Published in final edited form as:

Emotion. 2014 April ; 14(2): 251–262. doi:10.1037/a0036052.

Perceptions of Emotion from Facial Expressions are Not Culturally Universal: Evidence from a Remote Culture

Maria Gendron¹, Debi Roberson², Jacoba Marietta van der Vyver^{3,4}, and Lisa Feldman Barrett^{1,4,5}

¹Affective Science Institute and Department of Psychology, Northeastern University

²Department of Psychology, University of Essex

³School of Nursing and Public Health, University of Namibia

⁴Martinos Center for Biomedical Imaging, Massachusetts General Hospital

⁵Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School

Abstract

It is widely believed that certain emotions are universally recognized in facial expressions. Recent evidence indicates that Western perceptions (e.g., scowls as anger) depend on cues to US emotion concepts embedded in experiments. Since such cues are standard feature in methods used in cross-cultural experiments, we hypothesized that evidence of universality depends on this conceptual context. In our study, participants from the US and the Himba ethnic group sorted images of posed facial expressions into piles by emotion type. Without cues to emotion concepts, Himba participants did not show the presumed “universal” pattern, whereas US participants produced a pattern with presumed universal features. With cues to emotion concepts, participants in both cultures produced sorts that were closer to the presumed “universal” pattern, although substantial cultural variation persisted. Our findings indicate that perceptions of emotion are not universal, but depend on cultural and conceptual contexts.

Keywords

emotion; perception; facial expression; universality; relativity; culture

One of the most widely accepted scientific facts in psychology and human neuroscience is that there are a small number of emotions (anger, sadness, fear, etc.) that are expressed on the face and universally recognized (Ekman, 1972; Ekman & Cordaro, 2011; Izard, 1971, 1994; Matsumoto, Keltner, Shiota, Frank, & O'Sullivan, 2008; Tomkins, 1962, 1963). It is further assumed that since people all around the world can automatically recognize emotions in these expressions (e.g., scowls are recognized as anger, pouts as sadness, and so on),

Correspondence concerning this article should be addressed to Maria Gendron, 125 Nightingale Hall, Northeastern University, Boston, MA 02115, m.gendron@neu.edu.

M.G., D.R., J.M.V. & L.F.B. contributed to design, M.G., D.R. & J.M.V. contributed to data collection, M.G. & L.F.B. analyzed the data, M.G. and L.F.B. drafted the paper, and D.R. provided critical revisions. All authors approved the final version of the paper for submission.

facial expressions transcend concepts for emotion and the language that we use to refer to them, such that emotion “recognition” ability is pre-linguistic (Izard, 1994) and words are simply a medium for communicating already formed perceptions. We refer to this as the *universality hypothesis*. This view is standard curriculum in introductory psychology, is used by government agencies to train security agents (Hontz, 2009; Weinberger, 2010) and is a popular example when social science is communicated to the public (e.g., stories in National Geographic, Radiolab, etc.).

There are thousands of experiments that casually claim that emotions are universally perceived from the face, based on hundreds of studies that performed cross-cultural comparisons (e.g., comparing US and Japanese perceivers; for a meta-analysis see Elfenbein & Ambady, 2002) appearing to support this claim. Yet only a handful of studies – those that compare individuals from a Western cultural context (e.g., the US) to perceivers from remote cultural contexts with little exposure to Western cultural practices and norms – provide a strong test of whether people around the world *universally* recognize emotions in facial expressions. The strength of this “two-culture” approach is that cultural similarities cannot be attributed to cultural contact or shared cultural practice, suggesting that similarities are, in fact, due to the presence of a psychological universal (Norenzayan & Heine, 2005). Only six published experiments test universality of emotion perception using a two-culture approach (two in peer review outlets) and all but one were conducted nearly 40 years ago (see Table S1 in the Supplemental Material available online; Ekman, 1972; Ekman & Friesen, 1971; Ekman, Heider, Friesen, & Heider, 1972; Sorensen, 1975).

In addition, four of the experiments, (Ekman, 1972; Ekman & Friesen, 1971; Ekman, Heider, Friesen, & Heider, 1972) as well as virtually all of the studies that test participants that are not isolated from Western culture (see Elfenbein & Ambady, 2002),¹ contain an important methodological constraint – they ask perceivers to match a facial expression to an emotion word or description (i.e., they include a conceptual context in the experimental method). By examining Table S1, we can see that it is only the experiments that contain this conceptual context that produce robust evidence of universality. For example, participants might be presented with an orally translated story, such as “He is looking at something which smells bad” (describing a disgust scenario; Ekman & Friesen, 1971, p. 126) and asked to select the matching expression from two or three options. Yet, the two studies using isolated samples that did not provide a conceptual context (e.g., asking participants to freely label posed facial expressions; Sorensen, 1975) did not find evidence that emotion is perceived universally in facial expressions.

Recent experiments conducted with Western samples provide additional evidence for the presumed universal pattern of emotion perception is dependent on conceptual constraints. For example, studies that temporarily reduce the accessibility of emotion concept knowledge also impair emotion perception, such that the presumed universal pattern of emotion perception is not even obtained in a sample of homogeneous US undergraduate students (Gendron, Lindquist, Barsalou, & Barrett, 2012; Lindquist, Barrett, Bliss-Moreau, &

¹The forced-choice method in which a set of emotion word choices are provided to perceivers in emotion perception tasks is so prevalent that this meta-analysis was unable to use this as an attribute in their analysis.

Russell, 2006b; Roberson, Damjanovic, & Pilling, 2007; Widen, Christy, Hewett, & Russell, 2011). These studies indicate that evidence for “universality” appears to be conditional on the experimental methods being used (cf. Russell, 1994), pointing towards a more nuanced model of emotion perception in which culture and language are key to constructing emotional perceptions.

Our psychological constructionist model of emotion hypothesizes that emotion perception is dependent on emotion concepts that are shaped by language, culture, and individual experience (Barrett, 2009; Barrett, Lindquist, & Gendron, 2007; Barrett, Mesquita & Gendron, 2011; Barrett, Wilson-Mendenhall, & Barsalou, in press; Lindquist & Gendron, 2012). This model predicts that perceptions of discrete emotion are *unlikely* to be consistent across widely distinct cultural contexts because language, cultural knowledge, and situated action will all exert different influences on emotion perception. In our view, emotion categories themselves are flexible and embodied, and grounded in knowledge about situated action (Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011). Variability in perceptions of emotion even *within* a cultural and linguistic context is predicted by the flexible nature of emotion concepts. Yet some commonalities in perceptions of emotion across cultures are predicted to the extent that there are similar patterns of situated actions across cultural contexts. This type of cultural consistency is likely to be particularly evident in experimental tasks that explicitly invoke knowledge of situated action, such as “He is looking at something which smells bad” (Ekman & Friesen, 1971, p. 126), as is common in the previous universality literature. This stands in contrast to the adaptationist “basic emotion” approaches where perceptions of emotion are thought to be the product of a reflexive ability to “decode” non-verbal “signals” embedded in the face, voice and body, free from learning or language (e.g., Izard, 1994). As a consequence, in this “basic emotion” view, shared-culture, language and context should all be unnecessary to establish robust cross-cultural emotion perception.

