [http://dx.doi.org/10.1016/S0167-6393\(02\)00078-X](http://dx.doi.org/10.1016/S0167-6393(02)00078-X)
- Schwartz, J. W., Engelberg, J. W. M., & Gouzoules, H. (2019). Was that a scream? Listener agreement and major distinguishing acoustic features. *Journal of Nonverbal Behavior*. <http://doi.org/10.1007/s10919-019-00325-y>
- Schwartz, J. W., & Gouzoules, H. (2019). Decoding human screams: Perception of emotional arousal from pitch and duration. *Behaviour*, 156, 1283–1307. <http://dx.doi.org/10.1163/1568539X-00003566>
- Scott-Phillips, T. C. (2015a). Meaning in animal and human communication. *Animal Cognition*, 18, 801–805. <http://dx.doi.org/10.1007/s10071-015-0845-5>
- Scott-Phillips, T. C. (2015b). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology*, 56, 56–80. <http://dx.doi.org/10.1086/679674>
- Seyfarth, R. M., & Cheney, D. L. (2003a). Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Sciences*, 1000, 32–55. <http://dx.doi.org/10.1196/annals.1280.004>
- Seyfarth, R. M., & Cheney, D. L. (2003b). Signalers and receivers in animal communication. *Annual Review of Psychology*, 54, 145–173. <http://dx.doi.org/10.1146/annurev.psych.54.101601.145121>
- Seyfarth, R. M., & Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain and Language*, 115, 92–100. <http://dx.doi.org/10.1016/j.bandl.2009.10.003>
- Seyfarth, R. M., & Cheney, D. L. (2012). Animal cognition: Chimpanzee alarm calls depend on what others know. *Current Biology*, 22, R51–R52. <http://dx.doi.org/10.1016/j.cub.2011.11.050>
- Seyfarth, R. M., & Cheney, D. L. (2015). The evolution of concepts about agents: Or, what do animals recognize when they recognize an individual? In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 57–75). The MIT Press.
- Seyfarth, R. M., & Cheney, D. L. (2017). The origin of meaning in animal signals. *Animal Behaviour*, 124, 339–346. <http://dx.doi.org/10.1016/j.anbehav.2016.05.020>
- Seyfarth, R. M., & Cheney, D. (2018). Pragmatic flexibility in primate vocal production. *Current Opinion in Behavioral Sciences*, 21, 56–61. <http://dx.doi.org/10.1016/J.COBEHA.2018.02.005>

- Seyfarth, R. M., Cheney, D. L., & Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends in Cognitive Sciences*, 9, 264–266. <http://dx.doi.org/10.1016/j.tics.2005.04.001>
- Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, 80, 3–8. <http://dx.doi.org/10.1016/j.anbehav.2010.04.012>
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980a). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210, 801–803. <http://dx.doi.org/10.1126/science.7433999>
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980b). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, 28, 1070–1094. [http://dx.doi.org/10.1016/S0003-3472\(80\)80097-2](http://dx.doi.org/10.1016/S0003-3472(80)80097-2)
- Shettleworth, S. J. (2010). *Cognition, evolution, and behavior*. Oxford University Press. [http://doi.org/10.1016/S1364-6613\(99\)01409-6](http://doi.org/10.1016/S1364-6613(99)01409-6)
- Siebert, K., Langbein, J., Schön, P. C., Tuchscherer, A., & Puppe, B. (2011). Degree of social isolation affects behavioural and vocal response patterns in dwarf goats (*Capra hircus*). *Applied Animal Behaviour Science*, 131, 53–62. <http://dx.doi.org/10.1016/j.applanim.2011.01.003>
- Sievers, C., & Gruber, T. (2016). Reference in human and non-human primate communication: What does it take to refer? *Animal Cognition*, 19, 759–768. <http://dx.doi.org/10.1007/s10071-016-0974-5>
- Slocombe, K. E., Kaller, T., Call, J., & Zuberbühler, K. (2010). Chimpanzees extract social information from agonistic screams. *PLoS ONE*, 5, e11473. <http://dx.doi.org/10.1371/journal.pone.0011473>
- Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, 104, 17228–17233. <http://dx.doi.org/10.1073/pnas.0706741104>
- Stoeger, A. S., Charlton, B. D., Kratochvil, H., & Fitch, W. T. (2011). Vocal cues indicate level of arousal in infant African elephant roars. *The Journal of the Acoustical Society of America*, 130, 1700–1710. <http://dx.doi.org/10.1121/1.3605538>
- Struhsaker, T. T. (1967). Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In S. A. Altmann (Ed.), *Social communication among primates* (pp. 281–324). University of Chicago Press.
