Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________________________ _______________

Jonathan W. M. Engelberg Date

Listeners cannot discriminate acted from natural human screams

By

Jonathan W. M. Engelberg

Master of Arts

Psychology

Harold Gouzoules

Advisor

Jocelyne Bachevalier

Committee Member

Lynne Nygaard

Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Listeners cannot discriminate acted from natural human screams

By

Jonathan W. M. Engelberg
B.A., Johns Hopkins University, 2013

Advisor: Harold Gouzoules, Ph.D.

An abstract of a thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2017

Abstract

Listeners cannot discriminate acted from natural human screams

By Jonathan W. M. Engelberg

Are people able to distinguish naturally occurring screams from acted ones? Some authors have suggested that aspects of expression are difficult or impossible to reproduce voluntarily without the concomitant internal state, and that acted expressions are therefore likely to amount to overly-intense caricatures; thus it remains an empirical question the extent to which acted expressions are in fact representative of those that occur naturally. Within this context, and evolutionarily, screams are of particular interest. We used a forced-choice discrimination paradigm to test whether acted and natural screams are perceptibly distinct. The results indicated that listeners could not make this discrimination overall, suggesting that acted and natural screams do not differ significantly in acoustic form. Generalized linear mixed models on listeners' intensity ratings of screams revealed that acted exemplars were not perceived as more intense than natural exemplars. However, scream duration predicted both the likelihood that an exemplar was identified as acted and the likelihood that participants discriminated that scream accurately. These findings are promising with respect to the external validity of studies using acted screams, but future work should determine whether longer screams, or other presentations conveying added information (e.g., multimodal displays) are more easily recognized as acted or natural.

Listeners cannot discriminate acted from natural human screams

By

Jonathan W. M. Engelberg
B.A., Johns Hopkins University, 2013

Advisor: Harold Gouzoules, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2017

Table of Contents

Figures and Tables

Introduction

In his seminal text, *The Expression of Emotions in Man and Animals,* Darwin used acted emotional portrayals (along with electrically galvanized expressions) to explore the functional origins of nonverbal, communicative displays in humans (Darwin, 1872/1998). Noting, for example, the furrowed brows in an actor's simulation of grief, he posited that the visage took its ancestral form to protect the expresser's eyes while crying or screaming; it had since through habit become the facial gesture by which all genuine distress is made apparent. Thus it was that Darwin contemplated most emotional expressions as adaptive in design, automatic in production, and veridical in relation to their internal states (Owren & Bachorowski, 2003).

The modern study of emotional expression owes its foundations to Darwin's work (Russell, Bachorowski, & Fernandez-Dols, 2003). Researchers have embraced not only his ideas—adopting and modifying evolutionary frameworks to interpret patterns of expressive behavior (Tomkins, 1962; Ekman, 1992; Bachorowski & Owren, 2001)—but also elements of his methodology, including the use of acted exemplars (e.g., Ekman & Friesen, 1978; Banse & Scherer, 1996; Hammerschmidt & Jürgens, 2007). In fact, acted material now likely comprises the most commonly researched form of emotional expression (Juslin & Scherer, 2005; Briefer, 2012). It is not surprising, then, that some authors have questioned the basic assumption upon which this research is justified: that these reenacted emotional displays are representative of those that occur naturally (Douglas-Cowie, Campbell, Cowie, & Roach, 2003; Wagner, Trouvain, & Zimmerer, 2015).

In this paper, we will consider the validity of acted material as applied to research on emotional expressions. This is not simply a question of research design, but rather an empirical issue in its own right, invoking discussion as to the nature of emotions, how they are best conceptualized, and how selection has operated on their modes of expression. While at times we will appeal to the literature on facial affect—where the subject of authenticity has been more thoroughly reviewed (Gunnery & Hall, 2015)—we will focus our comments on the vocal modality, affording special attention to experiments that have directly juxtaposed acted vocal expressions with those naturally-produced. Finally, we will present the findings of an original research study, wherein human screams were used as a novel class of vocalization to test for any consistent and perceptible differences between acted and natural expressions of emotion.

Notes on terminology and scope

While *emotional expression*, or more simply *expression*, is the term we will most commonly use to refer to outward representations of internal states (whether or not these representations are honest), we will sometimes refer to them as *signals* or *displays*, aligning ourselves more broadly with general communication theory per the suggestions of Russell et al. (2003). These latter terms assume that expressions function in part to alter the behavior of other individuals (Maynard Smith & Harper, 1995). Additionally, the terms *signaler* and *receiver* may refer, respectively, to an individual producing an expression and to the individual(s) whose behavior it affects (Seyfarth & Cheney, 2003).

The term *acted* will refer to any volitional attempt to reproduce the form of an emotional expression. *Natural* will refer to any expression that occurs without explicit prompting, including those evoked in-lab through procedures designed to induce an emotional state (e.g., Velten, 1968). This is not a perfect dichotomy, for reasons we will examine later in this paper, but it is adopted to reflect the comparison that is made in our study and in similar studies of this nature (e.g., Jürgens, Grass, Drolet, & Fischer, 2015). Closely related to these terms is the idea of authenticity—whether or not an expression co-occurs with a genuine emotional experience, and thereby serves to indicate the presence of that internal state. It is often assumed that expressions produced volitionally are also inauthentic, or conversely, that natural expressions derive from truly-felt emotions (Ekman, Davidson, & Friesen, 1990). Although we suspect this is partially true in our experiment, the relationship between an expression's authenticity and its naturalness is a complicated one (Banse & Scherer, 1996), and caution is ever warranted in extrapolating between the two.

We will discuss the viability of acted material from a theoretical point of view; that is, we will ask whether humans possess the capacity to reproduce expressions of emotion at will, or whether certain physical and psychological constraints act to preclude such replications. This is not to say that every concern regarding the predominance of acted material could be addressed through an absence of inherent limitations. Many authors criticize the use of acted tokens because of methodological issues related to their acquisition, including a tendency to compartmentalize emotions while overlooking emotional blends (e.g., by requesting only basic emotional displays; Greasley, Sherrard, & Waterman, 2000; Cowie & Cornelius, 2003); to provide actors little to no establishing context (Greasley et al., 2000; Owren & Bachorowski, 2003); to obtain expressions produced solitarily rather than interactively (Douglas-Cowie et al., 2003); and, in the case of emotional prosody, to record actors reading from a script where reading itself can have effects on the melodic contour of speech (Douglas-Cowie et al., 2003). These concerns, while legitimate, are not intrinsic to acting *per se*; they can and have been mitigated, for example, by recording stimuli preceded by a short scenario or vignette (Banse & Scherer, 1996). Thus, the principal question that remains is whether actors, provided ideal conditions, are nonetheless unable to produce ecologically valid versions of natural expressions.

Background

It is notoriously difficult, among psychologists of emotion, to find consensus in regards to the object of their study (Scherer, 2003). Most researchers agree that emotions are at least in part biological responses to events internal or external to an organism, involving the synchronous organization of multiple, interrelated subsystems responsible for driving physiological changes as well as the subjective experience (Scherer, 2003; Russell et al., 2003; Drolet, Schubotz, & Fischer, 2014). Beyond this broad definition, however, there persist competing and often incompatible models to account for variation in emotional phenomena.

One well-known approach is that of *basic emotion theory*, championed prominently by authors like Ekman (1992) and Izard (1992). These researchers, continuing in Darwin's (1872/1998) functionalist traditions, contend that each emotion comprises a distinct and basally-adapted set of responses to some fundamental selective pressure (Izard, 1992). Ekman (1992) specifically postulated the existence of five, discrete emotional states—including joy, fear, anger, sadness, and disgust—each entailing its own, instinctively-determined pattern of motor, autonomic, and cognitive activity.

Facial and vocal expressions of emotion are, in this view, products of the same neurophysiological programs that regulate the rest of the emotional episode; that is, they are innate, and in large part involuntary (Hess & Thibault, 2009). As a corollary, any attempt to generate an expression without engaging the full, coordinated suite—effecting a smile, for example, absent the sensation of joy—will result in an expressive form differing from that of the genuine display (Ekman et al., 1990). These ideas are well-illustrated in the literature on *Duchenne smiles*, named for neurologist Duchenne de Boulonge, who described actions of the orbicularis and zygomatic muscles that could be achieved in combination only with "true feeling” (Duchenne, 1862/1990). Thus, it is maintained that this activity in smiles acts as an indicator of authenticity, whereas its omission belies an expression produced for social or strategic purposes, unlinked to a particular internal state (Ekman et al., 1990). Indeed, evidence from perceptual studies suggests that smiles are seen as more natural and authentic when the Duchenne markers have been mobilized (reviewed in Gunnery & Hall, 2015).

