Constructing Embodied Emotion with Language: Moebius Syndrome and Face-Based Emotion Recognition Revisited

Forthcoming at *Australasian Journal of Philosophy*
(This is a preprint. Please cite the published version.)

Hunter Gentry
Department of Philosophy
University of Wisconsin-Madison
hgentry@wisc.edu

Constructing Embodied Emotion with Language: Moebius Syndrome and Face-Based Emotion Recognition Revisited

Abstract: Some embodied theories of concepts state that concepts are represented in a sensorimotor manner, typically via simulation in sensorimotor cortices. Fred Adams (2010) has advanced an empirical argument against embodied concepts reasoning as follows. If concepts are embodied, then patients with certain sensorimotor impairments should perform worse on categorization tasks involving those concepts. Adams cites a study with Moebius Syndrome patients that shows typical categorization performance in face-based emotion recognition. Adams concludes that their typical performance shows that embodiment is false. Moebius patients must draw on amodal (non-embodied) emotion concepts. In this paper, I review face-based emotion recognition studies with Moebius patients yielding conflicting results, and diagnose these conflicts as a difference in experimental design. When emotion labels are provided, patients have typical performance, but when labels are not provided patients are severely deficient. I then show how an embodied, psychological constructionist view of emotions predicts and explains these performance differences. The upshot is that embodied theories of concepts are vindicated.

Keywords: Embodiment, Emotions, Concepts, Categorization, Language, Perception

Introduction

One hypothesis of embodied cognition states that cognition centrally involves sensorimotor processes (Newen, de Bruin, and Gallagher, 2018).[1] This hypothesis has received widespread attention in recent philosophical and psychological discourse. Many psychological studies with surprising results have been taken to support embodied cognition (for a comprehensive review and discussion see Farina 2021; Gallagher 2011; Shapiro 2019). Of special interest to this strand

[1] Following Shapiro (2019), there is probably not one thesis that unifies embodied cognition research. Indeed, there are at least three logically independent theses we can distinguish:
Conceptualization: the properties of an organism's body limit or constrain the concepts that that organism can acquire.
Replacement: Cognitive processes are not discrete, but continuous which makes the standard computational framework ill-suited to explain cognition. Additionally, an organism's body in interaction with the environment replaces the need for representational processes.
Constitution: The body or the world plays a constitutive rather than a merely causal role in cognitive processing (pp.4-5).
Replacement is most closely associated with anti-representational and dynamical systems approaches to cognition (e.g., Chemero 2011), whereas *Constitution* is most closely associated with extended cognition (e.g., Clark and Chalmers 1998). In this paper, I am focusing on *conceptualization*, and setting aside the latter two theses.

of embodied research is the suggestion that concepts are embodied.[2] Concepts are embodied, on this view, in the sense that deploying them involves sensorimotor simulations which activate sensorimotor cortex.[3]

Consider the following case from Bergen and Feldman (2008):

> Can you say how many windows there are in your current living quarters? Almost everyone simulates a walk-through to count them. Or consider a novel question—could you make a jack-o-lantern out of a grapefruit? (pg. 315).

According to Bergen and Feldman, in deploying your concept of a grapefruit, you engage in a task-dependent kind of sensorimotor simulation.[4] In this case, "reflecting on the carvability of a grapefruit involves creating internal motor and sensory experiences of carving a jack-o-lantern out of a grapefruit" (pg. 315). For these theorists (and many others), "Any time we use concepts, whether in performing categorization tasks, processing language about concepts, or reflecting on their features, we use mental simulation—the internal creation or recreation of perceptual, motor, and affective experiences" (pg. 315). The internal (re)creation of these experiences is thought to be implemented by activation of sensorimotor cortex.

---

[2] For a discussion of learning embodied concepts see Bergan and Feldman (2008). See also Núñez (2008) for a discussion of embodied mathematical concepts.

[3] Embodied concepts are sometimes referred to as "modal concepts" to contrast with "amodal concepts". They are called "modal" because the vehicle of representation is grounded in sensory modalities. Amodal concepts are abstract, symbolic representations, e.g., Language of Thought (Fodor 1975). However, see Michel (2021) for skepticism regarding the distinction between modal and amodal concepts.

[4] There is a wealth of literature on embodied simulation (Barsalou 1999; Bergen 2012; Shapiro 2019). For work on embodied simulation and metaphor see Gibbs, Jr. (2006); Lakoff and Johnson (2008).

Or consider the action-sentence compatibility effect (see Bergen 2012 for extended discussion). When asked to judge whether a sentence made sense, and to indicate their judgment via an action (here: pushing a button), subjects were faster at responding when the action was consistent with the target sentence. The hypothesized explanation is that interpreting the sentence requires sensorimotor activation, and concurrent inconsistent motor behavior interferes with this process. If interpreting sentences requires the deployment of concepts and interpreting sentences requires sensorimotor activation, then plausibly, what it is to deploy a concept is to activate sensorimotor cortex, i.e., embodied simulation.

Fred Adams (2010) has advanced an empirical argument against embodied concepts. Adams reasons that if concepts are embodied, and deployment of an embodied concept involves simulating sensorimotor processes, then patients with certain sensorimotor impairments should perform worse on categorization tasks involving those concepts. Moebius syndrome patients suffer from congenital bilateral facial palsy due to cranial nerve underdevelopment. As such, these patients cannot make facial expressions. The embodied concepts thesis predicts that these patients would have impairments in face-based emotional recognition tasks because they would not be able to simulate facial expressions. However, in Calder et al.'s (2000) study, Moebius patients did not significantly differ in their performance on such tasks compared to healthy controls. Adams concludes that their comparable performance shows that embodiment is false. Moebius patients must draw on amodal (non-embodied) emotion concepts.

In this paper, I will attempt to show a few things. For one, I want to complicate the picture that Adams (2010) has about Moebius syndrome performance in face-based emotion recognition tasks. The experimental literature has some studies showing comparable performance to healthy controls and others showing deficient performance. I think these mixed results are the reflection of a key experimental design difference--the use of labels. When experimenters provide labels, Moebius patients perform comparably to non-Moebius subjects, but when labels are not used, the patients are unable to complete the task. I will explain these results via research on the language-cognition-perception interface that shows a "label superiority effect" among subjects in categorization tasks. While the importance of labels seems to lend support to Adams' view that Moebius patients are utilizing amodal concepts, I will argue that it is also compatible with a more embodied view of emotions-- psychological constructionism. I will describe the basic philosophical commitments associated with this view, and then show how the view predicts and explains the performance differences in Moebius patients. The result is a view of embodied emotional concepts held together by the "glue" that is language.