- Sugiura, H. (2007). Effects of proximity and behavioral context on acoustic variation in the coo calls of Japanese macaques. *American Journal of Primatology*, 69, 1412–1424. <http://dx.doi.org/10.1002/ajp.20447>
- Suzuki, T. N. (2018). Alarm calls evoke a visual search image of a predator in birds. *Proceedings of the National Academy of Sciences*, 115, 1541–1545. <http://doi.org/10.1073/pnas.1718884115>
- Szipl, G., Ringler, E., Spreafico, M., & Bugnyar, T. (2017). Calls during agonistic interactions vary with arousal and raise audience attention in ravens. *Frontiers in Zoology*, 14, 1–13. <http://dx.doi.org/10.1186/s12983-017-0244-7>
- Taylor, A. M., Reby, D., & McComb, K. (2009). Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethology*, 115, 905–915. <http://dx.doi.org/10.1111/j.1439-0310.2009.01681.x>
- Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., Goncalves, I. B., Burkart, J. M., Flower, T., Gaunet, F., Glock, H. J., Gruber, T., Jansen, D. A. W. A. M., Liebal, K., Linke, A., Miklósi, Á, Moore, R., van Schaik, C. P., Stoll, S., ... Manser, M. B. (2017). Exorcising Grice's ghost: An empirical approach to studying intentional communication in animals. *Biological Reviews*, 92, 1427–1433. <http://dx.doi.org/10.1111/brv.12289>
- Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: The past, present and the future. *Ethology*, 119, 1–11. <http://dx.doi.org/10.1111/eth.12015>
- Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015). Characterizing vocal repertoires - Hard vs. soft classification approaches. *PLoS ONE*, 10, e0125785. <http://dx.doi.org/10.1371/journal.pone.0125785>
- Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology*, 21, 195–205. <http://dx.doi.org/10.1002/evan.21319>
- Wheeler, B. C., & Fischer, J. (2015). The blurred boundaries of functional reference: A response to Scarantino & Clay. *Animal Behaviour*, 100, 9–13. <http://dx.doi.org/10.1016/j.anbehav.2014.11.007>
- Wood, A., Martin, J., & Niedenthal, P. (2017). Towards a social functional account of laughter: Acoustic features convey reward, affiliation, and dominance. *PLoS ONE*, 12, e0183811. <http://dx.doi.org/10.1371/journal.pone.0183811>
- Yamaguchi, C., Izumi, A., & Nakamura, K. (2010). Time course of vocal modulation during isolation in common

- marmosets (*Callithrix jacchus*). *American Journal of Primatology*, 72, 681–688. <http://dx.doi.org/10.1002/ajp.20824>
- Zimmermann, E., Leliveld, L. M. C., & Schehka, S. (2013). Toward the evolutionary roots of affective prosody in human acoustic communication: A comparative approach to mammalian voices. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of emotional communication: From sounds in nonhuman mammals to speech and music in men* (pp. 117–132). Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199583560.001.0001>
- Zuberbühler, K. (2003). Referential signaling in non-human primates: Cognitive precursors and limitations for the evolution of language. *Advances in the Study of Behavior*, 33, 265–307. [http://dx.doi.org/10.1016/S0065-3454\(03\)33006-2](http://dx.doi.org/10.1016/S0065-3454(03)33006-2)
- Zuberbühler, K., Cheney, D. L., & Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology*, 113, 33–42. <http://dx.doi.org/10.1037/0735-7036.113.1.33>