In the vocal domain, it is similarly argued that acted and natural expressions are mechanistically distinct, and therefore distinct in acoustic configuration. Historically, researchers have demarcated two types of vocal expressive signaling in humans, each somewhat removed in function and subserved by partially dissociable neuro-productive pathways. *Emotional* (Marty, 1908) or *raw* (Hawk, Van Kleef, Fischer, Van der Schalk, 2009) expressions are characterized as affective leakages of sorts: they are concomitants of emotion, impulsive, unconventionalized, and autonomically impelled (Hawk et al., 2009). In basic emotion theory, these constitute the elemental, inherited patterns of expression: the Duchenne smile is their facial equivalent (Gervais & Wilson, 2005). There is evidence that these vocalizations are not initiated cortically but rather in limbic and mesencephalic regions, which together organize a reflex-like response through the eventual innervation of laryngeal muscles (Ruch & Ekman, 2001; Gervais & Wilson, 2005; Ludlow, 2005; Wattendorf et al., 2013). Thus, bilateral lesions of motor and premotor cortices are observed to spare production of these affect-driven utterances (Jürgens, 2009).

Emotive (Marty, 1908) or *emblematic* (Hawk et al., 2009) expressions, on the other hand, are considered by some almost word-like in nature, symbolic of emotion as opposed to symptomatic (Barrett, 2011). These vocalizations may resemble their emotional counterparts in form—e.g., strategically-produced laughter might mimic the respirational patterns of affective laughs—but they are volitionally produced, often involving finer, cortical control of the articulators, and are not obligately linked to any specific emotional states (Ruch & Ekman, 2001). Importantly, it is argued that emotive signals, and these signals alone, are the only vocal expressions of emotion that are furnished on command; that is, the circuits implicated in raw, emotional displays are not fully accessible to voluntary manipulation (Ruch & Ekman, 2001; Wattendorf et al., 2013). This segregation of pathways, in summation with the indirect effects of different emotions on the vocal apparatus (e.g., in fear, a decrease in salivation and an increase in respiratory activity; Scherer, 1986) could in principle lead to acoustic discrepancies between natural and reenacted expressions of the voice.

Many concerns relating to the use of acted material in empirical research are founded in these kinds of proximate explanations. To the extent that emotions and their expressions are beyond the jurisdiction of volitional control (Ekman et al., 1990), researchers fear that acted imitations might differ in structure from natural expressions (Auberge & Cathiard, 2003; Audibert, Auberge, & Rilliard, 2008; Wagner et al., 2015). One recurring critique is that acted portrayals are likely to be overly intense or exaggerated (Brown & Bradshaw, 1985; Greasley et al., 2000; Batliner et al., 2000; Douglas-Cowie et al., 2003; Barrett, 2011), yielding expressions that are easily discriminated relative to their naturally occurring counterparts (Russell et al., 2003; Owren & Bachorowski, 2003). Other authors have suggested that the differences are subtler and less systematic—deriving from whichever aspects of each emotion are those most difficult to reproduce (Jürgens, Hammerschmidt, & Fischer, 2011)—but conceivably detectable all the same (Auberge & Cathiard, 2003), and resulting in dissimilar responses from receivers.

Regardless of specifics, these arguments are staked in shared implications of ultimate causality. Evolutionary biologists have written at length about the necessary conditions for stable communication (i.e., communication that will endure through evolutionary time; Maynard Smith

& Harper, 1995). One theme is that a signal will not persist in a species unless, on the whole, it is reliable with respect to its internal (or external) referents; otherwise, selection would act on its receivers to ignore the dubious content made available (Gouzoules & Gouzoules, 2007). Assuming it is in a signaler's best interests to deceive (which certainly is not always the case; Maynard Smith & Harper, 1995), some selective principle must ensure that signals are not regularly misemployed. Of relevance here, Zahavi (1975) submitted that *handicap costs* are sustained in the production and transmission of honest signals, such that only individuals in possession of a fitness-related quality are capable of making that quality known. The elongated tail feathers of the male, scarlet-tufted malachite sunbird (*Nectarinia johnstoni*), for example, are so aerodynamically inefficient as to suggest that only birds in good health are able to bear the costs of maintaining them (Evans & Hatchwell, 1992). A similar concept is that of *index signals,* which are reliable because they are physically tied to a quality of interest (Maynard Smith & Harper, 1995); if the frequency range of a call is constrained by the length of the vocal cords, then frequency components may constitute direct indicators of an animal's size (Davies & Halliday, 1978; Vehrencamp, 2000).

The literature on human expressions of emotion has periodically alluded to both of these accounts in discussing the reliability of emotional displays. To wit, researchers have claimed that the production of natural expressions imposes certain costs (Mehu & Scherer, 2012) or else is linked compulsorily to underlying states (Rush & Ekman, 2001; Wattendorf et al., 2013), rendering them difficult or impossible to imitate faithfully. In either case, the implication is that these mechanisms must preserve discernable differences between acted and natural expressions, lest receivers evolve to stop responding altogether, and the entire system of emotional communication falls apart (Dezecache, Mercier, & Scott-Phillips, 2013).

The evolutionary story presented above is not, however, without its problems. First, it remains unclear what costs are actually entailed in substantiating natural expressions; even the canonical Duchenne markers may not prove as challenging to activate as previously thought (Krumhuber & Manstead, 2009), which by definition would suggest that they are not so generatively expensive after all. Second, and more importantly, it does not make sense that selection would shape a signal, nor a signaler's behavior, to the sole advantage of those receiving it (Seyfarth & Cheney, 2003). As Hinde (1981) and Fridlund (1991) among others have noted, signalers should not as a rule convey more information than that which is beneficial to themselves; automatic read-outs of emotion are if anything selected against (Smith, 1997). In fact, if we accept that duplicity is potentially adaptive—hence the need for accounts like costly signaling in the first place—then selection should favor individuals whose dishonest expressions closely resemble their genuine equivalents, the better to deceive effectively and reap whatever benefits are thereby conferred (Dezecache et al., 2013). Of course, receivers in turn should acquire ever more discriminating perceptual acuities to circumvent these attempts at dishonest signaling (Seyfarth & Cheney, 2003). What transpires, then, is an intraspecific coevolution not unlike the proverbial arms race that ensues between a predator and its camouflaging prey (Dawkins & Krebs, 1979; Bryant & Aktipis, 2014). In this dynamic framework, the question of emotional expressions and their relationship to internal states demands reexamination.

Some theories of emotion have deemphasized the link between affect and expression in favor of rather more diversified bases of behavior. Dimensional (Wundt, 1905; Russell & Mehrabian, 1977; Russell & Bullock, 1985) and appraisal processing (Scherer, 1986, 2003) accounts view emotions as determined at the confluence of several, continuously-varying internal scales, the most commonly identified of which include arousal, valence, and sometimes potency or control (Goudbeek & Scherer, 2010). Different emotions map onto different points within the same multidimensional space—such that anger, for example, is low in valence but high in arousal and control (Fontaine, Scherer, Roesch, & Ellsworth, 2007)—and hence are highly interrelated rather than discrete (Russell & Bullock, 1985).

These approaches, then, do not conceptualize expressions as fixed patterns emerging from basic modal states, but as nonspecific outcomes of underlying fluctuations across the set of shared affective axes (Russell, 2003). Further variation of expressive forms, it is proposed, reflects a considerable, top-down weighting of contextual factors, strategic goals, and sociocultural norms and expectations (Scherer & Banziger, 2010). Indeed, some advocates of these models have suggested abandoning altogether the conceit of expressions as simple windows into the inner milieu, emphasizing instead how they function in relation to social solicitation, influence, or appeal (Bachorowski & Owren, 2001; Russell et al., 2003; Jose-Miguel, Fernandez-Dols, & Crivelli, 2013). Scherer (1994, 2013) likewise rejected Marty's (1908) binary distinction of emotive as opposed to emotional expressions, on the basis that almost all expressions are resultant of combined, psychobiological sources, contextual modulators, and higher-level motivations.

The distinction between acted and natural expressions is made less obvious in turn. Insofar as expressions are consistently governed by some proportion of top-down effects (Scherer, 1994), it could be argued that perfectly non-acted prototypes rarely occur in our communicative landscape (Banse & Scherer, 1996). By a similar token, few acted expressions are likely to be entirely void of an honest, emotional contingent (Banse & Scherer, 1996; Scherer, 2013); popular techniques like the Stanislavski method (Moore, 1960) are built upon the self-induction of emotion as a means of delivering believable performances (and many studies have affirmed that self-induction is at least partially effective in marshaling the physiological correlates of emotion; Rein, Atkinson, & McCraty, 1995; Rainville, Bechara, Naqvi, & Damasio, 2006). Therefore it becomes difficult to identify any particular aspect of expression that is strictly linked to natural or to acted antecedents; rather, the overlap in causative processes would seem to suggest that the two expressive categories are more similar than not.

What, then, prevents unreliable signaling in a system where signals are readily recreated? It is possible that costs are not encompassed in the signals themselves but in the interpersonal consequences of deceptive communication (Gouzoules & Gouzoules, 2007). Humans are a species adapted to life in an intricate social meshwork, entailing repeated engagements with the same individuals over extended periods of time (Tibbetts & Dale, 2007). Much has been written of our capacity to predicate behavioral decisions on remembered outcomes of previous encounters (Nowak & Sigmund, 1998; Fehr & Gächter, 2000). A signaler who is revealed to misrepresent his states or intentions is one whom receivers are more likely to ignore in subsequent interactions (Dezecache et al., 2013). Already, we see in language a communicative schema where such socially-imposed costs play an important role in preserving the veracity of otherwise arbitrary signal-meaning associations (Lachmann, Szamado, & Bergstrom, 2001). If the same applies to emotional expressions, then we need not assume that natural displays are inherently irreproducible, especially when the risks of social punishment or neglect are effectively eliminated, for example, in the context of a film shoot or a laboratory study (where, in point of fact, faithful reproductions might well be encouraged).