While the emerging empirical picture challenges the widespread conclusion that the “universal” pattern of discrete emotion perception is achieved on the basis of the perceptual features of faces (and other non-verbal cues) alone, critical gaps in the literature still exist. No single study has explicitly manipulated the presence versus absence of emotion concept knowledge in an emotion perception task and examined the consequence for emotion perception in both a Western sample and in a sample from a remote cultural context. This paper fills that gap. We examined emotion perception in members of the Himba ethnic group, who live in the remote villages within the Keunene region of Northwestern Namibia and have relatively little contact with people outside of their immediate communities (and therefore limited exposure to emotional expressions outside their own cultural context). Also key, Himba individuals speak a dialect of Otji-Herero that includes translations for the English words “anger”, “sadness”, “fear”, disgust”, “happy” and “neutral” (Sauter, Eisner, Ekman, & Scott, 2010), allowing us to manipulate the presence or absence of similar concepts across cultures.

To manipulate the presence or absence of emotion concept information, we started with a relatively unconstrained perception task, where participants were asked to sort photographs of posed portrayals of emotion into piles. Half of the participants were asked to structure

their sort based on a set of emotion words, whereas the other half of participants freely sorted the faces. If scowling faces are universally perceived as angry, pouting faces as sad, smiling faces as happy, and so on, then participants in both cultures would freely sort the facial portrayals into six piles (based on the perceptual regularities in the stimulus set) and emotion concepts should not be necessary to observe this pattern. A strict universality view would predict that there should be strong consistency across cultures with virtually no variation across cultural contexts. More recent accounts that assume variability across cultures but a core innate mechanism for the production and perception of discrete emotions (e.g., dialect theory; Elfenbein, 2003, 2007; Tracy & Sharif, 2011) would predict only higher than chance consistency across the cultural contexts, even if there are significant differences in what they term “recognition accuracy”. Of course, rejecting the null hypothesis (no consistency at all) is not evidence for a discrete emotion universalist hypothesis; it would also be necessary to rule out the possibility that any consistency in perception could not be explained as universal perception of the affective properties of the faces (e.g., valence or arousal; cf. Russell, 1994), or as situated action.

Such a view, called *minimal universality* (Russell, 1995) is consistent with our constructionist approach, where affective events are perceived as emotions using emotion concepts and related emotion language; our view hypothesizes strong cultural variability in emotion perception. Specifically, we hypothesized that if emotion words impose structure on emotion perception, as suggested by recent research (Gendron et al., 2012; Lindquist, Barrett, Bliss-Moreau, & Russell, 2006a; Widen et al., 2011), then both Himba and US participants would differ in perceptions (measured as their sorting behavior) depending on whether participants were sorting into piles anchored by emotion words or not. We expected that US anchored sort responses would better resemble the expected “universal” solution when compared to US free sort responses. Furthermore, if emotion concept words are linked to universal perceptual representations (e.g., furrowed brow and pressed lips in *anger*) across distinct cultural contexts, then Himba participants would also produce a more “universal” (i.e., Western) looking solution in the anchored sort procedure. If the perceptual representations linked to emotion concept words vary across distinct cultural contexts as recent evidence suggests (Jack, Garrod, Yu, Caldara, & Schyns, 2012), however, then Himba participants’ sorting of facial expressions should differ with the introduction of emotion words, but would not be clearly in line with the presumed universal (i.e., Western) cultural model.

Methods

Participants

Participants were 65 individuals (32 male, 33 female; mean age= 30.84, SD=13.04) from the Himba ethnic group of Otji-Herero speakers and 68 individuals (30 male, 38, female; mean age=38.27; sd=12.77) from an American community sample. 11 Himba participants were excluded from analysis (resulting in a final sample of 54 participants) based on compliance (i.e., 4 participants decided not to finish sorting the stimuli), failure to sort into piles by emotion (2 participants) and feasibility (3 participants—one with low visual acuity, two with an inability to sort the stimuli at all and low forward digit span in a separate experiment). In

addition, 2 participants' data were dropped from further analysis due to data recording issues that were discovered during data entry (i.e., not all image placements were recorded).

Individuals from the Himba ethnic group constitute a strong sample for the two-culture approach (Norenzayan & Heine, 2005) to test the universality of emotion perception. Most of the Himba ethnic group lives in an ancestrally "traditional" culture, such that they do not take part in their country's political or economic systems, but rather live as semi-nomadic pastoralists, tending to herds of goats and cattle. In our study, Himba participant data was collected in two remote villages, both located in the Mountainous Northwest region of Namibia, Kunene.² The first village contained a mobile school as well as a missionary church; all members of the village otherwise lived traditionally and did not show significant signs of Westernization. The second village did not contain an established school or outside presence in the community. Both locations were far from the regional towns, and there was no evidence of tourism to these communities as is typical of "show" villages closer to the main regional town of Opuwo. While several of the participants (particularly in location 1) had some conversational English (greetings), none of the participants were English language speakers. All participants in this sample were native speakers of the Otji-Herero language.

Our US sample was tested at the Boston Museum of Science in the Living Laboratory environment (<http://www.mos.org/discoverycenter/livinglab>). One participant's data were dropped from further analysis due to data recording issues that were discovered during data entry (i.e., not all image placements were recorded). The Living Laboratory is a well-suited site for the collection of control data given the hectic environment that mirrors the busy and often social context of testing out in the field. In addition, the Living Laboratory affords the collection of data from a community sample, such that individuals from a range of ages and backgrounds can be included.

Stimuli

Stimuli were 36 4×6 cards containing photographs of facial expressions of emotion posed by African American individuals. Stimuli were selected based on the following criteria. (1) The facial structure of the identities included in the stimulus set were rated as closest to that of several example within gender identities from the Himba ethnic group (as judged by a set of 120 perceivers from a largely Western cultural context using Amazon's Mechanical Turk system); the final set of face stimuli (three male and three female identities) were selected from a larger set of 23 male and 20 female identities from Gur (Gur et al., 2002), IASLab (www.affective-science.org), and NimStim (<http://www.macbrain.org/resources.htm>) face sets. (2) A given identity posed a facial expression depicting anger, fear, sadness, disgust, neutral affect, and happiness, with a final stimulus set of 36 faces. Each face stimulus was edited in Adobe Photoshop such that visual background information was removed. A uniform white collar (as is found on NimStim identities) was added to all photographs in order to make the stimuli uniform and remove variation that might distract from facial actions. (3) Additional norming conducted on Amazon's Mechanical Turk confirmed that

²Studying emotion perception in the Himba is additionally important given a recent paper concluding that Himba and British individuals have universal emotion perception in non-word human vocalizations (Sauter et al., 2010). Participants in our study were from different villages than those tested in prior research on emotion perception with the Himba.

the face stimuli were indeed rated as depicting the expected emotions, with only a few exceptions (Table S2 and Table S3 in the SOM).