Here's a roadmap for the rest of the paper. In section 1, I review a few different Moebius syndrome studies of face-based emotion recognition, paying special attention to the performance differences when labels are used or not used. I will then provide a brief analysis of the findings. In section 2, I will argue that the use of labels is the key difference in experimental design by reference to the "label superiority effect". In section 3, I will get specific about what role labels are playing in the Moebius syndrome studies such that Moebius patients can perform comparably to controls. Here, I will also introduce my favored theory of emotions: psychological constructionism, and show how this view predicts and explains the performance differences.

1. Face-Based Emotion Recognition in Moebius Patients

Adams' (2010) argument against embodied concepts can be formalized as follows:

1. If Moebius patients succeed on face-based emotion recognition tasks, then concepts are not embodied.
2. Moebius patients do succeed on such tasks.
3. So, concepts are not embodied

Premise 1, Adams thinks, is what embodied theorists are committed to saying about Moebius patients. In particular, if the embodied theorist thinks that simulation of facial expressions is necessary for successful categorization of facial expressions, then because Moebius patients cannot simulate facial expressions, they will not succeed. Premise 2, Adams thinks, is supported by Calder et al. 's (2000) study showing successful categorization in Moebius patients. And I suppose he thinks that the results must generalize. The conclusion, then, follows deductively.

I take issue with both premises of this argument. I will start with premise 2 by reviewing a few faced-based emotion recognition studies with Moebius patients that yield conflicting results. I will ultimately argue that the performance differences are due to differences in experimental design-- the use of emotion labels. This difference leads to two different kinds of studies-- both probative, but eliciting different capacities. Importantly, these differences in performance can be predicted and explained by my favored view of emotions: psychological constructionism; therefore, premise 1 is false. But before I get to all of that, let us take a closer look at the complicated picture of Moebius patients' performance on face-based emotion recognition tasks.

1.1 A Complicated Picture of Moebius Patient Performance

Originally described by von Graefe and Saemisch (1880) and named by Moebius in 1888, Moebius syndrome is congenital bilateral facial palsy (and/or bilateral abducens palsy) resulting from cranial nerve underdevelopment (more specifically cranial nerves VII and VI).[5] Incidence of the syndrome is complicated due to comorbidities and inconsistencies of associated symptoms and features (Bell et al. 2019); however it is estimated to occur in 1 in 50,000 to 1 in 500,000 live births (Rasmussen et al. 2015). Some researchers report that there is equal incidence among genders (De Stefani et al. 2019; Nicolini et al. 2019), but Bell et al. (2019) report slightly higher incidence in males. Additionally, most patients are of "normal" (i.e., typical) intelligence (Nicolini et al. 2019) and cognitive development (De Stefani et al. 2019) .

Due to facial palsy, patients cannot make facial expressions. As such, patients often report difficulties in building social relations. A deficit in emotional processing is thought to explain this difficulty. In particular, patients are thought to be deficient in attributing emotional categories to facial expressions.[6]

To test this claim, Giannini et al. (1984) had a single Moebius patient watch videotapes of people at slot machines playing for one cent, ten cent, or twenty-five cent jackpots. The videotapes recorded the facial expressions at selection of the jackpot and the payout (or lack thereof). After watching, the patient was asked to determine how much money was at risk. To do this, the patient was instructed to press the button corresponding to their answer (either 1¢, 10¢, or 25¢).

[5] Perhaps a better name is "Moebius sequence", however, because to date, there are no strict diagnostic criteria due to inconsistencies of associated symptoms (for a review see Bell et al. 2019; De Stefani et al. 2019).
[6] Caution must be urged in interpreting the studies reviewed here due to extremely small sample sizes. It is difficult to draw any definitive conclusions. This cuts both ways-- against my own analysis and argument, as well as Adams (2010). However, I am attempting to provide a more holistic analysis of the available data on Moebius syndrome and face-based emotion recognition than Adams.

Because the video only showed facial expressions, the judgment was made based purely on facial cues.

Despite having a typical IQ and master's level education, "the patient could not respond at all to the task" (Giannini et al. 1984; pg.174). The patient described the experience as like trying to understand a foreign language. Approximately 300 controls had performed the task without any issues. The authors hypothesize that because the controls did not suffer from any neurological conditions, the patient's inability to complete the task is due to her Moebius syndrome. That is, the fact that the patient cannot make facial expressions is supposed to explain her inability to recognize them.

More recently, Nicolini et al. (2019) were primarily interested in testing autonomic responses in Moebius patients, but to get a baseline for emotion recognition, they showed video clips of cartoon characters acting out emotions to be identified by the subjects.[7] Initially, they asked the patients what emotional state they thought the character in the video was in. Importantly, this was a free-response question with no cueing or labels provided. Similar to the patient in Giannini et al.'s study, the patients could not respond to the questioning.

To overcome this difficulty and achieve their baseline measure, Nicolini et al. administered the Test of Emotion Comprehension (TEC) (Pons et al. 2002). The TEC typically involves the experimenter reading a story to the subjects or showing a video and then presenting the subjects

[7] Nicolini et al. (2019) say that the "Stimuli were comprised of short clips taken from the Internet in which the main character of the scene was in a happy, sad, or scary situation" (pg. 3). Apparently, the video clips included more than just facial expressions which is significant because the subjects could not respond to the questioning. This might be a reflection of children's underdeveloped emotional conceptual system in general, which could be severely underdeveloped in Moebius children. See Widen (2013) for a discussion and review of children's emotional conceptual development.

with four facial expressions, e.g., sad, happy, angry, neutral. Nicolini et al. presented the four facial expressions and asked the patients to select which one matched the cartoon character's emotional state in the video they watched. Again, crucially, the presented facial expressions did *not* have emotion labels. The patients' performance on TEC was significantly impaired compared to controls (mean control scores: ~5 versus mean patient scores: ~2.5).

However, in 2000, Calder et al. conducted a face-based emotion recognition study with Moebius patients that showed comparable performance to healthy controls. Calder et al. (2000) tested three Moebius patients on a battery of tasks. Most important for our purposes, however, is the Ekman facial recognition task. The researchers used the Ekman and Friesen (1976) series of 60 photos. There were 10 models who each assumed 6 different facial affectations (happiness, sadness, anger, fear, disgust, and surprise). The subjects were asked to identify which emotion was being modeled in each picture from a list of 6 emotion labels. Patient performance did not significantly differ from 40 healthy controls (mean patient score: 47.67/60.00 versus mean control score: 50.35/60.00). However, Calder et al. do note that patients' performance was slightly deficient.