Indeed, no single profile of naturalness nor the lack thereof has revealed itself in research devoted to finding these cues in vocal emotional expressions. Studies of speech-embedded prosody have focused primarily on fundamental frequency (F0)-related parameters, which correspond perceptually to pitch and its dynamic contour (Gerhard, 2003). Several authors have found that F0 variability is increased in acted relative to natural prosody (Williams & Stevens, 1972; Audibert, Auberge, & Rilliard, 2010; Jürgens et al., 2011); the former might also possess higher values of F0 in general (Williams & Stevens, 1972; Auberge, Audibert, & Rilliard, 2004; Audibert et al, 2010). These characteristics have been associated with elevated arousal levels (Briefer, 2012), lending modest support to the position that acted expressions are artificially intense (Greasley et al., 2000). However, Scherer (2013) failed to report the same F0-related contrasts, while Auberge et al. (2004) in fact described reduced F0 modulation in an acted condition. Jürgens and colleagues (2011, 2015) also noted that, F0 variability aside, acted prosody lacked any of the other hallmarks of heightened arousal; rather, its differences from natural prosody seemed to stem from minor aberrations across the full harmonic spectrum (e.g., more dominant lower harmonics in acted expressions).

A handful of studies have investigated markers of naturalness in nonverbal vocal expressions, and the majority of these have concentrated exclusively on laughter (Anikin & Lima, 2016). Interestingly, the prevailing trend here is that volitional laughter is lower in pitch (McGettigan et al., 2013; Bryant & Aktipis, 2014; Lavan, Scott, & McGettigan, 2016) and pitch variability (Bryant & Aktipis, 2014; Anikin & Lima, 2016) than laughter that is spontaneously evoked. It is not clear, however, that all of the acted laughs produced in these studies were meant to represent perfect imitations of the tickle- or stimulus-induced laughs to which they were compared. Bryant and Aktipis (2014), in particular, requested laughter from participants without any instructions regarding a target emotional state, whereas Anikin and Lima (2016) used acted exemplars from preexisting published corpora.

Results from perceptual studies of vocal expressions are, on the whole, no less equivocal. In line with some of the acoustic analyses, several authors have reported that, in comparison to natural prosody, acted expressions received higher listener ratings of arousal or emotional intensity (Auberge et al., 2004; Shahid, Krahmer, & Swerts, 2008; Audibert et al., 2008; Laukka, Audibert, & Auberge, 2012), possibly rendering each expression more recognizable than its natural manifestation (Batliner et al., 2000). Conversely, Drolet, Schubotz and Fischer (2012, 2013, 2014) showed that, with the exception of anger, acted emotions were not better recognized than natural exemplars; if anything, the reverse was true (Drolet et al., 2012). Once again, studies of non-speech expressions have further complicated the emerging picture: volitionally-produced laughter is routinely perceived as less rather than more aroused than natural laughs (McGettigan et al., 2013; Lavan et al., 2016; Lavan & McGettigan, 2016), and Anikin and Lima (2016) demonstrated that arousal-related F0 measurements were strong, positive predictors that a participant would judge a nonverbal vocalization to have originated from a natural source. Yet this same research team, in assembling a recent corpus, discovered no correlation between ratings of arousal and stimulus authenticity (Lima, Castro, & Scott, 2013).

Taking together the evidence from acoustic comparisons and subjective ratings, it remains ambiguous as to what extent, and by which reliable cues, acted and natural vocal expressions are perceptibly distinct. It seems likely, given the heterogeneity of results so far, that if markers of naturalness or of acting exist at all, they are both context- and emotion-dependent, as well as conditional on the vocal modality through which that emotion is carried (e.g., emotion-laden speech versus laughter). It is therefore essential that investigators continue to consider this question across a wide variety of vocalizations and eliciting scenarios, the better to account for prevailing variability and arrive at a thorough understanding of any constraints attending the use of acted expressions.

The present experiment

Our study is the first to explore the possibility that differences between acted and natural expressions are manifest and perceptible in human screams. Screams constitute a unique category of vocalization (Green, Whitney, & Potegal, 2011), exemplified by high mean fundamental frequencies (F0s), broad F0 variations, and substantial harmonic energy in high-frequency bands (Green et al., 2011), in addition to rapid temporal modulations that often lend them a subjective quality of "roughness" (Arnal, Flinker, Kleinschmidt, Giraud, & Poeppel, 2015). From the point of view of our discussion, screams afford a compelling opportunity to test predictions that may follow from the previously-outlined arguments. For example, if acted expressions are indeed regarded as more intense or aroused than natural expressions (Auberge et al, 2004), screams could represent a revealing exception: given that they already co-occur with states of escalated physiological arousal (Green et al., 2011), their further exaggeration may be limited. On the other hand, it has been theorized that at least some screams are more compulsory than controlled—that is, they are more so reflexive, emotional outpourings than they are deliberate, emotive signals—and therefore they might prove especially difficult to imitate voluntarily (Scherer, 1994).

As yet, research in human screams is sparse, especially relative to an extensive literature on agonistic screams in non-human primates (Gouzoules, Gouzoules, & Marler, 1984). The screams emitted by rhesus macaques (*Macaca mulatta*), chimpanzees (*Pan troglodytes*), and other primate species during their aggressive encounters serve to recruit aid from allied conspecifics (de Waal & Van Hooff, 1981), and in that function may convey information about the caller's identity (Cheney & Seyfarth, 1980), the identity of the opponent (Gouzoules et al., 1984), and the severity of the aggression (Slocombe, Townsend, & Zuberbühler, 2008), among other details of feasible relevance to their far-flung confederates (Gouzoules et al., 1984). Yet, the potential for human screams to inform on comparably multifarious dimensions remains to this day virtually unaddressed, likely owing in part to reservations concerning the use of acted material. Certainly, neither recording screams in a naturalistic context nor conjuring them forth through induction procedures would seem wholly practical, or ethical if that (Juslin & Scherer, 2005). The few experimental programs that have incorporated screams—usually intermixed among other nonlinguistic utterances—have tended to rely on acted exemplars to comprise their stimulus sets (Hawk et al., 2009; Banissy et al., 2010; Sauter et al., 2010; Dellacherie, Hasboun, Baulac, Belin, & Samson, 2011; Aube, Angulo-Perkins, Peretz, Concha, & Armony, 2014). Thus it is for pragmatic as well as theoretical reasons that the validity of research using acted screams must undergo formal evaluation.

To that end, we used a forced-choice paradigm to test whether listeners could discriminate acted screams from naturally-produced exemplars. In this procedure, participants were required to choose whether each delivered expression was acted or natural, allowing the calculation of discrimination scores as an indicator of perceptual differences between these two encoding conditions. Of the previous studies that have employed this paradigm with acted and natural vocal expressions of emotion, most have reported discriminatory scores slightly to moderately above chance levels (e.g., 59% mean accuracy: Auberge & Cathiard, 2003; 59%: Jürgens, Drolet, Pirow, Scheiner, & Fischer, 2013; 60%: Jürgens, Grass, Drolet, & Fischer, 2015; 65%: Anikin & Lima, 2016). However, only Anikin and Lima (2016) used nonverbal expressions in their design (as opposed to emotional speech), and these researchers did not investigate screams as a case of special empirical interest.

Beyond our primary goal of ascertaining whether acted and natural screams are audibly distinct, we pursued evidence towards addressing two additional concerns. First, a number of studies have documented a broad spectrum of individual abilities on tasks comparable to our own (Auberge & Cathiard, 2003; Audibert et al., 2008; McGettigan et al., 2013). Auberge and Cathiard (2003), for example, reported that 8 out of 20 participants achieved above-chance levels of discrimination between acted and natural prosody, ranging from 67 to 90% accurate response rates, whereas the other 12 participants failed to perform at levels exceeding the chance criterion. It is therefore of interest to determine what personal or demographic traits might underlie some of this considerable variation. One candidate predictor is the sex of the listener, which previous research has suggested could affect the perception and recognition of emotional expressions (Vigil, 2009). Females might in general detect emotional cues more proficiently than males (Belin, Fillion-Bilodeau, & Gosselin, 2008; Vigil, 2009; but see Fugate, Gouzoules, & Barrett, 2009), or there may prevail an interaction effect whereby listeners of one sex are more (or less) discerning in attending to vocalizers of the other (Belin et al., 2008; McKeown, Sneddon, & Curran, 2015). Other differences in emotional perspicacity have been described in relation to a participant’s age (e.g., impairments in older adults; Hunter, Phillips, & MacPherson, 2010), spoken language (as congruent or not with that of the vocalizer, where the former is at an advantage; Mesquita and Frijda, 1992), and trait empathy (defined here as a tendency to understand the internal states of others through affective as well as cognitive representations; Preston & de Waal, 2002). We therefore collected data on each of these factors, along with more direct measures of socioemotional acuity (Golan, Baron-Cohen, & Hill, 2006), in order to determine whether these participant-related attributes might explain significant variability in task performance.