Procedure

All participants were consented prior to performing the experiment. Participants from the Himba ethnic group were verbally consented with the use of a translator. Participants within each culture were randomly assigned to either the free sorting or anchored sorting condition. In the free sorting condition, participants were instructed to sort the faces by feelings so that each person in a given pile was experiencing the same emotion. Participants were allowed to create as many piles as needed with unlimited time. The participant was then allowed to freely sort the images into piles. Participants were reinstructed if initial sorting appeared to be based on identity, rather than emotion (this took place in less than 25% of Himba participants). All but 2 participants were able to sort by emotion following reinstruction.

Participants in the word-anchored sort condition heard slightly modified instructions where they were told that they might find certain emotions on the faces of people. The labels for those emotions (“anger”, “fear”, “disgust”, “sadness”, “happy”, and “neutral” or their translation in the Himba dialect of Herero as “okupindika”, “okutira”, “okujaukwa”, “oruhoze”, “ohange” and “nguri nawa”, respectively) were then provided.³ Label anchors were delivered verbally to all participants, rather than written, because the Himba participants we tested are from a pre-literate culture. To ensure that the emotion concepts remained accessible throughout the task, participants were reminded of the six emotions every 6 cards that they placed down. If the participant initially looked through the stimuli before beginning to sort, they were also reminded every six stimuli that were closely examined by the participant. For American participants, a few modifications to the anchoring protocol were set in place based on piloting. American participants were given the option of waiving further repetition of the list of emotion terms if they 1) indicated that they did not wish to hear it again and 2) were able to repeat the list back to the experimenter without error.

Once sorting was complete, all participants (regardless of culture or task version) were asked to provide a label to describe the content of each pile. Labels were directly translated into English by an interpreter who worked on a prior study with Himba participants supporting the universality hypothesis (Table S1, Row 3; Sauter et al., 2010). This question was initially phrased in an open-ended manner (i.e., “What is in this pile?”) in order to minimize expectancy effects regarding what type of content to produce. The experimenter recorded any behavioral descriptors or mental state descriptors that the participant produced. All participants were also asked specifically for facial expression and emotion labels that were not spontaneously provided by participants during open-ended questioning. Since many Himba participants failed to provide additional information in response to these

³The translation of English-language emotion terms into Otji-Herero was originally achieved in Sauter et al (2010) and the same translator was used in our own work. Kemuu Jakara conducted extensive translation work in the area, with over 12 years of experience working as a guide and translator for various academics. In addition, the translations that were provided were consistent with provided translations in the newly developed Otji-Herero/English dictionary by Nduvaa Erna Nguaike.

prompts, prompted responses (both for American and Himba samples) were not subjected to further analysis.

Results

New evidence of cultural relativity rather than universality in emotion perception

Labeling Analysis—The words that participants freely offered to name their piles indicated, as we predicted, that Himba participants did not perceive the posed, facial expressions according to the “universal” pattern (Figure 1). As expected, US participants tended to use discrete emotion words such as “sadness” or “disgust” to label their piles ($M = 4.45$, $SD = 1.091$) more so than did Himba participants ($M = .81$, $SD = .939$), $t(55) = 2.644$, $P < .001$. Overall use of each discrete emotion term by Himba and US participants can be found in Table S4.

Interestingly, US participants more generally named their piles with a wide array of additional mental state words such as “surprise” or “concern” ($M = 2.32$, $SD = 1.447$) at greater frequency than did Himba participants ($M = .69$, $SD = .838$), $t(55) = 1.630$, $P < .001$. In contrast, Himba participants were more likely to label face piles with descriptions of physical actions such as “laughing” or “looking at something” ($M = 2.38$, $SD = 1.098$) when compared to US participants ($M = .84$, $SD = 1.098$), $t(55) = -1.546$, $P < .001$, suggesting that individuals in Himba culture used *action descriptors* more frequently than *mental state descriptors* to convey the meaning of facial actions. Indeed, extensive research on Action Identification Theory (Vallacher & Wegner, 1987) indicates that it is possible to describe another person’s actions in either physical or mental terms (Kozak, Marsh, & Wegner, 2006). For example, perceiving a face as “fearful” requires mental state inference because an internal state is assumed to be responsible for the observable “expression”; this stands in contrast to perceiving actions on the face in more physical (and less psychological) terms (e.g., looking) that might be observed in any number of emotional or non-emotional instances (e.g., see Tables S5, S6 in the SOM).

Cluster analysis—Using a cluster analysis approach, we examined whether participants’ sorts mapped on to the Western discrete emotion categories from which the photographs were sampled (i.e., *anger*, *fear*, *sadness*, *disgust*, *neutral* and *happiness*). A hierarchical cluster analysis (Sokal & Michener, 1958) produces a set of nested clusters organized in a hierarchical tree. Unlike other clustering procedures (e.g., k-means; Hartigan & Wong, 1979), the number of clusters can be discovered rather than being pre-specified. We used an agglomerative approach, starting with each item as its own cluster and progressively linking those items together based on an estimate of their distance from one another (computed from the number of times face stimuli appeared in the same vs separate piles; see description of co-occurrence matrix below). We employed a cluster distance measure of average linkage because it balances the limitations of single and complete linkage methods, which can be influenced by noise and outliers or force clusters with similar diameters, respectively (Aldenderfer & Blashfield, 1984). The average linkage clustering method uses information about all pairs of distances to assign cluster membership, not just the nearest or the furthest item pairs. We performed a cluster analysis on the data from each cultural group separately,

further broken down by sorting condition, resulting in four different MDS solutions: Himba Free Sort, Himba Anchored Sort, US Free Sort, and US Anchored Sort.

To accomplish the cluster analysis, for each of the four conditions, we computed a cooccurrence matrix (Coxon & Davies, 1982). Each co-occurrence matrix contained a row and column for each of the 36 items in the set, resulting in a 36×36 symmetric matrix. Each cell in the matrix represented the number of times a given pair of face items was grouped by participants into the same pile (i.e., across participants). The larger the number in a cell, the more frequently those two items co-occurred, and thus the higher perceived similarity between those items at a group level. We then converted this co-occurrence similarity matrix into a distance matrix, where a higher cell value was an indication of less similarity between items. The cluster analysis was then performed on each dissimilarity matrix and we examined the resulting dendrogram for each (Figure S1 & Figure S2). Based on large within-cluster average item distance when the solution contained less than six clusters, (i.e., as items were grouped into larger, increasingly inclusive clusters, the clusters became less coherent; see dendrograms in Figure S1 in the SOM), and for theoretical reasons (since 6 discrete expression portrayal types were included in the set), we report the results of the six-cluster solutions.