Bogart and Matsumoto (2010) also found that Moebius patients did not significantly differ from controls in face-based emotion recognition. At the time of my writing this, these authors have the title of largest sample size with Moebius patients-- clocking in at n=37. Bogart and Matsumoto used a variant on the traditional Ekman set of photos, called the "Multi-Ethnic Facial Expression Set" (Matsumoto and Ekman 2006).[8] Procedurally, participants are presented with a photo of a

[8] The original Ekman task used only white models posing in 6 basic emotions. This variant used black, brown, and white models posing in 7 basic emotions, adding "contempt".

face expressing one of seven emotions. They have unlimited time to view the photo and then choose the correct emotion label. The control group had a performance range of 52-93 (M=75.44) while the Moebius group ranged from 60-90 (M=73.67).

1.2 Taking Stock

So here's the complicated picture of Moebius patient performance on face-based emotion recognition tasks: Calder et al. (2000) and Bogart and Matsumoto (2010) report comparable performance to healthy controls (as do Vannuscorpus et al. 2020), but Giannini et al. (1984) and Nicolini et al. (2019) report deficient performance.[9] These mixed results, whatever the explanation, cast doubt on Adams' 2nd premise. It is not the case that Moebius patients always succeed on emotion recognition tasks. They sometimes do; and in the studies where they do succeed, the experimental design is different from the studies where they do not. In particular, Calder et al. (2000), Bogart and Matsumoto (2010), and Vannuscorpus et al. (2020) provided a list of emotion labels to choose from, whereas Giannini et al. (1984) and Nicolini et al. (2019) did not use emotion labels-- subjects free-named the facial expressions.

In addition, Nicolini et al. (2019) tested young children with Moebius syndrome as opposed to adults. Here's what they say about this decision:

[9] Giannini et al. did not make use of *emotional* labels. The patient had to recognize the emotional expressions in the videos and then infer how much money was at stake. It is doubtful that labels for the amount of money would interact with emotional processing to produce an advantage, unless the patient knew the meaning of the emotional expressions already.

> Our study is the first to use a relatively large sample of very young patients to investigate the effects of facial muscle paralysis on both autonomic responses and emotion recognition. The investigation of these issues early in development *is critical for the detection of emotional processing mechanisms at a stage where more complex cognitive strategies might not yet compensate for their deficits* (2019, pg.8; emphasis added).[10]

The claim here seems to be that typical (as opposed to deficient) adult Moebius performance on emotional recognition tasks might be a reflection of compensatory mechanisms. It is possible that adult patients have developed strategies for emotional recognition to compensate for their inability to make and simulate facial expressions (Michael et al. 2015). So success on these tasks could have other explanations.

Correlatively, *deficits* on emotion recognition tasks could have other explanations as well. For example, the difficulties we see in social interaction between Moebius subjects and non-Moebius subjects could perhaps be due to the non-Moebius subjects having difficulty in the interaction. As Michael et al. (2015) put it, “Since people are accustomed to receiving information about others’ mental states from their facial expressions, the absence of this expected information may interrupt an interaction partners’ facial mimicry and cause him or her to feel uncomfortable or confused about what the person with [Moebius syndrome] is thinking or feeling” (pg.2). Furthermore, as an anonymous reviewer pointed out, Moebius subjects may have received less social input as children because others deemed them to be uninteresting or unresponsive.

[10] De Stefani et al. (2019) found similar findings with children with Moebius syndrome. In a forced choice emotion categorization task (with stylized faces rather than labels), Moebius children were less accurate than controls. The authors hypothesize that the children might rely on more rule based strategies to identify the target emotion. For example, if the stimuli has feature F, then it is emotion E. The problem with such a strategy is that if F is shared with other emotional categories, then F is not diagnostic of a *single* category. As we will see in later sections, there is reason to suspect that features of emotional facial expressions are shared with many emotional categories.

This is all to say that success and failure on these face-based emotion recognition tasks are hard to interpret. However, this is a significant point because it puts pressure on Adams' premise 1. In particular, *adult* patients succeeded in the studies reported above, and this could be due to compensatory mechanisms that are themselves embodied. Without ruling this possibility out, Adams cannot conclude that embodied theories as a whole are mistaken. That is, Moebius patients can succeed on emotion recognition tasks and embodiment can still be true. So, given that when labels are used patients succeed, and when labels are not used they fail, could it be that the labels are functioning as a compensatory mechanism? Further, could it be that labels interact with embodied processes?

2. The Label Superiority Effect and Language-Cognition Interface

In this section, I will begin to make the case that the use of labels in the Moebius syndrome studies is a key difference. Specifically, it could be the case that the use of labels explains how Moebius patients are able to successfully categorize emotional facial expressions. Moreover, as I will argue later, this explanation, although involving labels, is still embodied. I will start by discussing the importance of labels in categorization-- including emotional categorization-- in healthy populations. I will then review some evidence that emotional categories are perceptually fuzzy and complex, and that language helps to simplify distinctions between categories.

2.1 How and When Labels are Cognitively Advantageous

Carruthers (2002) distinguishes between strong and weak theses concerning the interface between cognition and language. Strong theses have it that language is somehow necessary for *all* human cognition. Weak theses have it that language is not necessary for cognizing, but that it "scaffolds" cognition-- language is necessary to acquire some (but not all) concepts and/or language reduces computational burdens. I will not be taking a stance on whether the strong thesis is true, but I will be assuming that the weak thesis is true.[11] Everyone can agree that a human with language and another human without language will differ cognitively. The question is: how different will they be?

A perceptually fuzzy category is a category whose perceptual presentation overlaps with other categories. Importantly, this does not mean that the underlying reality of the category is indeterminate, it just means that evidence for the correct application of the corresponding concept is fuzzy or complex. Here's an example: COVID-19 and the common cold are distinct diseases, but they have overlapping symptoms. This makes it hard to perceptually distinguish the two, but this does not mean that there is no boundary between the two disease categories.