Finally, from a subset of participants, we requested ratings of emotional intensity for each scream in the stimulus set. Thus, we were able to test directly the hypothesis that acted expressions are perceived as more emotionally intense than natural expressions (Auberge et al., 2004), as well as the logical corollary that the perceived intensity of a scream could predict a participant's likelihood of identifying that scream as acted (Audibert et al., 2008).

Methods

Testing took place from July, 2016 to February, 2017 at the Emory University Psychology Department's Bioacoustics Laboratory. This research was approved by and conducted in compliance with Emory's Institutional Review Board.

Participants

One-hundred thirty six participants (101 females) took part in the study. Ages ranged from 16 to 42 years (*M* = 18.96, *SE* = .20). One-hundred four participants listed English as their first spoken language (31 listed a different first language, while one did not provide this information). Most participants were undergraduates recruited via an online portal system who received class credit for completion of the study; those otherwise recruited were not compensated in any way. All participants provided their voluntary and informed consent.

Materials

Apparatus. Experimental procedures were performed on a Dell OptiPlex 755 computer. Sounds were delivered through a pair of headphones (JVC G-Series model HA-G55, JVCKENWOOD USA Corporation, Long Beach, CA, or a Beyerdynamic DT 770 PRO, Beyerdynamic GmbH & Co., Heilbronn, Germany). E-Prime 2.0 software (Psychology Software

Tools, Inc., Pittsburgh, PA) was used to present stimuli and collect data during the forced-choice discrimination task and the emotional intensity rating task. Participants provided inputs on the former via a peripheral serial response box (model 200a, Psychology Software Tools, Inc., Pittsburgh, PA), whereas intensity ratings were selected using the computer mouse. E-Prime recorded all responses and their latencies with millisecond-precision timing. A third task, the Cambridge Mindreading (CAM) Face-Voice Battery (Golan et al., 2006), was presented using DMDX Experiment Software (University of Arizona; Forster & Forster, 2003). Participants produced their responses on this task using the 1-4 numerical keys on the computer's keyboard.

Stimuli. Acted stimuli were compiled from scripted television programs, commercial movies, and advertisements, as well as from publicly available online sound banks (Human Sound Effects, Partners In Rhyme, Inc., Santa Monica, CA; The Nightingale Voice Box, Nightingale Music Productions, North York, Ontario, Canada). Natural stimuli were collected from YouTube clips and unscripted television programs. Screams were not added to the natural set until two experimenters agreed that they were likely produced without explicit prompting based on the surrounding context (e.g., the vocalizer was reacting to a surprise visitor and evidently not aware of the camera while screaming). Online videos were downloaded using Total Recorder version 8.0 (High Criteria, Inc., Richmond Hill, Ontario, Canada) and WinXHD Video Converter Deluxe (Digiarty Software, Inc., Chengdu, China), while DVD media were extracted using WinXDVD Ripper Platinum (Digiarty Software, Inc., Chengdu, China). All source videos were saved, converted to the MPEG file format, and cropped at timestamps surrounding the target vocalizations for future reference regarding their eliciting scenarios. Name and sex of the vocalizer were recorded whenever possible, in addition to relevant contextual details.

Audio files were isolated and converted to mono 16-bit 22.05 kHz WAV files using Adobe Audition CC (Adobe Systems, San Jose, CA) and Audacity version 2.1.2 (http://audacity.sourceforge.net). Edits were applied only when necessary to mitigate noise, delete clicks or pops, and, in the case of some DVD sources, remove a track containing background music; any file requiring further modifications was not included in the stimulus set. Short segments of silence were retained before and after each sound so as not to introduce any artifacts due to abrupt starts or stops. RMS amplitude normalization was applied across files to achieve relative uniformity of presentation volume. Final waveforms and spectrograms were carefully inspected to ensure no evidence of clipping or other serious distortions.

The final stimulus set was selected according to the following criteria: (1) to include only the highest-quality sounds, as agreed upon by two experimenters, though some low-level background noise was audible in a few stimuli of each condition; (2) to represent a variety of vocalizer attributes (e.g., sex, relative age) and eliciting contexts (e.g., opening gifts, hurting oneself, riding a rollercoaster); and (3) to achieve as much of a balance as possible in these characteristics between acted and natural conditions. In all, 72 exemplars were used as stimuli in the experiment, comprising 36 acted screams and 36 natural screams. Almost every scream was produced by a unique vocalizer, though two acted screams were produced by the same individual. Fifty seven screams were produced by females, and the proportion of male and female vocalizers was approximately equal between the two conditions, $\chi^2(1) = 0.08$, $p = .772$. Scream durations, not including surrounding periods of silence, ranged from 0.29 s to 3.81 s ($M = 1.17$ s) and did not differ significantly between conditions, $t(70) = -0.23$, $p = .819$.

Procedure

Forced-choice discrimination task. At the beginning of the task, participants were informed that they would listen to a series of human screams, and that these screams were collected from a variety of different sources. Participants were told that they would have to judge whether each presented scream was acted or natural. Acted screams were described as those that "came from television or film," whereas natural screams "came from real-life situations"; no further information was provided regarding the sources of exemplars. Participants were told to wait until a stimulus had played completely through, and then to input their responses as quickly and as accurately as possible using buttons labelled "1" and "2" on the serial response box. The mapping of these buttons to "acted" or "natural" was counterbalanced between participants.

Trials were cued with the words "Please Listen Now" appearing in the center of a white visual display. After a period of 2 s, a stimulus was delivered through the headphones while on-screen instructions reminded participants which button corresponded to which response. The next trial began immediately after an input was received.

Participants completed trials in two blocks. A short practice block, consisting of five trials and using exemplars not included in the primary experimental task, served to acquaint participants with the procedure; their practice answers were not preserved. In the subsequent experimental block, participants heard and made judgments regarding the 72 exemplars that comprised the stimulus set, during which all responses as well as response latencies were recorded by E-Prime. In both blocks, stimuli were presented in a randomized order. An experimenter remained unobtrusively on hand to make sure instructions were followed, and to provide clarification as necessary, but otherwise no feedback was offered in either block. Any participant who demonstrated continued lack of understanding, attention, or interest in following directions would have been excluded from analysis, but this did not prove necessary.

Emotional intensity rating task. A subset of participants (n = 36) completed a task in which they rated the emotional intensity of each scream in the stimulus set. Ratings were provided on a scale from 1 to 5, with 1 corresponding to "Not intense at all," and 5 corresponding to "Very intense." Participants used the computer mouse to select a number on screen, after which the selection was displayed for 1 s before moving on to the next stimulus. All stimuli were presented in a randomized order. Additionally, the timing of this task relative to the discrimination task (i.e. prior to or after) was pseudo-randomized between participants.

CAM Face-Voice Battery. Following the perceptual judgment tasks, the CAM Face-Voice Battery (Golan et al., 2006) was used to assess abilities in emotion recognition. This task required participants to watch short, silent clips of actors (facial recognition task) and listen to emotionally-intoned, but semantically neutral sentences (vocal recognition task) and determine which mental or emotional states were being portrayed, given the choice of four possible options in each trial. The battery returns four scores for each participant: a *facial emotion recognition* score (out of 50), a *vocal emotion recognition* score (out of 50), an *overall emotion recognition* score (out of 100; the sum of the previous scores), and a *concepts correctly recognized* score (out of 20), based on how many distinct emotional states the participant was able to identify accurately. A more detailed description of the CAM Face-Voice Battery and its scoring can be found in Golan et al. (2006).

Questionnaires. Additional participant information was collected with the use of self-report questionnaires. The first was a ten-item survey developed specifically for this study. Seven items were designed to estimate the participant's prior exposure to screams in the media, based on his or her knowledge and consumption of video games, television, and film genres likely to contain screams. All answers were provided on a five-point scale, where higher ratings indicated greater knowledge or familiarity, and responses were summed to derive a participant's single *Experience Index*. The remaining three items gauged the participant's confidence in making judgments during the experiment and in reading emotions in general; these responses were summed to provide the participant's *Confidence Index*. Data on sex, age, first spoken language, and handedness were collected on this questionnaire as well.

Empathy was measured using the Cambridge Behavior Scale (Baron-Cohen & Wheelwright, 2004), a 40-item questionnaire that assesses cognitive as well as emotional aspects of empathic ability. The resultant *Empathy Quotient* (EQ) ranges from 0 to 80, with higher EQs corresponding to greater levels of empathy.

Participants also completed the second generation Online Alexithymia Questionnaire (OAQ-G2; Thompson, 2007). Individuals with alexithymia or alexithymic tendencies have difficulty identifying and processing their own emotions, a deficit commonly associated with a difficulty in identifying the emotions of others (FeldmanHall, Dalgleish, & Mobbs, 2013). The OAQ-G2 has a maximum score of 185, with higher scores suggesting the presence of more alexithymic traits.

Statistical Analysis

Data preparation. Reaction times for each trial were calculated from the offset time of the stimulus. Data from trials where responses were provided before at least 50% of the sound had played were excluded from all analyses (resulting in the exclusion of 15 trials from a total of 9,792 trials across all participants). Additionally, for model building, responses from seven participants were omitted due to incomplete data (remaining n = 129).