Using cluster analysis, we found, as expected, the US participants (Figure 2A) grouped together smiling (happy), scowling (angry), wide-eyed (fearful) and neutral faces into distinct piles. The resulting clusters did not recapitulate separate categories for pouting (sad) and nose-wrinkled (disgusted) faces, indicating that these photographs were not sorted by US participants into their own piles.⁴ Himba participants, in contrast, (Figure 2B) grouped smiling (happy) faces together, wide-eyed (fearful) faces together. Resulting clusters did not recapitulate separate categories for scowling (angry), pouting (sad), and nose-wrinkled (disgusted) faces, indicating that these items were not sorted into their own piles. Furthermore, when compared to the Himba sample, the six-cluster solution provides a better “fit”, as demonstrated graphically by the shorter bracket lengths in the dendrogram (Figure S1) for the US sorts.

Multidimensional scaling—While the cluster analyses reveal whether the emotion perception sorts from each culture neatly conform to the presumed universal solution, they do not reveal the underlying properties that are driving sorting in each culture. Posed expressions could be placed into the same pile because they share an emotional meaning (e.g., fearful) or because they share a behavioral meaning (e.g., looking). This is a particularly relevant distinction given that the majority of Himba participants identified their pile contents with behavioral and situational descriptors rather than with mental state descriptors. We conducted multidimensional scaling (MDS) analyses because this approach can reveal a cognitive map of the sorting pattern for each group of participants, along with the underlying dimensions that best represent how participants perceived the similarities and differences among the face stimuli (Kruskal & Wish, 1978).

⁴Importantly, these piles are not accounted for by the items included in the set that were relatively “poor” exemplars. Four images, three of which were portrayals of anger and one portrayal of fear had forced-choice ratings that diverged from the expected category. Yet the cluster that most clearly diverged with the “discrete emotion” and presumed universal model was comprised of disgust and sadness portrayals.

Multidimensional scaling (MDS) analyses of the free sort data (Figure 3A, B) confirmed that Himba participants possessed a cognitive map of the posed facial expressions that was anchored in action identification, whereas the US participants' cognitive map also contained information about mental state inferences of discrete emotions. To conduct the MDS analyses, each co-occurrence matrix (i.e., the similarity matrix used in the hierarchical cluster analysis) was also subjected to an ALSCAL multidimensional scaling procedure (Young & Lewycky, 1979). MDS provides an n-dimensional map where the dimensions themselves can reveal the properties that perceivers use to structure their sorts. Based on the stress-by-dimensionality plots for solutions between one and six dimensions (Figure S2), a four dimensional solution was selected as the best fit for the two US data (free sorting and anchored) and a three dimensional solution was selected as the best fit for the two Himba data (free sorting and anchored).

To empirically identify the MDS dimensions (rather than to label them subjectively as most experiments do), we opted for an empirical approach (as in Feldman, 1995), in this case using hierarchical regression to identify the dimensions. To empirically identify the MDS dimensions, we required that the face stimuli be normed on a variety of properties. Fourteen independent samples of participants rated the 36 images in the stimulus set to estimate the extent to which each face depicted a given behavior (crying, frowning, laughing, looking, pouting, scowling, smelling, and smiling; Table S5) or a given emotion as a mental state (anger, fear, disgust, sadness, neutral and happiness; Table S2). Each sample of 40 participants rated the 36 faces on one property only, resulting in a total of 560 participants. Ratings for each property were on a Likert scale ranging from 1 "not at all" to 5 "extremely". The images were rated by participants recruited on Amazon's Mechanical Turk (Buhrmester, Kwang, & Gosling, 2011) and restricted to IP addresses located in North America. Since our experiments on Mechanical Turk would not allow for us to collect responses by Himba individuals, we decided to limit our normative ratings to the US cultural context. This allows us to test the applicability of the US cultural model of emotions (captured by these ratings) to both US and Himba cultural contexts. We computed the mean property rating for each image resulting in 14 attribute vectors with 36 means, which we then used to empirically identify the MDS dimensions (following Feldman, 1995). In a series of hierarchical multiple regression analyses, we examined the extent to which behavior and emotion vectors explained variance in the MDS dimension coordinates for each dimension. By entering behavior vectors as the first block of predictors, and the emotion vectors as the second block, we were able to estimate the extent to which each dimension captured mental state (emotion) inferences (e.g., sadness) over and above merely describing the physical action of a face (e.g., pouting). Following action identification theory (Kozak, Marsh, & Wegner, 2006; Vallacher & Wegner, 1987), behavior vectors were always entered as the first block of predictors, and emotion vectors as the second. We performed a series of F-tests (Table S7) examining whether the behavior vectors (see Table S5), sufficiently explained a given MDS dimension, and whether discrete emotion vectors (see Table S2) accounted for that dimension over and above behaviors. A summary of the dimension variance accounted for by the behavioral and emotion vectors is presented in Table 1. We predicted, and found, that the MDS dimensions for the Himba solution reflected behavioral distinctions (e.g., smiling, looking), whereas the dimensions for the US solution

largely reflected mental state distinctions (over and above behavior; see Table 1 for summary, Table S7 for F-tests).

Taken together, MDS solutions for Himba and US free sort data confirmed strong cultural differences in emotion perception. It is unlikely that these findings were simply due to “poor” task-related performance in Himba participants, who have performed well on other psychological testing with significant demands (de Fockert, Caparos, Linnell, & Davidoff, 2011). Furthermore, stress was relatively low for the MDS solutions themselves and the R^2 values for identification of the MDS dimensions were large for Himba participants’ data (R^2 values ranged between .597 and .865), indicating that there was robust and meaningful group-level consistency in how Himba participants sorted the face stimuli.