One function that labels play is to transform perceptually fuzzy or complex categories into perceptually discrete, unified categories. This is a function that language plays when learning the target concept as well as deploying the target concept after it has been learned. When a category is perceptually fuzzy, graded, or probabilistic, it might not be possible to learn the corresponding concept without language. Extreme examples might include 'democracy' and 'charity'-- these highly abstract concepts do not have simple perceptual extensions. At the very least, it will be

[11] Carruthers dismisses the strong thesis as it is formulated here.

easier to learn these concepts with labels, than without.[12] Secondly, given that one has learned a perceptually fuzzy concept, having a label associated with it makes categorization easier. The label becomes associated with the most diagnostic features of that category, making salient boundaries between categories that are otherwise perceptually obscured, and serving as a direct route to category knowledge. In learning a word, one learns to chunk together certain, perhaps diverse, mental representations which facilitates concept learning and deployment (Pietroski 2005).

There are many views on how labels have this transformative effect (I am going to remain agnostic on this).[13] One view, call it "the attentional view", says that labels guide feature attention to the most diagnostic features of the category.[14] As an example, consider the shape bias. Starting around 18 months old, children will attend to shape as an indicator of category membership.[15] The idea here is that objects in the environment do not frequently change shape without also changing category. For example, I can't turn into a wolf without changing my shape. Children as early as 18-months old learn that despite perceptual dissimilarity along many

[12] Ramscar et al. (2010) argue that learning to predict a category label from exemplar features (FL learning) (rather than predicting features from a label (LF learning)) is more effective when the category is perceptually fuzzy. This is because, "LF learning tends to produce representations in which a number of competing outcomes are all highly probable" (pg.922). However, in FL learning, the subject discriminates between competing outcomes. The authors ultimately argue that language learning and understanding is a predictive process-- it's about tightly coupling the exemplar features that best predict the category label. In language understanding then, words are cues to predict meaning (see also Elman 2009).

[13] Another view says that language works to reduce the dimensionality of representations effectively making them more categorical. See Blouw et al. (2016; especially pg. 1138) , DiCarlo and Cox (2007), and Stewart et al. (2011) for explanations at the neural level. For views like this based on deep convolutional neural networks and transformer models see Buckner (2018), Gregor et al. (2016), and Valeriani et al. (2023). For a view like this based on psychological findings see Perry and Lupyan (2014). Finally, see Edelman and Intrator (1997a,b) for early descriptions of this view based on psychophysical findings from color and shape perception.

[14] Green (2020) explains the attentional view as a modulation effect: "A modulation effect makes a difference to the perceptual representation of feature values along a fixed dimension or set of dimensions (e.g., color, size, distance, contrast) that a perceptual process already has at its disposal" (pg.327).

[15] It is worth highlighting that it seems the shape-bias plays a role in Lupyan et al.'s (2007) alien study too (reviewed below). The fuzzy perceptual distinction between the two aliens consists in differences in shape.

dimensions between two stimuli, if shape stays constant, then they are likely of the same category.

In Linda Smith's "dax" experiments (Smith and Heise 1992; Jones and Smith 1993), young children are shown a "dax", and fail to extend the name to objects that do not share the same shape as the training exemplars. This suggests that shape matters to whether an item belongs to the dax category or not. However, they will extend the name if, for example, the size or texture changes, suggesting that these features are not important for membership.[16] Here, shape is not only useful in learning novel categories, it is also useful in deploying concepts after they have been learned too. If shape, in general, is a diagnostic feature of category membership across multiple categories, then attending to specific shapes will be useful in learning new concepts, as well as, deploying already learned concepts. Attending to shape, at the exclusion of other features of a particular stimulus, simplifies the representation of that stimulus. And when the attended-to feature is particularly diagnostic of a category, inferring category membership becomes easier (in learning or in deployment of learned categories).

In discussing these advantages, Clark (1998) says that "[T]he presentation of the same label accompanying a series of slightly different perceptual inputs (e.g., different views of dogs)...flags the presence of some further underlying structure and thus invites the network to seek the perceptual commonality" (pg.7). McClamrock (1995) makes a similar point when he says that labels thrust stable properties onto a complex environment effectively reducing our computational burden. In short, labels serve as a kind of "glue" to bind together perceptually

[16] They will not extend the name if the texture changes when the training sample has eyes. Smith and colleagues conclude that texture matters to extension of names if the exemplars are animate objects (the eyes are supposed to denote animate object).

complex stimuli, making them perceptually simple.[17] This makes concepts easier to learn while also facilitating categorization once the concept has been learned. This is one thing we mean when we say that "language scaffolds cognition".[18]

2.2 Experimental Evidence of the Label Superiority Effect

To help illustrate both of these advantages, let us consider the following study by Lupyan et al. (2007) which showed that subjects were faster and more accurate at categorizing novel stimuli when provided with a nonsense label than subjects not provided a label.[19] Lupyan et al. trained subjects to associate two kinds of "aliens" with either of two behavioral responses: "approach" or "flee". The two kinds of aliens were subtly different in shape (see figure 1). There were two training conditions: a label condition, in which the presented alien was labeled (either "leebish" or "grecious") and a no-label condition. Subjects were given feedback after each answer. After training, subjects were tested on how well they learned the associations. The results showed that subjects in the label condition, despite having the same amount of experience with the stimuli as subjects in the no-label condition, were faster and more accurate in both the training phase and at test. Lupyan et al. conclude that "learning nonsense verbal labels facilitated categorization more than did learning nonverbal associations" (2007, pg.1081). Lupyan et al. provide a possible

---

[17] Some researchers have thought that labels "chunk" complex perceptual representations into discrete, unified representations (Huang and Awh 2018). This view might not be a competitor to the attentional view because attending to one perceptual feature could be an instance of chunking the perceptual representation. Alternatively, it could be that attending to category diagnostic features across perceptually dissimilar category exemplars allows the subject to chunk representations of those exemplars under a common label. See Pietroski (2005).

[18] It is worth noting that Clark is heavily influenced by Dennett (1984; 2015) and Deacon (1997). Early versions of this kind of view of language's influence on cognition can be found in these works. Thanks to an anonymous reviewer for this point.

[19] They were more accurate in the sense that there was a matter of fact about similarity that made a particular stimulus a member of one group and not the other. This is because Lupyan et al. created the stimuli to divide into two categories.

explanation for the performance advantage: labels become associated with the most diagnostic information for that category. In their words:

> [R]ather than being fixed features, category names modulate item representations on-line through top-down feedback. According to this account, as a label is paired with individual exemplars, it becomes associated with features most reliably associated with the category. When activated, it then dynamically creates a more robust category attractor (pg.1082).