Overall performance. Response accuracy rates and discrimination levels (*d'* scores) were calculated for each participant. *D'* is a statistic used in signal detection theory (Macmillan & Creelman, 2004) which, as applied here, provides an estimate of a participant's sensitivity to the difference between two stimulus groups (i.e., acted and natural) while separating out the effects of individual response biases. In this analysis, responses were conceptualized as *Hits* and *False Alarms*, where a Hit entailed correctly identifying an acted scream as such, and a False Alarm entailed incorrectly judging a natural scream as acted. A participant's *d'* score was then calculated as the difference between her Z-transformed *Hit Rate* (the proportion of acted trials that yielded Hits) and her Z-transformed *False Alarm Rate* (the proportion of natural trials that yielded False Alarms).

Factors affecting response. Generalized linear mixed models (GLMMs) were constructed to examine the relationship between participant attributes and response parameters. GLMMs allow for the incorporation of random effects related to each participant and stimulus, thereby controlling for the nonindependence of multiple observations associated with these factors (Barr, Levy, Scheepers, & Tily, 2013). Initial models were fitted with all candidate predictors as fixed factors and participant and stimulus as random intercepts. Insignificant fixed terms were sequentially removed until only those that explained a significant portion of the variability remained in a minimal model. Removed terms were then reentered individually to ensure that their inclusion did not significantly improve model fit according to a likelihood-ratio test (Barr et al., 2013). Minimal models were evaluated with respect to a null model (in which all fixed factors were omitted, leaving only the intercept and random effects) using a likelihood-ratio test. The significance of beta estimates for each remaining term was assessed using a Wald test as well as a likelihood-ratio test in which the model was compared to a nested model omitting only the predictor of interest (Dautriche & Chemla, 2014); in every case, results from these tests converged, and so only Wald *Z*-statistics (for logistic mixed models), χ^2-statistics (for linear mixed models), and their corresponding *p*-values are reported. Tables of fixed factors for all significant minimal models are included in the Appendix.

Emotional intensity ratings. To explore the relationship between emotional intensity ratings and the acted or natural condition of a stimulus, as well as between these ratings and the participant's response, generalized linear mixed models (GLMMs) were specified using the subset of data for which these ratings were available (n = 36). Likelihood-ratio tests were used to determine whether the inclusion of the intensity ratings as predictors improved the model's fit relative to a null.

Data were prepared and signal detection analyses were conducted using SPSS Statistics version 24 (IBM Corp., Armonk, NY). Models were implemented and tested in the R Statistical Environment (R Core Team, 2013).

Results

An $\alpha = .05$ was adopted for all statistical tests.

Overall performance

Participants achieved a mean accuracy rate of 50.52% (*SE* = 0.49) and a mean *d'* score of 0.03 (*SE* = 0.03). In signal detection analysis, *d'* = 0.0 designates chance-level performance (Macmillan & Creelman, 2004); therefore, neither accuracy rates (*t(*135) = 1.07, *p* = .289) nor *d'* scores (*t*(135) = 1.06, *p* = .293) significantly exceeded the chance criterion. In fact, taking the empirical threshold of *d'* > 1.0 as indicative of moderate sensitivity to the tested perceptual dimension (Macmillan & Creelman, 2004), no participants showed discrimination between acted and natural screams (max *d'* = 0.79). Paired sample T-tests revealed no differences between

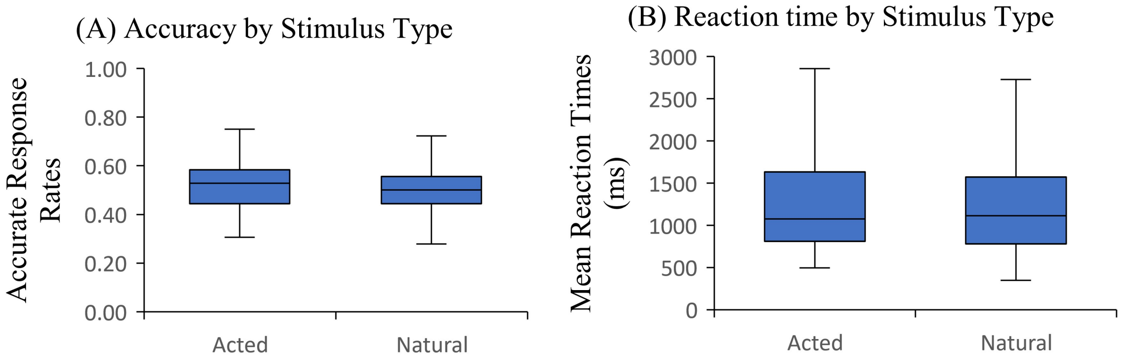

Figure 1. No significant differences were observed between (A) participants' accuracy rates nor (B) their reaction times in response to acted and natural exemplars. Outliers were omitted from the plot of reaction times.

acted and natural conditions in accuracy rates ($t(135) = 0.68$, $p = .501$) nor in reaction times on accurate responses ($t(135) = -0.53$, $p = .589$), indicating that participants did not perform any better at correctly identifying natural relative to acted screams, or vice versa (**Figure 1**).

Factors affecting response

A logistic GLMM assessed the relationship between participant attributes and the likelihood of responding accurately on any given trial. Fixed factors entered into the initial model included listener sex (male = 1); age; first language, as a binary variable to capture match (0) or non-match (1) with the vocalizer's language; Experience Index; Confidence Index; EQ; total CAM score; and alexythemia score, as well as sound duration, trial number, vocalizer sex (male = 1), and the listener by vocalizer sex interaction. After stepwise elimination of insignificant terms, only sound

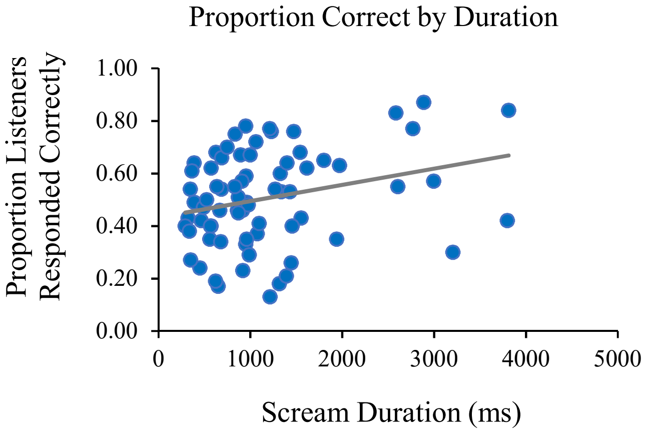

Figure 2. The proportion of listeners who responded correctly on any given scream increased with the duration of that scream, $R^2 = .08$, $p = .019$.

duration remained as a significant predictor (Table A1 in the Appendix), suggesting a slight increase in accuracy as duration increased ($\beta = 0.29$, $SE = 0.11$, $Z = 2.55$, $p = .019$). To confirm this finding, the proportion of participants who responded accurately to each sound was computed and linearly regressed on sound duration, producing results consistent with the GLMM ($R^2 = .08$, $\beta = 0.06$, $SE = 0.03$, $p = .02$; **Figure 2**). No participant attributes significantly improved the GLMM's fit when added back into the model (likelihood-ratio tests, $p > .11$ in every case), suggesting that these factors did not account for variability in accuracy of response.

A second logistic GLMM assessed the extent to which participant attributes predicted the likelihood of identifying screams as acted. The same fixed factors were entered as candidate predictors and removed according to insignificance by the Wald test. The final model included as significant predictors EQ ($\beta = -0.01$, $SE = 0.002$, $Z = -2.23$, $p = .026$) and sound duration ($\beta = 0.51$, $SE = 0.10$, $Z = 4.90$, $p < .001$; Table A2 in the Appendix); no other factors significantly improved model fit (likelihood-ratio tests, $p > .07$ in every case). Note the very small effect size for EQ, particularly relative to that for duration. This is also evident in **figure 3**, which shows simple linear regressions of (a) the proportion of screams judged as acted on participants' EQ scores ($R^2 = .03$, $\beta = -0.001$, $SE = .001$, $p = .039$), and (b) the proportion of participants who judged screams as acted on scream durations ($R^2 = .25$, $\beta = .11$, $SE = .02$, $p < .001$).