Emotion words impact emotion perception in a culturally relative manner

As predicted, the introduction of emotion word anchors prior to and during sorting influenced both US and Himba participants’ perceptions of the facial expressions. Participants from the US (and to a limited extent, Himba participants) conformed more to the presumed universal discrete emotion pattern, suggesting that it is emotion concepts that help to structure perceptions into discrete categories. The hierarchical cluster analysis produced a much clearer “universal” pattern for the US participants in the conceptually anchored condition. Each of the six clusters appeared to represent a distinct discrete emotion category, with nose-wrinkled (disgust) and pouting (sadness) expressions now appearing in separate clusters (Figure 2C, clusters 4 and 6, respectively) compared to the free sorting solution where those expressions ended up in the same cluster (Figure 2A, cluster 5). For Himba participants, conceptually anchored sorting did not produce a cluster solution that recapitulated the discrete emotion portrayals embedded in the set. Only two of six clusters (see Figure 2D) contained a predominant emotion portrayal type: Cluster 1 (Figure 2D) contained only smiling faces and cluster 3 (Figure 2D) contained only wide-eyed faces. While at first blush this might appear to provide support for the universal perception of happiness and fear, an alternative interpretation is that there was universal perception for valence and arousal, respectively. For example, fear expressions are higher in arousal level than the other expressions included in the set and happy expressions were the only positive expressions included in the set. This is only speculation, however, because the stimulus set did not include other posed, positive expressions or surprised expressions which are equivalently high in arousal to fear expressions, which would be needed to more clearly test between discrete emotion vs affect views. There was also a difference in the composition of Cluster 2 (the predominantly neutral cluster) of Himba participants between the free and conceptually anchored conditions, with the cluster becoming more “inclusive” in the conceptually anchored condition, such that there were a number of negative emotion portrayals grouped with the neutral portrayals.

MDS analyses more clearly revealed how emotions words influenced the sorting behavior of Himba participants. Specifically, we found that facial expressions depicting emotion categories were more tightly clustered together in multidimensional space for Himba participants exposed to emotion word anchors compared to Himba participants who freely sorted the faces into piles (Figure 3A, C); this increase in clustering was most apparent for

smiling and neutral faces. For US participants (Figure 3B, D), there was little change in the distances between within-category facial expressions.

As in the free sorting data, the MDS dimensions anchoring the US cognitive map were identified in emotion terms, reflecting that US participants were making mental state inferences about the facial expressions (see Table 1, Table S7 for F-tests in the SOM). The Himba cognitive map for the anchored sort data was described more in terms of mental states when compared to the map representing the freely sorted faces. One of the three MDS dimensions was empirically identified in emotion terms, and two remained in behavioral terms (although, one of these dimensions was trending for emotion; $P = .064$). Taken together, these data indicate that emotion words have a powerful effect on emotion perception, even under relatively unconstrained task conditions.

Computing “Accuracy”

In typical studies of the universality hypothesis, researchers compute the extent to which participants choose a word or description for each facial expression, or freely label that expression, in a way that matches the experimenter’s expectations, and this is reported as “recognition accuracy”. We were not able to compute a similar “recognition accuracy” score because we do not have a model of what “chance” responding might look like in the sorting task. Nonetheless, we were able to compute a “consistency” score as the number of faces from a given emotion category that comprised the “dominant” content in a given pile. This score allowed for items from the same category to be broken up into multiple piles and also allowed singletons (i.e., images that are placed alone and not grouped with other items) to be counted as “consistent”. “Consistency”, then, was defined as the percentage of items representing a discrete emotion category that were grouped together and made up the majority of the content in a given pile.

A mixed-model analysis of variance, with “consistency” as the dependent variable, emotion category as the repeated measure (portrayals of anger, fear, sadness, disgust, neutral and happiness) and cultural group as the between participants factor (American or Himba) produced a main effect for cultural group, such that Himba participants had lower “consistency” for sorting discrete emotion faces together than did US participants, with a mean “consistency” score of 49.615% ($SD=33.559$) compared to US participants’s mean of 67.101 % ($SD=22.350$), ($F_{1,56}=53.412$, $P < 0.001$; see Table S8). Lower discrete emotion sorting “consistency” was found regardless of emotion category. In both Himba and US cultures, discrete sorting “consistency” was marginally higher for participants exposed to words ($M=58.576$, $SD=24.669$) than participants who were not ($M=54.223$, $SD=25.158$), ($F_{1,118} = 3.422$, $P = 0.067$; for means see Table S8). The three way interaction was not statistically significant, ($F_{5,280}=0.816$, $P = 0.510$, Greenhouse-Geisser corrected)

The overall level of consistency that we computed in this study should not be compared to those obtained in traditional emotion perception tasks (which are provided for interested readers in Table S1), because of computational differences between these scores and traditional recognition accuracy scores. In addition, “consistency” is not the only way that an approximate “accuracy” score can be computed. For an additional approach called “discrimination” see supplementary materials.

Discussion

The Western cultural script for what it means to be human includes the ability to “recognize” emotional expressions in a universal way. We show that small changes in experimental procedure disrupt evidence for universal emotion perception. By comparing emotion perception in participants from maximally distinct cultural backgrounds, US participants and Himba participants from remote regions of Namibia with limited exposure to Western culture, we demonstrated that facial expressions are not universally “recognized” in discrete emotional terms. Unlike prior experiments, we used a face-sorting task that allowed us to manipulate the influence of emotion concepts on how the faces were perceived. Without emotion concepts to structure perception, Himba individuals perceived the facial expressions as behaviors that do not have a necessary relationship to emotions, whereas US participants were more likely to perceive the expressions in mental terms, and as the presumed universal emotion categories. When words for emotion concepts were introduced into the perception task, participants from both cultures grouped the facial expressions in a distinct manner. For US participants, it was in a way that better resembled the presumed “universal” solution. For Himba participants, the difference between free and conceptually anchored solutions was subtle. Part of the sorting pattern of Himba participants was better accounted for by mental state ratings, consistent with the view that the presence of discrete emotion words produces emotion perception closer to the presumed “universal” model. Yet, the data did not neatly conform to this presumed universal model. This nuanced pattern of effects is consistent with recent evidence demonstrating that people from different cultures vary in their internal, mental representations of emotional expressions, even for emotion terms that have available translations across cultures (Jack et al., 2012). That is, linguistically relative emotion perception manifests in a culturally relative way.

The present findings also advance our understanding of what aspects of perception may be consistent across cultures, what Russell (1995) has referred to as sources of *minimal universality*. In our data, smiling faces, as well as wide-eyed faces, were clustered into distinct piles by perceivers from both Himba and US cultural contexts. Thus, it is possible that perceivers in both cultural contexts understood those same perceptual features as indicative of the same discrete emotional meaning. This would be consistent with a view in which smiles and widened eyes are signals of “basic” and universally endowed states of happiness and fear, respectively. Notably, this would constitute a much more limited set of universal expressions than hypothesized in most universality views. An alternative explanation, however, is that observed cross-cultural consistency in sorting behavior is driven by the universal ability to perceive some other psychological property, such as affect. The present task was not designed to test for this type of cross-cultural consistency since examining affect perception would require a set of stimuli including multiple portrayal types for each quadrant of affective space (i.e., positive low arousal, positive high arousal, negative low arousal, negative high arousal). In our experiment, smiling faces were the only positive affective expression available to present to participants. In a separate line of work with the Himba culture, we did test this hypothesis using vocalizations, since there were multiple exemplars from the different quadrants of affective space that were available. In both free-labeling and forced-choice experimental designs, we found evidence that Himba

participants perceived affect in Western style vocal cues, with the most robust cross-cultural perception for the valence (positive and negative) dimension (Gendron, Roberson, van der Vyver & Barrett, 2013).