The idea here, echoing Clark (1998) and McClamrock (1995), is that labels facilitate learning and categorization by becoming strongly tied to the properties that distinguish that category from others.[20] Moreover, though the distinction between two categories might turn out to be perceptually complex, labels allow one to represent that distinction in a simpler way: "...labels may have allowed subjects to more easily represent the somewhat fuzzy perceptual distinction between categories (‘‘more rounded and smooth’’ vs. ‘‘less rounded, with ridges’’) in terms of a simpler verbal distinction (‘‘leebish’’ vs. ‘‘grecious’’)" (Lupyan et al. 2007, pg.1082).[21]

---

[20] See Casasanto and Lupyan (2011).

[21] See also Lupyan and Bergen (2016) and Lupyan and Thompson-Schill (2012).

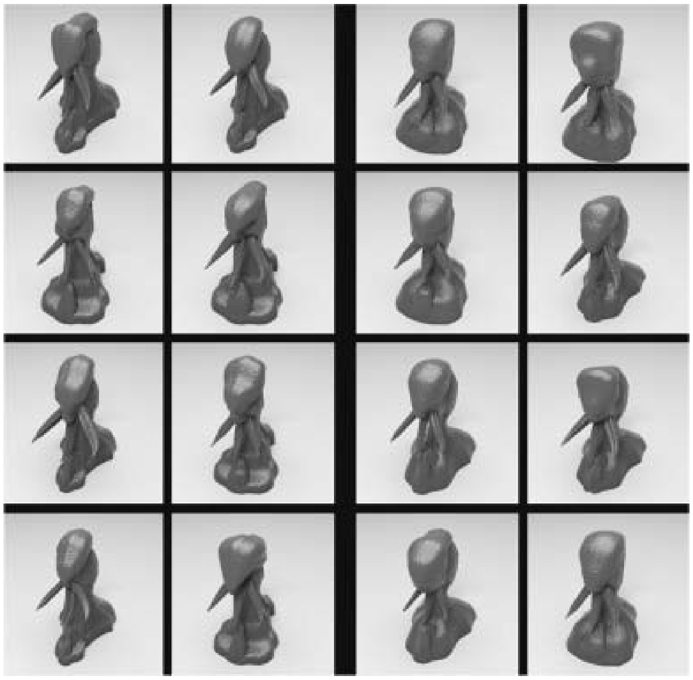

FIGURE 1. The "aliens" on the left have flatter bases and a small ridge on the head. The "aliens" on the right have rounder bases and smoother heads. From Lupyan et al. (2007), pg. 1078.

2.3 Are Emotional Categories Discrete?

Dating back to Darwin (1872), emotions were thought to have a discrete, unified reality in nature. For example, Darwin thought that certain facial expressions and movements corresponded to specific emotional categories because they were adaptive (see Allport 1924). Inspired by Darwin, basic emotion theories (Ekman and Cordaro, 2011; Panksepp, 2011; Shariff and Tracy, 2011; for discussion see Deonna and Teroni, 2012, ch.2) state that emotions and emotional perception are innate, pre-linguistic, and universal.[22] Indeed, the inventor of the

[22] There is ambiguity in the term "basic emotion". Scarantino and Griffths (2011) distinguish between three notions of "basic": conceptual, biological, and psychological. Something is conceptually basic just in case it occupies the basic level of a conceptual taxonomy. Something is biologically basic just in case it has an evolutionary origin and distinctive biological markers. Finally, something is psychologically basic just in case it does not contain another psychological component as a part. These notions of basic-ness are independent of one another. An emotion can be conceptually basic, but this does not entail that it is biologically or psychologically basic, and vice-versa. The basic emotion theorists, I think, have in mind biological basic-ness. It can be said that the Stoics and later early modern

Ekman (1992) facial recognition task was a basic emotion theorist who thought that there were 6 basic emotions that corresponded one-to-one to certain facial expressions. It's not a coincidence that the Ekman facial recognition task has 6 different emotions represented by 6 distinct facial expressions (contempt has been added as a 7th basic emotion; see Matsumoto and Ekman 2004). If basic emotion theories are right, then emotions are *not* perceptually fuzzy categories. Rather, they are discrete, unified natural kinds. Further, if emotions correspond one-to-one to facial expressions, then, by hypothesis, seeing an emotional expression on a face should be sufficient information to know what emotional state that person is in.

Basic theories have faced challenges going back at least 50 years (Lacey 1967; for a review and discussion see Barrett 2006; Colombetti 2014). In what follows, I will review 2 main lines of experimental evidence that allegedly cast doubt on the basic emotion theorists' claim that emotions have a discrete, unified reality.[23] This evidence supports the claim that emotions are perceptually fuzzy categories. I will ultimately argue that because emotions are perceptually fuzzy categories, emotion labels are highly influential for everyone (healthy subjects and certain patient populations) in emotional recognition tasks. This should be sufficient to show that the use of emotion labels is a key difference in experimental design of the Moebius syndrome studies I discussed above.

---

philosophers (e.g., Descartes, Spinoza, Hume, and Locke) thought that there were psychologically basic emotions (see Colombetti 2014, chapter 2).

[23] Above, I distinguished between a category being perceptually fuzzy and a category itself being fuzzy. I do not think the evidence canvassed here shows the latter unless one thinks that facial expressions, vocal signals, and various physiological responses are *constitutive* of emotional categories. If one thought that facial expressions, vocal signals, etc. were merely expressions of emotional categories, then the evidence canvassed here only shows that emotions are perceptually fuzzy. I will not be weighing in on the debate about whether emotional behaviors are constitutive or expressions of emotional categories. All I need to show for my argument is that emotions are perceptually fuzzy. However, if it turned out that emotional categories are *themselves* fuzzy, my argument would still go through.

To pump the reader's intuitions, I invite you to try the Ekman facial recognition task for yourself without labels (see figure 2). Take a moment and try to figure out what emotion is being expressed in each of the six photos.

FIGURE 2. Example photos from the Ekman Facial recognition task without labels.

From left to right starting at the top, we have anger, fear, disgust, shock, happiness, and sadness. How did you do? If you were barely above chance, you are not alone, and for good reasons. Barrett (2006) provides a review of emotion research that allegedly challenges the basic emotion theorists' claims that (i) emotions are discrete, unified categories (see footnote 23) and that (ii) emotion recognition proceeds automatically and passively. First, if basic theories are right, then there should be strong correlations between emotional categories and various measurable responses (e.g., subjective experience, facial expressions, and vocal signals). For example, when anger erupts in a person, there should be specific and distinct bodily responses: an increase in blood pressure, a scowl on the face, and perhaps the urge to yell or hit something. Unfortunately, these strong correlations, as predicted by the basic theories, have failed to materialize. Barrett concludes, “Taken together, enough evidence has accumulated for some theorists to conclude that lack of response coherence within each category of emotion is empirically the rule rather than the exception” (2006, pp.33-34).[24] Essentially, these data suggest that there is variability between subjects in their bodily responses and subjective experiences of emotions, instead of consistent and distinct responses and experiences.