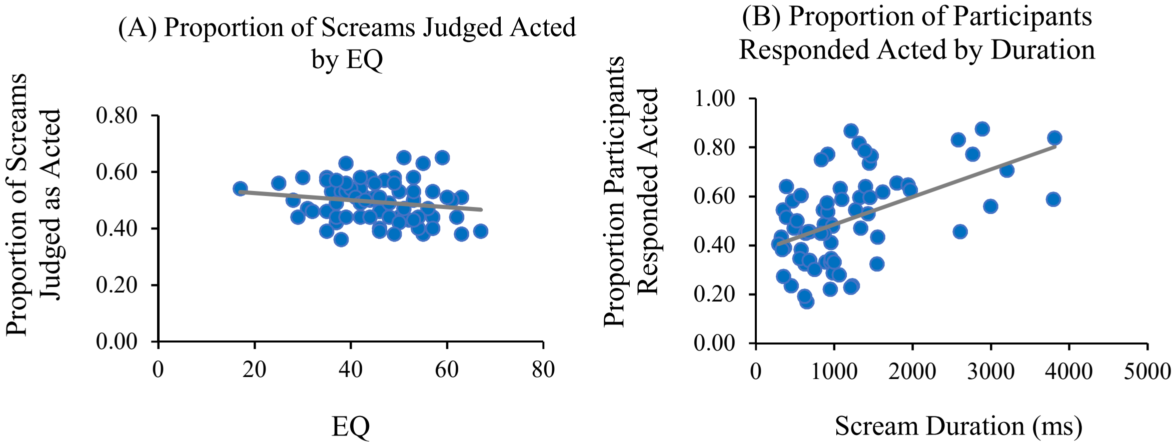

Figure 3. (A) The proportion of screams that a participant judged as acted decreased slightly as EQ increased, $R^2 = .03$, $p = .039$; (B) The proportion of participants who judged a given scream as acted increased with the duration of the scream, $R^2 = .25$, $p < .001$.

Finally, a linear GLMM explored the relationship between reaction times on accurate trials and the same potential factors. Emerging as significant predictors were listener sex (β = 0.50, *SE* = 0.16, $\chi^2(1) = 8.36$, $p = .004$), indicating that females responded more quickly than males; first language (β = 0.52, *SE* = 0.17, $\chi^2(1) = 9.83$, $p = .002$), indicating that English speakers responded more quickly than non-English speakers; sound duration (β = -140.46, *SE* = 21.71, $\chi^2(1) = 41.84$, $p < .001$), indicating that longer sounds elicited quicker reaction times; and the interaction of listener by vocalizer sex (β = -0.18, *SE* = .06, $\chi^2(1) = 7.82$, $p = .005$; Table A3 in the Appendix), indicating that male listeners reacted slightly more quickly to male screams than to female screams, whereas female listeners did not show this trend.

Emotional intensity ratings

A logistic GLMM was fitted with scream type as a dichotomous outcome variable (acted = 1) and emotional intensity ratings as fixed factors (dummy coded so that ratings 2, 3, 4, and 5 were entered as separate variables with reference to rating 1). A likelihood-ratio test comparing this model to a model without intensity ratings revealed no differences in model fit ($\chi^2(4) = 0$, p = 1), indicating that intensity ratings did not predict whether or not a scream was acted. A second logistic GLMM used the same model specifications but with participants' responses as the dichotomous outcome variable (acted = 1). Again, this model did not fit the data better than the null ($\chi^2(4) = 7.22$, $p = .125$), suggesting that listeners did not respond as a function of the perceived intensity of the scream.

Discussion

We tested whether participants could distinguish acted from natural screams in a forced-choice discrimination task. Signal detection analyses revealed that participants were no more sensitive to the difference between acted and natural exemplars than would be expected by chance, suggesting that listeners are not capable of making this discrimination. These results indicate that acted and natural screams are not perceptibly distinct, and therefore imply that screams produced voluntarily can convincingly represent those that occur in nature.

We had predicted that known factors pertaining to emotional perception might be associated with variation in performance. For the most part, we did not find evidence to support this hypothesis; no measured participant-related attributes affected the listener's accuracy of response. Empathy levels, as assessed by the Cambridge Behavior Scale (Baron-Cohen & Wheelwright, 2004), negatively correlated with the probability that a listener would perceive any given scream as acted, which may suggest that high-empathy individuals more readily trust in the sincerity of expressions; however, this effect was very small and requires replication. Possibly, the discrimination task was so difficult that meaningful variation between participants could not emerge, as we might expect if listeners were reduced to guessing on trials more often than not. This explanation seems unlikely, however, as despite the low levels of discrimination overall, we observed a range of accuracy rates (36-65%) as well as the total proportion of screams that each participant identified as acted (35-69%). Alternatively, our choice of participant-related factors may not have appropriately captured the sources of variability on this task. Anikin and Lima (2016) noted that cuing listeners as to the emotion that a nonverbal sound conveyed did not facilitate their abilities to discriminate acted from natural vocalizations. It may be that this judgment is not directly contingent upon decoding specific underlying states; thus advantages in vocal emotion recognition may not transfer to improved performance on the task.

That said, we did observe reaction time-related effects consistent with tendencies broadly reported in the emotion recognition literature. Namely, females responded more quickly than males on accurate trials, as did native relative to non-native language speakers, findings which could be interpreted as congruent with evidence that each of these groups is at an advantage when deciphering emotional expressions (sex differences reviewed in Vigil, 2009; language and cultural differences in Elfenbein & Ambady, 2002). Of additional interest in our data was an interaction between listener and vocalizer sex, such that males reacted more quickly to male- than to female-vocalized screams, whereas female listeners did not show this bias. McKeown et al. (2015) previously reported that in laughter, male listeners performed best when discriminating male vocalizers; these researchers suggested that men were unduly inclined to believe that female-produced expressions were genuine. It is difficult, however, to draw the same conclusions from our results, provided that none of these factors affected the accuracy of discrimination nor the probability of identifying screams as acted. Future research on screams should determine whether these effects on response latency are specific to this task, or whether they simply are indicative of general, behavioral inclinations elicited by these vocalizations.

Though we were unable to characterize much of the variation between participants, we did discover moderate to large effects of scream duration on participants' responses. Response accuracy increased as a function of duration, perhaps suggesting that listeners could discriminate more precisely when afforded the additional acoustical information of a longer and protracted stimulus. Audibert et al. (2008) reported a comparable enhancement on an acted-natural discrimination task when listeners heard extended prosodic utterances as opposed to monosyllabic sounds. It seems possible that listeners in our study similarly benefited from added details, perhaps conveyed through melodic contour or otherwise made discernable across the time course of longer screams. A more prominent effect, however, was evident in the positive relationship between a scream's duration and the probability that it was perceived as acted. In fact, this correlation was so robust as to explain 25% of the variation in the proportion of participants who identified particular screams as acted. Sustained vocal production clearly operated as an important cue by which participants judged the artificiality of a scream, though it is unclear as to what extent this reflects a legitimate marker of acted expression versus the participants' prepossessed stereotypes regarding screams that occur in scripted film or television. Note that in this study, we matched the mean duration of screams between acted and natural conditions, but it is possible that in frequency of occurrence, acted screams are more typically prolonged.

Our findings stand in qualified opposition to previous studies assessing the discriminability of acted and natural vocal expressions, which have reported mean accuracy rates ranging from 59% (Auberge & Cathiard, 2003; Jürgens et al., 2013) for emotional prosody to 65% (Anikin & Lima, 2016) for a variety of nonlinguistic utterances. It is possible that natural sounding screams are relatively easy to produce—and therefore less distinct from naturally occurring exemplars—compared to emotion-laden speech or other non-verbal vocalizations. One argument to that effect is that screams are already so emotionally intense that actors are less prone to exaggerating their features to the point of artificiality. Indeed, we found that the perceived intensity of screams did not differ between acted and natural exemplars, suggesting that, on average, actors did not over-express their intended emotional states. However, we also found that these intensity ratings did not predict listeners' judgments as to whether or not a scream was acted; it is unlikely, then, that the lack of discrimination on this task can be attributed to similarities relating to the intensity of the stimuli, given that listeners did not use perceived intensity to make their judgments. Nonetheless, our results would indicate that the criticisms of acted material as problematically intense (Greasley et al., 2000) do not apply to all instances of acted expression.

If acted and natural screams differ in other ways, it may be that listeners are not attuned to these contrasts. While most individuals are likely to have heard and produced screams from a very young age (Green et al., 2011), certain variants used in this study (e.g., fear or pain screams) are almost certainly not as ubiquitous in the everyday interactions of our participants as expressions such as laughter or emotional speech. Thus, listeners may not possess much experience with certain kinds of natural screams—at least relative to other vocalizations—nor much practice in particular at segregating acted from natural exemplars. Bear in mind, however, that screams in other natural contexts are probably not so rare, and pending ongoing research in our laboratory, it remains to be seen whether context even matters with respect to a scream's acoustic profile. Perhaps a more parsimonious account, in light of the observed correlation between duration and response accuracy, is that screams are generally shorter than other vocalizations and therefore furnish fewer cues by which to judge their naturalness.

From a comparative perspective, it is likely that screams were heavily shaped in the context of kin communication. Analogous screams in non-human primates function to solicit agonistic aid from the caller's relatives, if not otherwise bonded allies (Cheney, 1977; de Waal &

Van Hooff, 1981; Gouzoules et al., 1984). If both signaler and receiver in these interactions benefit from the same behavioral outcomes—as, for example, when a mother profits indirectly through her offspring's agonistic success (Silk, 2002)—then it could be argued that there is scant selective advantage in producing these calls deceptively (Maynard Smith & Harper, 1995); in other words, reliability might derive not from physical restrictions but from the simple fact that callers have little reason to deceive. This explanation is not completely satisfactory, however, as even between potential allies, chimpanzees have been observed to exaggerate the severity of an encounter in their recruitment screams (Slocombe & Zuberbühler, 2007); indeed, from a theoretical standpoint, selection in favor of kin cooperation constrains but does not preclude individual manipulation in between-kin interactions (see, for example, parent-offspring conflict theory; Trivers, 1974). For humans, at least, it seems more plausible that the possible benefits of a deceptive scream are typically outweighed by the risks of social punishment or other extrinsically delivered costs (Dezecache et al., 2013). Further research should explore the circumstances in which screams are produced and the means by which their reliability is maintained. At present, it is apparent that the costs of scream production itself are not sufficient to prevent their willful replication.