Another possibility is that both cultures were able to consistently perceive situated action. Two results are consistent with this view. First, the free-labeling responses of participants from Himba culture revealed that posed facial expressions were understood in terms of situated actions—e.g., “smiling” and “looking at something”. Second, external perceiver ratings of faces based on situated actions (i.e., “smiling”, “looking”, etc.) were better predictors of how the faces were sorted by Himba perceivers than were external ratings of discrete emotional states, and these dimensions, while not sufficient to explain US sorts, appeared to be important to them. Furthermore, neuroimaging evidence indicates that during instances of emotion perception, both the “mirroring” network (important to action-perception) and “mentalizing” networks (also called the theory of mind network, the semantic network, or the default mode network; Barrett & Satpute, 2013; Lindquist & Barrett, 2012) are engaged (Spunt & Lieberman, 2012). Thus consistency across Himba and US participants sorting may have been driven by similar behavioral understanding of those facial muscle configurations as “smiling” and “looking” even if a mental state representation of “happiness” and “fear” additionally occurred for US participants. Alternatively, it is possible that emotion concepts were implicated for both cultures, but that these concepts differed in the extent to which they are abstract concepts with mentalizing properties (and requiring language), as opposed to more concrete concepts that are anchored in action identification. This is a matter for future research.

Our findings underscore that facial actions themselves, regardless of their emotional meaning, can be meaningful as instances of situated action, an observation that is receiving increasing attention in the literature. For example, recent work has argued that the wide eyes of a posed fear expression evolved to increase sensory intake (Susskind, Lee, Cusi, Felman, Grabski & Anderson, 2008), whereas the nose wrinkle of disgust evolved from oral-nasal rejection of sensory input (Chapman, Kim, Susskind & Anderson, 2009). It is consistent with a psychological construction approach to hypothesize that specific facial actions evolved to serve sensory functions, and later acquired social meaning and the ability to signal social information (Barrett, 2012). Although our findings do not speak to this hypothesis directly, they are consistent with the idea that facial actions are not arbitrary and instead may reflect the type of situated action that occurs during episodes of emotion. Yet as we have seen in our norming data, facial actions and emotions are *not* redundant for US perceivers in emotion perception, pointing to the importance of considering them separately in any discussion of evolution.

Our findings also help to reveal the limited ability of traditional “accuracy” approaches to quantify meaningful cultural variation in emotion perception. These results simply compare the task performance of participants to a single model of emotion perception. Our own computed consistency and discrimination scores which are based on the “fit” of the discrete emotion model to sorting behavior did not reveal the shifts in sorting that accompanied the presence of emotion words that were revealed by cluster and MDS analyses. In particular, our multidimensional scaling approach represents a significant innovation in modeling

cultural variation in emotion perception because it allowed us to test multiple explanations for the sorting behavior in each cultural context. With this approach, in conjunction with the explicit labeling data, it was revealed that individuals from Himba culture rely on behavioral, rather than mental categories to make meaning of facial behaviors.

Our findings are consistent with a growing body of evidence that emotions are not “recognized” but are perceived via a complex set of processes (Barrett, Mesquita, & Gendron, 2011) that involve the interplay of different brain networks, such as those that support action identification and mental state inference (Spunt & Lieberman, 2012; Zaki, Hennigan, Weber, & Ochsner, 2010). Our finding that the presumed universal pattern of emotion perception appears to be linguistically relative is consistent with the pattern of published results (see Table S1), as well as our own laboratory research. In prior work, we demonstrated that experimentally decreasing the accessibility of emotion words’ semantic meaning, using a procedure called “semantic satiation” (Tian & Huber, 2010), reduces the accuracy with which participants produce the presumed universal pattern of emotion perception (Lindquist et al., 2006a) because words help to shape the underlying perceptual representation of those faces (Gendron et al., 2012). Our current findings are consistent with research on patients with semantic deficits due to progressive neurodegeneration (i.e., semantic dementia) or brain injury (i.e., semantic aphasia), who do not perceive emotions in scowls, pouts, smiles and so on (Calabria, Cotelli, Adenzato, Zanetti, & Miniussi, 2009; Roberson, Davidoff, & Braisby, 1999). Even research in young children points to the importance of emotion words in emotion perception, because the presumed universal pattern of emotion perception emerges in young children as they acquire conceptual categories, anchored by words, for emotions (Widen & Russell, 2010). Taken together, these findings challenge the assumption that facial expressions are evolved, universal “signals” of emotion, and instead suggest that facial expressions are culturally-sensitive “signs” (Barrett, 2012; Barrett et al., 2011). These findings have implications for the broad-scale adoption of a set of canonical, caricatured facial expressions into other disciplines such as cognitive neuroscience and clinical applications that has taken place over the last several decades. To properly understand how humans perceive emotion, we must attempt to move past the Western cultural model as the assumed model for people everywhere.

The present findings are not readily explained by a universality account of emotion recognition, even those that admit some minor cultural variability. It has been suggested, for example, that in addition to universal recognition abilities cultures have “display rules” (Ekman, 1972) that allow people to regulate their expressions. Cultures are also thought to have “decoding rules” (Matsumoto & Ekman, 1989) that govern how people report on their perceptions of emotion to maintain a culturally appropriate response (e.g., in Japanese culture, individuals will discount the intensity of emotion perceived and report that another person is feeling less intensely). Neither display nor decoding rule accounts have the predictive power in explaining the language-based effects we observed in the present study, or the publishing findings demonstrating that, in US samples, the presumed universal solution can be disrupted by interfering with emotion word processing. Display and decoding rules, as instantiations of culture, are relatively rigid and would not predict that emotion perceptions vary based on momentarily accessible concepts. Further, our findings fit with other growing evidence that emotions such as anger and fear are not natural kinds,

particularly in light of emerging evidence that central (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012) and peripheral (Kreibig, 2010) measurements do not reveal consistent, diagnostic markers of distinct emotional states (Barrett, 2006), and consistency in brain and bodily activation is driven by situations/methods rather than emotions themselves.

Limitations and Future Directions

Our own study is not without limitations. Although our stimuli were posed by individuals of African origin, it is not known whether the expressions that they portray are isomorphic with the facial actions that Himba individuals typically make in everyday life, since they were portrayals based on the Western set of emotion categories. While these posed expressions set limits to what we can learn about the nature of emotion representations in Himba individuals, they were well suited to the hypothesis at hand, allowing us to test the extent to which the presumed universal (i.e., Western or US cultural) model describes perceptions of emotion in individuals from a remote culture. It is possible that we observed more “action-based” perceptions of these faces in Himba individuals precisely because these presumed facial expressions are not universal at all, and are less meaningful as emotions in Himba culture. Future research should investigate whether Himba culture contains its own unique set of expressions, and whether Himba individuals use mental state inferences when presented with them.