Another line of evidence allegedly against basic emotion theorists are production studies of emotion. In this kind of study, researchers measure bodily responses during emotionally triggering events to determine if there are signature responses for particular emotion categories. For example, researchers might be interested in whether there are signature facial muscle movements that underlie specific emotions. To do this, they try to elicit the target emotion and measure facial muscle movements using EMG. If basic emotion theories are right, then there should be distinct and consistent bodily responses that are produced for each emotion. Commenting on production based studies of the face, Barrett notes “that the bulk of existing evidence has failed to support the hypothesis of distinct patterns of automatic facial EMG activity for anger, sadness, fear, and other emotion categories” (2006, pg.39). Strikingly, this evidence is consistent with non-human animal studies that show that these creatures “rarely produce involuntary, reflexive displays” (Barrett 2006, pg.39).

On reflection, these findings are fairly intuitive. People do not always smile when they are happy, and people can be happy while furrowing their brow. A single facial expression can be

[24] Barrett cites (Bradley & Lang, 2000; Russell, 2003; Shweder, 1993, 1994) on this point.

associated with multiple emotional categories.[25] This observation has led Seyfarth and Cheney (2003) to conclude that because a single facial expression, in isolation, underdetermines what emotion the emoter is in, facial expressions (as well as vocal behaviors) have low informational value and low referential specificity. Importantly, the claim is not that facial expressions underdetermine what valenced state (i.e., negative or positive affect) the emoter is in. People are good at reading valence from facial expressions. They are not good at reading emotion, and that's because specific facial expressions do not correspond one-to-one to particular emotional categories. Taken together, this evidence shows why the reader (and the general population!) has difficulty on the Ekman facial recognition task (figure 2). Emotions are perceptually fuzzy and complex, and this makes emotional concepts especially good candidates for linguistic influence. In particular, labels allow everyone to represent perceptually fuzzy distinctions between emotional categories in a simpler way. By being associated with the most diagnostic features of emotional categories, labels become robust category attractors that bind together disparate perceptual information into a discrete and unified whole, while also serving as direct routes to category knowledge. I now turn to reviewing two main lines of evidence of linguistic influence on emotional processing.

2.4 Label Superiority Effect in Emotional Recognition Tasks

Russell and Widen (2002) provide good evidence that there is a label superiority effect in emotional recognition. Children ages 2 to 7 were asked to categorize various facial expressions under three conditions: label, face, or both. In each condition, a box was assigned a label, a face,

[25] This does not mean that there is no connection between facial expressions and emotions. In general, there is a family resemblance inter- and intra-personally among facial expressions for particular emotional categories (see Barrett 2006, pp.37-39), but family resemblance is not what basic emotion theories predict.

or both and the children had to decide if the presented image went into the box or not. For example, if the box was labeled “happy”, the child had to decide if the image of a smiling person went in the box. In the face only condition, the boxes were assigned pictures of facial expressions and the children had to decide which box another picture of a face went into. Children in the label-only condition outperformed those in the other two conditions. The authors conclude, similar to Lupyan et al. (2007), that emotional labels facilitate categorization. Moreover, Russell and Widen comment: “These results invite further research on the assumption that facial expressions play a key role in the development of children’s understanding of emotion. For instance, facial expressions might play an important role, *but only in combination with situational, vocal or behavioral cues*” (2002, pg.48). The idea here is that other sources of information, including linguistic information, very likely play an important role in emotional concept development. This last point is consistent with the evidence reviewed above against basic emotion theories.[26]

Another line of evidence comes from an observation from Lindquist and Gendron (2013):

> Although response options are typically considered an innocuous feature of the task, it has been shown that including emotion words in the experiment inflates participants’ “accuracy” at identifying the emotion on the face…participants are generally better than chance at “accurately” identifying the emotion on the face when words are available in the experiment as response options (>63%). Studies that do not include emotion words in the task find substantially lower “accuracy” rates, however. For instance, the “accuracy”

[26] It’s also consistent with the larger theoretical point about labels-- that they bind together complex perceptual stimuli making concept learning and later conceptual deployment easier.

> of responses is quite low when participants are asked to freely label an emotional caricature without being given a set of words to choose from (e.g., between 7.5% and 54%)... More strikingly, providing labels can even cause participants to perceive a face as an instance of an "incorrect" emotion (e.g., participants perceive a scowling face as "disgust" rather than "anger" when the word "disgust" is available but "anger" is not (pg.68).

Lindquist and Gendron are talking about *healthy* subjects' performance here. With labels, healthy subjects are above 63% accurate, but without labels healthy subjects' performance ranges from a staggeringly low 7.5% to 54%. Again, echoing Barrett (2006), face-based emotional recognition is really hard for everyone! Taken together, these two lines of evidence seem to weigh in favor of the hypothesis I floated above: given the data suggesting that emotional categories are perceptually fuzzy, one would expect emotion concepts to be highly susceptible to linguistic influence, e.g., label superiority effect. Language scaffolds emotional processing by transforming perceptually fuzzy boundaries between categories into perceptually discrete boundaries, and they do this by binding together perceptually disparate information into unified wholes.

2.5 Taking Stock

At this point, the reader is likely wondering how this discussion bears on the mixed results of the Moebius syndrome studies. Here's the upshot: healthy subjects are not great at face-based emotional recognition tasks even with labels, but much worse without. Moebius patients are worse still without labels. Given that both populations have access to language, the only difference between them that could explain Moebius patients' worse performance, when labels are *not* provided, is the fact that they cannot simulate facial expressions.[27]

However, even if Moebius patients are benefitting from a label superiority effect, and that's what explains their ability to succeed when labels are provided, it is not clear that that helps the embodied theorist. This is because labels seemingly activate amodal representations, not sensorimotor reenactments. After all, labels are themselves amodal representations-- they are symbolic, abstract stand-ins that arbitrarily pick out worldly things. So it's not clear how this shows that Adams' premise 1 is false. Indeed, one might think that this lends *prima facie* support to Adams' premise 1. If Moebius patients succeed on face-based emotion recognition tasks *via labels activating amodal representations*, then embodiment is false.