It is critical to point out that the absence of explicit discrimination as demonstrated here does not necessarily imply that acted and natural screams are equivalent in every respect. Notably, cues may exist that elude conscious detection but still elicit differential patterns of neural response. Drolet et al. (2012, 2013), for example, found significant differences in the fMRI-BOLD signal evoked by acted and natural prosody, with the latter marked by greater activation in areas associated with social cognition, despite observing poor discrimination in the forced-choice paradigm. This team also reported an interaction between the acted or natural condition of a stimulus and participants' performance on an emotion recognition task, such that anger was better recognized in acted exemplars whereas sadness was more discernable when naturally conveyed. One future direction, then, is to determine through imaging and through implicit tasks whether acted and natural screams are disambiguated in listeners' responses even if they are not themselves aware of the distinction. That said, we are doubtful whether acted and natural screams would evoke manifestly dissimilar response patterns, given the total lack of differences in accuracy rates as well as in response latencies between our task conditions. A more promising avenue of research might be to determine whether participants benefit from multimodal presentation; research in affective speech, for example, suggests that audiovisual displays are better discriminated than either modality presented alone (Audibert et al., 2008). (Note, however, that screams are likely to have functioned evolutionarily in long-distance contexts where only auditory cues were available to receivers.)

One potential limitation of this study is that natural screams were by necessity procured through sources made available on public video sharing sites. It could be argued that some of the eliciting scenarios were not naturalistic in the strictest sense, and that the camera could alter the expressive behavior of subjects who knew they were being filmed. Efforts were made, however, to alleviate these issues. Many screams were extracted from contexts in which the vocalizers were reacting to unanticipated events (e.g., when surprised by a visiting relative); did not know that they were on film (e.g., when pranked on a hidden camera video), and/or were not apparently disposed towards attending to the camera in that moment (e.g., when suffering the pain from a fall). If anything, we believe the conservativeness according to which we conducted our scream selection process may have made the task easier to perform, as we excluded from consideration any scream where the authenticity of the surrounding clip was in question. Because participants failed to discriminate regardless, we doubt that this matter is of preeminent concern.

Research on human screams is in its infancy, and it is apt to advance principally through the continued use of acted material. Our findings are somewhat reassuring with respect to the generalizability of scream research thus far (Huang et al., 2010; Arnal et al., 2015), and it is our hope that upcoming studies will further illuminate the functional and informative capacities of these exceptional vocalizations. We do not suggest, however, that all acted screams are equally suitable for empirical implementation. That we demonstrated no perceptual differences between acted and natural screams implies that screams are believably reproducible, at least by professional actors; that is, there is nothing that intrinsically prohibits the production of deliberate and credible facsimiles. Nonetheless, it is probable that screams, like other expressions of emotion (Scherer, 1994), fall on a continuum of voluntary to involuntary determination, and it is possible that those occupying the extremes are in fact discriminable to listeners. Caution should remain the rule as investigation presses forward.

References

Anikin, A., & Lima, C. F. (2016). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *The Quarterly Journal of Experimental Psychology*, (just accepted), 1-53.

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*(15), 2051-2056.

Aubé, W., Angulo-Perkins, A., Peretz, I., Concha, L., & Armony, J. L. (2014). Fear across the senses: Brain responses to music, vocalizations and facial expressions. *Social Cognitive and Affective Neuroscience*, *10*(3), 399-407.

Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Communication*, *40*(1), 87-97.

Aubergé, V., Audibert, N., & Rilliard, A. (2004). Acoustic morphology of expressive speech: What about contours? In *Proceedings of Speech Prosody 2004* (pp. 201-204). Nara, Japan.

Audibert, N., Aubergé, V., & Rilliard, A. (2008). How we are not equally competent for discriminating acted from spontaneous expressive speech. In *Proceedings of Speech Prosody 2008* (pp. 693-696). Campinhas, Brazil.

Audibert, N., Aubergé, V., & Rilliard, A. (2010). Prosodic correlates of acted vs. spontaneous discrimination of expressive speech: a pilot study. In *Proceedings of Speech Prosody 2010*. Chicago, United States.

Bachorowski, J. A., & Owren, M. J. (2001). Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychological Science*, *12*(3), 252-257.

Banissy, M. J., Sauter, D. A., Ward, J., Warren, J. E., Walsh, V., & Scott, S. K. (2010). Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. *Journal of Neuroscience*, *30*(41), 13552-13557.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614-636.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163-175.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Barrett, L. F. (2011). Was Darwin wrong about emotional expressions? *Current Directions in Psychological Science*, *20*(6), 400-406.

Batliner, A., Fischer, K., Huber, R., Spilker, J., & Nöth, E. (2000). Desperately seeking emotions or: Actors, wizards, and human beings. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion* (pp. 195-200). Belfast, Ireland.

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531-539.

Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1-20.

Brown, B.L., & Bradshaw, J.M. (1985). Towards a social psychology of voice variations. In H. Giles & R. N. St. Clair (Eds.), *Recent Advances in Language, Communication and Social Psychology* (pp. 144-181). London, England: Lawrence Erlbaum Associates.

Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior*, *35*(4), 327-335.

Cheney, D. L. (1977). The acquisition of rank and the development of reciprocal alliances among free-ranging immature baboons. *Behavioral Ecology and Sociobiology*, *2*(3), 303-318.

Cheney, D. L., & Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, *28*(2), 362-367.

Cowie, R., & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, *40*(1), 5-32.

Darwin, C., Ekman, P., & Prodger, P. (1998). *The expression of the emotions in man and animals*. London, England: Harper Collins.

Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 892.

Davies, N. B., & Halliday, T. R. (1978). Deep croaks and fighting assessment in toads *Bufo bufo*. *Nature*, *274*(5672), 683-685.

Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society of London B: Biological Sciences*, *205*(1161), 489-511.

De Waal, F. B., & Van Hooff, J. A. (1981). Side-directed communication and agonistic interactions in chimpanzees. *Behaviour*, *77*(3), 164-198.

Dellacherie, D., Hasboun, D., Baulac, M., Belin, P., & Samson, S. (2011). Impaired recognition of fear in voices and reduced anxiety after unilateral temporal lobe resection. *Neuropsychologia*, *49*(4), 618-629.

Dezecache, G., Mercier, H., & Scott-Phillips, T. C. (2013). An evolutionary approach to emotional communication. *Journal of Pragmatics*, *59*, 221-233.

Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*,*40*(1), 33-60.

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(1), 140-150.

Drolet, M., Schubotz, R. I., & Fischer, J. (2013). Explicit authenticity and stimulus features interact to modulate BOLD response induced by emotional speech. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(2), 318-329.

Drolet, M., Schubotz, R. I., & Fischer, J. (2014). Recognizing the authenticity of emotional expressions: F0 contour matters when you need to know. *Frontiers in Human Neuroscience*, *8*, 144.

Duchenne de Boulogne, G. B. (1990). *The mechanism of human facial expression.* (R. A. Cutherbertson, Ed. and Trans.). Cambridge, England: Cambridge University Press. (Original work published 1862).

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169-200.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, *58*(2), 342.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin*, *128*(2), 203.

Evans, M. R., & Hatchwell, B. J. (1992). An experimental study of male adornment in the scarlet-tufted malachite sunbird: II. The role of the elongated tail in mate choice and experimental evidence for a handicap. *Behavioral Ecology and Sociobiology*, *29*(6), 421-427.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The Journal of Economic Perspectives*, *14*(3), 159-181.

FeldmanHall, O., Dalgleish, T., & Mobbs, D. (2013). Alexithymia decreases altruism in real social decisions. *Cortex*, *49*(3), 899-904.

Fernández-Dols, J. M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, *5*(1), 24-29.

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050-1057.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116-124.

Fridlund, A. J. (1991). Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology*, *32*(1), 3-100.

Fugate, J. M., Gouzoules, H., & Barrett, L. F. (2009). Separating production from perception: Perceiver-based explanations for sex differences in emotion. *Behavioral and Brain Sciences*, *32*(05), 394-395.

Gerhard, D. (2003). *Pitch extraction and fundamental frequency: History and current techniques* (Technical Report TR-CS 2003-06). Department of Computer Science, University of Regina, Canada.

Gervais, M., & Wilson, D. S. (2005). The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly Review of Biology*, *80*(4), 395-430.

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders*, *36*(2), 169-183.

Goudbeek, M., & Scherer, K. (2010). Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America*, *128*(3), 1322-1336.

Gouzoules, H., & Gouzoules, S. (2007). The conundrum of communication. In C. J. Campbell, A. Fuentes, K. C. MacKinnon, M. Ranger, & S. K. Bearder (Eds.), *Primates in perspective* (pp. 621-635). New York, NY: Oxford University Press.

Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Animal Behaviour*, *32*(1), 182-193.

Greasley, P., Sherrard, C., & Waterman, M. (2000). Emotion in language and speech: Methodological issues in naturalistic approaches. *Language and Speech*, *43*(4), 355-375.

Green, J. A., Whitney, P. G., & Potegal, M. (2011). Screaming, yelling, whining, and crying: categorical and intensity differences in vocal expressions of anger and sadness in children's tantrums. *Emotion*, *11*(5), 1124-1133.

Gunnery, S. D., & Hall, J. A. (2015). The expression and perception of the Duchenne smile. In A. Kostic & D. Chadee (Eds.), *The social psychology of nonverbal communication* (pp. 114-133). Basingstoke, United Kingdom: Palgrave Macmillan.

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, *21*(5), 531-540.

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*(3), 293-305.

Hess, U., & Thibault, P. (2009). Darwin and emotion expression. *American Psychologist*, *64*(2), 120-128.

Hinde, R. A. (1981). Animal signals: ethological and games-theory approaches are not incompatible. *Animal Behaviour*, *29*(2), 535-542.

Huang, W., Chiew, T. K., Li, H., Kok, T. S., & Biswas, J. (2010). Scream detection for home applications. In *The 5th IEEE Conference on Industrial Electronics and Applications* (pp. 2115-2120). Taichung, Taiwan.

Hunter, E. M., Phillips, L. H., & MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychology and Aging*, *25*(4), 779-787.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99*(3), 561-565.

Jürgens, R., Drolet, M., Pirow, R., Scheiner, E., & Fischer, J. (2013). Encoding conditions affect recognition of vocally expressed emotions across cultures. *Frontiers in Psychology*, *4*, 1-10.

Jürgens, R., Grass, A., Drolet, M., & Fischer, J. (2015). Effect of acting experience on emotion expression and recognition in voice: Non-actors provide better stimuli than expected. *Journal of Nonverbal Behavior*, *39*(3), 195-214.

Jürgens, R., Hammerschmidt, K., & Fischer, J. (2011). Authentic and play-acted vocal emotion expressions reveal acoustic differences. *Frontiers in Psychology*, *2*, 1-11.

Jürgens, U. (2009). The neural control of vocalization in mammals: a review. *Journal of Voice*, *23*(1), 1-10.

Juslin P. N., & Scherer K.R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65-135). New York, NY: Oxford University Press.

Krumhuber, E. G., & Manstead, A. S. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion*, *9*(6), 807-820.

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the national academy of sciences*, *98*(23), 13189-13194.

Lavan, N., & McGettigan, C. (2016). Increased discriminability of authenticity from multimodal laughter is driven by auditory information. *The Quarterly Journal of Experimental Psychology*, *70*(10), 1-10.

Lavan, N., Scott, S. K., & McGettigan, C. (2016). Laugh like you mean it: Authenticity modulates acoustic, physiological and perceptual properties of laughter. *Journal of Nonverbal Behavior*, *40*(2), 133-149.

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: a corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods*, *45*(4), 1234-1245.

Ludlow, C. L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology*, *147*(2), 205-222.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Cambridge University Press.

Marty, A. (1908). Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie. *Kant-Studien*, *13*(1-3), 457-460.

Maynard Smith, M. J., & Harper, D. G. (1995). Animal signals: models and terminology. *Journal of Theoretical Biology*, *177*(3), 305-311.

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2013). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex, 25*(1), 246-257.

McKeown, G., Sneddon, I., & Curran, W. (2015). Gender differences in the perceptions of genuine and simulated laughter and amused facial expressions. *Emotion Review*, *7*(1), 30-38.

Mehu, M., & Scherer, K. R. (2012). A psycho-ethological approach to social signal processing. *Cognitive Processing*, *13*(2), 397-414.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: a review. *Psychological Bulletin*, *112*(2), 179-204.

Moore, S. (1960). *The Stanislavski method: The professional training of an actor.* New York, NY: Viking Press.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*(6685), 573-577.

Owren, M. J., & Bachorowski, J. A. (2003). Reconsidering the evolution of nonlinguistic communication: The case of laughter. *Journal of Nonverbal Behavior*, *27*(3), 183-200.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(01), 1-20.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, *61*(1), 5-18.

Rein, G., Atkinson, M., & McCraty, R. (1995). The physiological and psychological effects of compassion and anger. *Journal of Advancement in Medicine*, *8*(2), 87-105.

Ruch, W., & Ekman, P. (2001). The expressive pattern of laughter. In A. W. Kaszniak (Ed.), *Emotion, Qualia, and Consciousness* (pp. 426-443). Tokyo, Japan: Word Scientific.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145-172.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *48*(5), 1290-1298.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(3), 273-294.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology*, *54*(1), 329-349.

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, *99*(2), 143-165.

Scherer, K. R. (1994). Affect bursts. In S. van Goozen, N. Van de Poll, & J. Sergeant (Eds.), *Emotions: Essays on emotion theory* (pp. 161-193). Psychology Press.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1), 227-256.

Scherer, K. R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech & Language*, *27*(1), 40-58.

Scherer, K. R., & Bänziger, T. (2010). On the use of actor portrayals in research on emotional expression. In K. R. Scherer, T. Bänziger, & E. Roesch, *Blueprint for affective computing: A sourcebook and manual* (166-176). New York, NY: Oxford University Press.

Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*(1), 145-173.

Shahid, S., Krahmer, E., & Swerts, M. (2008). Real vs. acted emotional speech: Comparing Caucasian and South-Asian speakers and observers. In *Proceedings of Speech Prosody 2008*. Campinhas, Brazil.

Silk, J. B. (2002). Kin selection in primate groups. *International Journal of Primatology*, *23*(4), 849-875.

Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, *104*(43), 17228-17233.

Slocombe, K. E., Townsend, S. W., & Zuberbühler, K. (2009). Wild chimpanzees (*Pan troglodytes schweinfurthii*) distinguish between different scream types: evidence from a playback study. *Animal Cognition*, *12*(3), 441-449.

Smith, W. J. (1997). The behavior of communicating, after twenty years. In D. H. Owings, M. D. Beecher, & N. S. Thompson, *Perspectives in ethology: Volume 12: Communication* (pp. 7-53). New York, NY: Plenum.

Thompson, J., (2007). Online Alexythymia Questionnaire – G2. Retrieved from http://oaq.blogspot.com/

Tibbetts, E. A., & Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, *22*(10), 529-537.

Tomkins, S. (1962). *Affect imagery consciousness: Volume I: The positive affects*. New York, NY: Springer Publishing Company.

Trivers, R. L. (1974). Parent-offspring conflict. *American Zoologist, 14*(1), 249-264.

Vehrencamp, S. L. (2000). Handicap, index, and conventional signal elements of bird song. In Y. Espmark, T. Amundsen, & G. Rosenqvist, *Animal signals: Signalling and signal design in animal communication* (pp. 277-300). Trondheim, Norway: Tapir Academic Press.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*(4), 473-482.

Vigil, J. M. (2009). A socio-relational framework of sex differences in the expression of emotion. *Behavioral and Brain Sciences*, *32*(05), 375-390.

Wagner, P., Trouvain, J., & Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, *48*, 1-12.

Wattendorf, E., Westermann, B., Fiedler, K., Kaza, E., Lotze, M., & Celio, M. R. (2013). Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging. *Cerebral Cortex*, *23*(6), 1280-1289.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *The Journal of the Acoustical Society of America*, *52*(4B), 1238-1250.

Wundt (1905). *Grundzüge der physiologischen Psychologie* [Fundamentals of physiological psychology]. Leipzig, German: Engelmann. (Original work published 1874).

Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205-214.

Appendix

Table A1

GLMM on factors affecting the probability that participants would respond accurately on a trial, based on 9,273 trials across 129 participants

	Coefficient	*SE*	Wald statistic (*Z*)	*p*
Minimal model				
Intercept	-0.32	0.16		
Duration	0.29	0.11	2.55	0.011

This model was fitted with a logit-link function, response accuracy (1 = correct) as the binary output variable, and random intercepts for participant and stimulus.

Table A2

GLMM on factors affecting the probability that participants would identify a scream as acted, based on 9,273 trials across 129 participants

	Coefficient	*SE*	Wald statistic (*Z*)	*p*
Minimal model				
Intercept	-0.28	0.20		
EQ	-0.01	0.00	-2.23	0.026
Duration	0.51	0.10	4.90	<.001

This model was fitted with a logit-link function, participant response (1 = acted) as the binary output variable, and random intercepts for participant and stimulus.

Table A3

Factors affecting the reaction time to a stimulus on accurate-response trials, based on 4,674 trials across 129 participants

	Coefficient	*SE*	Wald statistic (χ^2)	*p*
Minimal model				
Intercept	1250.54	94.35		
Sex	499.03	160.34	3.11	0.004
Vocalizer Sex	58.46	47.16	1.24	0.759
Language	523.59	167.94	3.14	0.002
Duration	-140.46	21.71	-6.47	<0.001
Sex*Vocalizer Sex	-176.54	63.13	-2.80	0.005

This model was fitted linearly with reaction time as the output variable and random intercepts for participant and stimulus.