A second way in which our study fails to fully characterize emotion perception in Himba individuals relates to the fact that we used translations of US emotion category words, and additional characterization of Himba lexical categories is needed. It is possible that the Herero translation for emotion words used in our study are recently borrowed, as appears to be the case for the color term “burou”, borrowed from the German ‘blau’ also meaning blue. This might explain why Himba participants infrequently used the Otji-Herero discrete emotion words to label their piles, as well as the reduced potency of words to shape the perception of facial expressions. To that end, it would also be important to know the actual frequency with which Himba speakers use emotion words in everyday discourse (which could speak to how functional those words are for shaping perception), and the extent to which Himba individuals engage in action identification vs mental state inference more broadly. Our results suggest that Himba speakers might effectively use behavior categories for making predictions about future actions and their consequences, whereas US individuals rely on mental state inferences for those purposes.

Finally, the present study was only designed to test for action-perception as an alternative source of *minimal universality*. Future research should examine under what conditions affect perception (e.g., dimensions of pleasure-displeasure and arousal) also accounts for cross-cultural consistency in emotion perception.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Preparation of this manuscript was supported by the NIH Director's Pioneer Award (DP1OD003312) to L.F.B. We thank K. Jakara for his translation work. We thank J. Vlassakis, A. Matos, J. Fischer, S. Ganley, P. Aidasani, O. Bradeen and E. Muller for their help with data collection, entry or coding. We also thank S. Caparos and J. Davidoff for use of their supplies and support regarding data collection in Namibia.

References

- Aldenderfer, MS.; Blashfield, RK. Cluster analysis. Newbury Park: Sage Publications; 1984.
- Barrett LF. Are Emotions Natural Kinds? Perspectives on Psychological Science. 2006; 1(1):28–58. [PubMed: 26151184]
- Barrett LF. The future of psychology: Connecting mind to brain. Perspectives in Psychological Science. 2009; 4(4):326–339.
- Barrett LF. Emotions are real. Emotion. 2012; 12(3):413–429. [PubMed: 22642358]
- Barrett LF, Mesquita B, Gendron M. Context in emotion perception. Current Directions in Psychological Science. 2011; 20(5):286–290.
- Barrett LF, Satpute A. Large-scale brain networks in affective and social neuroscience: Towards an integrative architecture of the human brain. Current Opinion in Neurobiology. 2013; 23:361–372. [PubMed: 23352202]
- Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? Perspectives on Psychological Science. 2011; 6(1):3–5. [PubMed: 26162106]
- Calabria M, Cotelli M, Adenzato M, Zanetti O, Miniussi C. Empathy and emotion recognition in semantic dementia: a case report. Brain Cogn. 2009; 70(3):247–252. [PubMed: 19339096]
- Chapman HA, Kim DA, Susskind JM, Anderson AK. In bad taste: Evidence for the oral origins of moral disgust. Science. 2009; 323:1222–1226. [PubMed: 19251631]
- Coxon, APM.; Davies, P. The user's guide to multidimensional scaling : with special reference to the MDS(X) library of computer programs. Exeter, NH: Heinemann Educational Books; 1982.
- de Fockert JW, Caparos S, Linnell KJ, Davidoff J. Reduced Distractibility in a Remote Culture. PLoS One. 2011; 6(10):e26337. [PubMed: 22046275]
- Ekman, P. Universals And Cultural Differences In Facial Expressions Of Emotions. In: Cole, J., editor. Nebraska symposium on motivation. Lincoln, Nebraska: University of Nebraska Press; 1971. p. 207-283.
- Ekman P, Cordaro D. What is Meant by Calling Emotions Basic. Emotion Review. 2011; 3(4):364–370.
- Ekman P, Friesen WV. Constants across cultures in the face and emotion. J Pers Soc Psychol. 1971; 17(2):124–129. [PubMed: 5542557]
- Ekman P, Heider E, Friesen WV, Heider K. Facial expression in a preliterate culture. 1972
- Elfenbein HA, Ambady N. On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychol Bull. 2002; 128(2):203–235. [PubMed: 11931516]
- Feldman L. Valence focus and arousal focus: individual differences in the structure of affective experience. Journal of Personality & Social Psychology. 1995; 69(1):153–166.
- Gendron M, Lindquist KA, Barsalou L, Barrett LF. Emotion words shape emotion percepts. Emotion. 2012; 12(2):314–325. [PubMed: 22309717]
- Gendron M, Roberson D, van der Vyver JM, Barrett LF. Perceiving emotion from vocalizations of emotion is not universal. Manuscript under revision. 2013
- Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L, et al. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. J Neurosci Methods. 2002; 115(2):137–143. [PubMed: 11992665]
- Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society, Series C (Applied Statistics). 1979; 28(1):100–108.