I think this move is too hasty. In the next section I will be laying out my favored embodied theory of emotions-- psychological constructionism-- which is compatible with the label superiority effect, but rejects the idea that labels act to activate amodal representations. In other words, it could be true that Moebius patients succeed on face-based emotional recognition tasks using labels, but that does not show that embodiment is false. In fact, psychological constructionism predicts and explains Moebius patients' mixed performances.

3. Psychological Constructionism about Emotions

[27] Although see the end of section 1.

In this section, I will be introducing psychological constructionism in three parts. I will then show how this view explains the Moebius patients' mixed performance. I will begin with a brief (selective) survey on the philosophy of emotions.

3.1 Selective Survey of the Philosophy of Emotions

A natural objection to raise to Ekman-style basic emotion theory is that there are many more emotions than anger, fear, sadness, happiness, surprise, disgust, and contempt. What about love? Schadenfreude? Some basic emotion theorists distinguish between basic and non-basic emotions. There are, broadly speaking, two views on the relation between the basic and non-basic emotions: the unity thesis and the disunity thesis (Prinz 2004a; for review and discussion see Colombetti 2014). The unity thesis says that basic emotions are constitutive parts of non-basic emotions, and in virtue of this constitutive relation, basic and non-basic emotions are unified under the natural kind "emotions". The disunity thesis denies this-- basic emotions are different in kind from non-basic emotions such that different theories must be given for each.

A representative disunity theorist would be Griffiths (2008) who thinks that basic emotions are homologues to certain structures in non-human animals. Non-basic emotions, on the other hand, are uniquely human. This fundamental distinction, for Griffths, means that basic and non-basic emotions form two distinct natural kinds for which two different theories must be given.

A representative unity theorist would be Prinz (2004a) who thinks that all emotions are embodied appraisals. Embodied appraisals are perceptions of changes in the body that have the function of representing specific properties in the environment, e.g., danger in the case of fear, loss in the case of sadness. For Prinz, basic emotions are products of natural selection. That is, basic emotions are perceptions of bodily changes that have been naturally selected to have the function of representing a restricted set of properties. Non-basic emotions, on the other hand, are either "blends" of basic emotions or "recalibrated" basic emotions. By blending fear and surprise, for example, you might get the nonbasic emotion of awe. Prinz thinks that schadenfreude is a recalibrated basic emotion, namely joy. Joy has been recalibrated to represent others' suffering, instead of what it was naturally selected to represent.

Prinz's embodied view is inspired by William James (1884) and Carl Lange (1885). Both James and Lange thought that "emotions occur when the perception of an exciting fact causes a collection of bodily changes, and 'our feeling of the same changes as they occur IS the emotion'" (Prinz 2004b, pg.1). Prinz agrees that emotions are perceptions of bodily changes, but disagrees in that he thinks emotions are amenable to rational assessment. Hence, he thinks emotions are somatic, but also semantic-- they represent contents as embodied appraisals.[28] The account I will offer later-- psychological constructionism-- is an embodied account that rejects basic emotions.

Prinz's view, especially the notion of "blending" basic emotions to make non-basic ones, is interesting because it could perhaps accommodate the data that constructionists levy against basic theories. The idea is that Prinz can agree that many emotions are perceptually fuzzy-- the non-basic ones. The blending of basic emotions entails the overlap of emotional categories. Here, Prinz would agree that emotions are embodied, he would even agree that many emotions are

[28] Slaby (2014) offers an account of extended emotion as an extension of the extended cognition hypothesis.

perceptually fuzzy, but he would disagree that this means we should abandon basic emotion theory.

However, the perceptual fuzziness of emotional categories is not restricted to merely non-basic emotions. Constructionists think that perceptual fuzziness trickles down to even the supposedly basic emotions. In reaction to the findings canvassed in section 2.3, constructionists, "concluded that emotions are states tailor-made to a given context, which emerge when more elemental processes such as basic hedonic feelings or feelings of arousal are made meaningful using cognitive interpretation" (Lindquist 2013, pg. 357). Emotions are not basic, valence and affect are.

As evidence for this claim, consider the following. When researchers gather self-reports of subjective experiences of emotions and project them into a geometric space they do *not* find discrete clustering around the 6 putative emotional categories that would be indicative of distinct experiences among emotional categories. Rather, they find clustering around similarly valenced states, e.g., negatively valenced states like anger, fear, sadness. This suggests, comments Barrett (2006), that emotions are not experientially primitive (i.e., basic). That is, contra basic emotion theories, emotional experience is composed out of more basic psychological components, e.g., affect and arousal. Barrett goes further in her analysis:

> The main evidence that experiences of emotion can be broken down into more elemental bits is that when projected into geometric space, self-reports conform more or less to a circumplex structure (Feldman, 1995b; Remington, Fabrigar, & Visser, 2000; Russell,

> 1980; for a review, see Russell & Barrett, 1999). A circumplex structure emerges only when the objects in a correlation matrix (in this case, reports of emotion experience) are heterogeneous and psychologically reducible to a more basic set of properties (Guttman, 1957) (Barrett 2006, pg. 35).

There are two main takeaways from this line of evidence. The first is that if emotions are perceptually discrete categories with distinct and consistent behavioral and psychological responses, then subjective experience of emotions should reflect that. But self-reports do not conform to this categorical structure. Instead, and this is the second point, self-reports conform to more basic psychological components including affect and arousal. Self-reports cluster when projected into geometric space, just not in the way predicted by basic theories.

3.2 What is Psychological Constructionism? Part 1: Embodiment

In what follows, I will be discussing and motivating a particular kind of psychological constructionism about emotions that comes from Maria Gendron, Lisa Feldman Barrett, and Kristin Lindquist-- the conceptual act theory (or CAT). CAT is an embodied view that supplies a necessary role for language. I will argue that CAT provides a plausible explanation of the Moebius patients' performance. Hence, if true, the CAT account would constitute a counterexample to Adams' premise 1. In this sense, embodiment actually gains strong support from the Moebius syndrome studies.