- Hontz CR, Hartwig M, Kleinman SM, Meissner CA. Credibility Assessment at Portals, Portals Committee Report. 2009
- Izard, CE. The face of emotion. New York: Appleton-Century-Crofts; 1971.
- Izard CE. Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychol Bull.* 1994; 115(2):288–299. [PubMed: 8165273]
- Jack RE, Garrod OGB, Yu H, Caldara R, Schyns PG. Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America.* 2012; 109(19):7241–7244. [PubMed: 22509011]
- Kozak MN, Marsh AA, Wegner DM. What do I think you're doing? Action identification and mind attribution. *J Pers Soc Psychol.* 2006; 90(4):543–555. [PubMed: 16649854]
- Kreibig SD. Autonomic nervous system activity in emotion: A review. *Biol Psychol.* 2010; 84(3):394–421. [PubMed: 20371374]
- Kruskal, JB.; Wish, M. Multidimensional scaling. Beverly Hills, Calif: Sage Publications; 1978.
- Lindquist KA, Barrett LF, Bliss-Moreau E, Russell JA. Language and the perception of emotion. *Emotion.* 2006; 6(1):125–138. [PubMed: 16637756]
- Lindquist KA, Barrett LF. A functional architecture of the human brain: Insights from Emotion. *Trends in Cognitive Sciences.* 2012; 16(11):533–540. [PubMed: 23036719]
- Lindquist KA, Gendron M. What's in a word? Language constructs emotion perception. *Emotion Review.* 2013; 5(1):66–71.
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. The brain basis of emotion: a meta-analytic review. *Behav Brain Sci.* 2012; 35(3):121–143. [PubMed: 22617651]
- Matsumoto D, Ekman P. American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion.* 1989; 13(2):143–157.
- Matsumoto, D.; Keltner, D.; Shiota, M.; Frank, M.; O'Sullivan, M. Facial expressions of emotions. In: Lewis, M.; Haviland-Jones, JM.; Barrett, LF., editors. *Handbook of emotion.* New York: Macmillan; 2008. p. 211-234.
- Norenzayan A, Heine SJ. Psychological universals: what are they and how can we know? *Psychol Bull.* 2005; 131(5):763–784. [PubMed: 16187859]
- Roberson D, Damjanovic L, Pilling M. Categorical perception of facial expressions: evidence for a "category adjustment" model. *Mem Cognit.* 2007; 35(7):1814–1829.
- Roberson D, Davidoff J, Braisby N. Similarity and categorisation: neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition.* 1999; 71(1):1–42. [PubMed: 10394708]
- Sauter DA, Eisner F, Ekman P, Scott SK. Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences of the United States of America.* 2010; 107(6):2408–2412. [PubMed: 20133790]
- Sokal RR, Michener CD. A statistical method for evaluating systematic relationships. *The University of Kansas Scientific Bulletin.* 1958; 38:1409–1438.
- Sorensen, ER. Culture and the expression of emotion. In: Williams, TR., editor. *Psychological Anthropology.* Paris: Mouton Publishers; 1975. p. 361-372.
- Spunt RP, Lieberman MD. An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *Neuroimage.* 2012; 59(3):3050–3059. [PubMed: 22019857]
- Susskind JM, Lee DH, Cusi A, Feiman R, Grabski W, Anderson AK. Expressing fear enhances sensory acquisition. *Nature Neuroscience.* 2008; 11:843–850. [PubMed: 18552843]
- Tian X, Huber DE. Testing an associative account of semantic satiation. *Cogn Psychol.* 2010; 60(4): 267–290. [PubMed: 20156620]
- Tomkins, SS. *Affect, Imagery, Consciousness.* Vol. Vol. 1. New York: Springer; 1962.
- Tomkins, SS. *Affect, Imagery, Consciousness.* Vol. Vol. 2. New York: Springer; 1963.
- Vallacher RR, Wegner DM. What Do People Think They're Doing - Action Identification and Human-Behavior. *Psychological Review.* 1987; 94(1):3–15.
- Weinberger S. Airport Security: Intent to deceive? *Nature.* 2010; 465:412–415. [PubMed: 20505706]
- Widen SC, Christy AM, Hewett K, Russell JA. Do proposed facial expressions of contempt, shame, embarrassment, and compassion communicate the predicted emotion? *Cogn Emot.* 2011; 25(5): 898–906. [PubMed: 21824027]

- Widen SC, Russell JA. Differentiation in preschooler's categories of emotion. *Emotion*. 2010; 10(5): 651–661. [PubMed: 21038948]
- Wilson-Mendenhall CD, Barrett LF, Simmons WK, Barsalou LW. Grounding emotion in situated conceptualization. *Neuropsychologia*. 2011; 49:1105–1127. [PubMed: 21192959]
- Young FW, Lewycky R. *ALSCAL User's Guide*. 1979
- Zaki J, Hennigan K, Weber J, Ochsner KN. Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *J Neurosci*. 2010; 30(25):8481–8488. [PubMed: 20573895]

Figure 1.

Pile labels used by Himba and US participants. Mean number of words produced by each group ($\pm$ standard error) is plotted on the y-axis broken down by word type. We observed cross-cultural differences in label use when participants were asked to freely sort facial expressions, ($F_{1,55}=24.952$, $P<.001$, $\eta_p^2=.312$), and this effect was qualified by the type of label produced, ($F_{2,110}=56.719$, $P<.001$, $\eta_p^2=.508$).

Figure 2.

Six-cluster solutions derived from a hierarchical cluster analysis. The cluster analyses are plotted in A–D with cluster on the x-axis. The y-axis represents the number of items grouped into a given cluster, with contents stacked by the emotion portrayed in each posed facial expression. Stacked bars containing several different colors indicate that faces portraying different discrete emotions were clustered together. Bars with a single color (or predominance of a given color) indicate relatively clean clustering of faces depicting one emotion category. The US free sort (A) cluster solution contains discrete emotion (i.e., the presumed “universal”) clusters with the exception of cluster five, which appears to contain portrayals of both disgust and sadness. The Himba free sort (B) has three clear clusters (one through three) that map on to discrete emotion (i.e., the presumed “universal”) pattern. Both US (C) and Himba (D) conceptually anchored sorting appears to yield relatively distinct cluster solutions compared to free sorting.

Figure 3.

Multidimensional scaling (MDS) solutions for free sorting of facial expressions. Free sort data are plotted in two-dimensional space for Himba (**A**) and US (**B**) participants. Anchored sort data are also plotting in two-dimensional space for Himba (**C**) and US (**D**) participants. Items are plotted by emotion type. Clearer evidence of the “universal solution” (closer clustering of facial expressions within the same emotion category) is more evident in the US solutions (**B, D**) than in the Himba solutions (**A, C**). We quantified the clustering of items

within a category across all dimensions and plotted these mean distances ($\pm$ standard error) for each cultural group in each task (E).

Table 1
Multidimensional Scaling (MDS) Dimensions Identified using Hierarchical Multiple Regressions

Sort Solution	Dimension 1		Dimension 2		Dimension 3		Dimension 4	
	Behavior	Emotion	Behavior	Emotion	Behavior	Emotion	Behavior	Emotion
Himba Free	.826 ***	.064	.732 ***	.060	.597 ***	.122		
US Free	.938 ***	.021	.885 ***	.078 ***	.913 ***	.028	.776 ***	.105 *
Himba Anchored	.865 ***	.055	.803 ***	.100 *	.704 ***	.091		
US Anchored	.854 ***	.098 ***	.926 ***	.032 *	.786 ***	.137 ***	.851 ***	.080 **

Note. The first block of regressors was comprised of eight sets of independent ratings where the 36 face stimuli were rated for the extent to which they depicted specific behaviors (e.g., smelling, crying; see supplementary Tables S5, S6 for details); R^2 values are presented in the Behavior column under each dimension. The second block of regressors was comprised of six sets of ratings where each of the 36 faces was rated for the extent to which they represented emotions (e.g., disgust, sadness); the change in R^2 due to these ratings is presented under each dimension labeled Emotion. The change in R reflects the extent to which the dimension reflects mental state inferences about emotion over and above the variance accounted for by mere action identification. Himba sorting data was captured by a lower dimensionality solution compared to US sorting data (3 dimensions instead of 4). As a result, a fourth dimension was only identified for US participants' data. Significant effects are bolded;

* p .05,
** p .01,
*** p .001.