If Barrett (2006) (and other constructionists) are right that emotions are composed out of more basic components, then how do emotions come about? CAT states that emotional experience and perception arise out of conceptual acts. A conceptual act, "is synonymous with categorization and relies on representations of prior experiences (i.e., knowledge, concepts, episodic memories). These prior experiences are represented as situation-specific reenactments of emotion in the brain's sensorimotor cortices" (Lindquist 2013, pg.362). Here, CAT is committed to the embodied approach to emotional concepts (that I outlined in the introduction). That is, what people know about emotion is "represented in part by the same neural substrates that have increased activity when a person actually experiences core affective feelings, engages in behaviors, and perceives sensory stimuli" (Lindquist 2013, pg. 362). This means that when people have an emotional experience, say of anger, they are engaging in a conceptual act that is embodied. They categorize their experience as "anger" by sensorimotor simulation.[29]

The locus of the conceptual act need not be limited to the subject's emotional experience. One can also turn the locus of conceptualization outward onto the world as when one reads emotion on someone else's face. In both cases, "emotions emerge in consciousness when people categorize ambiguous internal and external sensations as instances of discrete emotion categories" (Lindquist 2013, pg.360). And in both cases, the conceptual act will be an embodied sensorimotor reenactment.

It's worth contrasting CAT with Prinz's view. In principle, I think embodied appraisals are compatible with CAT. For example, consider the following illustration from Prinz:

[29] One might object here that categorization is different from an emotional experience, so CAT is mistaken. But this is to beg the question against CAT because that is exactly what CAT denies. It is only through categorizing your experience (or your perception of another) as a particular emotion (say, anger), that one experiences anger.

> Consider the chain of events leading to fear. Something dangerous occurs. That thing is perceived by the mind. The perception triggers a constellation of bodily changes. These changes are registered by a further state: a bodily perception. The bodily perception is directly caused by bodily changes, but it is indirectly caused by the danger that started the whole chain of events. It carries information about danger by responding to changes in the body. That further state is fear (2004a, pg.69).

CAT theorists can accept everything that Prinz says here with a caveat. Insofar as the subject engages in a conceptual act that results in their categorizing the experience as fear, "that further state is fear". Because Prinz is committed to basic emotion theory, he thinks that fear arises independent of categorization. But this is not the case for psychological constructionists, and CAT in particular (more on this below).

3.3 Part 2: Sensorimotor Reenactments are not Sufficient

As I reviewed in section 2.3, there is mounting evidence that emotions are not basic, discrete kinds. Instead, emotional experience and perception is highly variable both inter- and intra-personally across time. Moreover, the elements that make up a token emotional experience are many, including affect, facial muscle movements, vocal signals, body language, and context. The perceptual fuzziness of emotions has consequences for the human emotional conceptual system: "if the body and the situation help to constitute the mind, then the structure and content of the conceptual system for emotion should be grounded in the structure and content of emotional events as they naturally occur" (Barrett and Lindquist 2008, pg.17).

Because emotional experience and perception are fuzzy, emotional concepts will be too.[30] This is due to those concepts being embodied simulations of naturally occurring emotional experiences and perceptions. So just as experiencing anger on a particular occasion may involve, for example, shaking, furrowing the brow, shouting, and the urge to break stuff, activation of the concept "anger" might involve simulating these responses by subthreshold reactivation of the neural substrates that underlie those responses. I say "may" and "might involve" here because, as already noted, these responses can vary. If this is right, then sensorimotor reenactments will not be sufficient information to identify the target emotional category. Or, more cautiously, it might be enough information, but there is too much information (e.g., too many perceptual properties or too many subtle variations between reenactments) to identify the target category.

Contrast this view with the basic emotion theorists' view of emotional concepts. The basic emotion theorist hypothesizes, for example, that there are one-to-one correspondences between facial muscle configurations and emotional categories. If this is true, then knowing those specific configurations (or simulating those configurations) will be sufficient to identify the target emotional category. If emotional experience and perception are discrete, then so too will be emotional concepts. But as I argued in section 2.3, this is not the case.

3.4 Part 3: Labels as Binders

[30] I actually think that this explanation runs in the other direction too: that because our concepts are fuzzy, emotional experience and perception is fuzzy too. To foreshadow the next subsection, labels serve to make our emotional concepts discrete as well as emotional perception. Lupyan (2012) calls this bi-directional linguistic modulation "the label feedback hypothesis".

In section 2.4 I discussed a finding from Lindquist and Gendron (2013) that showed that without labels, healthy participants range from 7.5% to 54% accurate at emotional recognition tasks. However, with labels, healthy participants jump to above 63% accurate. CAT hypothesizes that language serves as a glue to bind together the various bodily responses into discrete categories.[31] In particular, emotional labels (e.g., "anger", "fear", "happy") organize humans' emotional conceptual system by creating a kind of filing system (see Prinz 2005, pg. 683). When a child undergoes an emotional experience (e.g., cries because big brother took away her rattle) and the parent asks "Why are you sad?," the parent is labeling the child's very complex experience as an instance of "sadness". As Lindquist (2013) puts it:

> Indeed, emotion categories might be acquired in childhood by bootstrapping situations and core affective feelings to the words used by adult caregivers (e.g., when mom and dad tell Joey not to be "sad" because of a broken toy, Joey learns that negative feelings following a loss are associated with the category "sadness" in his culture) (pg.362).

As the child begins to make strong associations between situations, bodily responses, and emotional labels, the child begins a filing system with folders for each emotional category.[32] Inside these folders are packets of embodied simulations that are strongly tied to that emotion.

---

[31] A similar point is made by Prinz (2005): "verbal labels serve as placeholders for ideas that are too complex to hold in one's mind all at once" (pg. 692). Due to emotional experiences and perceptions being so complex and variable, labels simplify. But importantly, as Prinz notes, the label by itself is not enough. The label "must be pinned down to the senses in order to be applied in the world" (pg.692). Hence why the labels serve as file folders that contain embodied representations.

[32] Widen (2013) reviews data suggesting that children's emotional conceptual system begins with two broad categories: good and bad. As these children learn emotional labels, their emotional conceptual systems begin to sharpen into discrete categories. It appears that the maturation of the emotional conceptual system happens concurrently with the expansion of emotional vocabulary.

For example, in the “sad” folder, there might be simulations for crying, hunching the shoulders, pain in the chest, frowning, etc.[33]

Moreover, as Clark (1998) and Lupyan et al. (2007) have argued, labels not only facilitate learning of categories, they also facilitate categorization even after that concept has been learned. The filing system metaphor will be helpful here. If one is trying to identify what emotion a person is experiencing, it helps when the conceptual landscape is organized into discrete folders with labels for each. When labels are provided on a task, those labels serve as direct routes to conceptual knowledge by activating embodied simulations. Consistent with the evidence canvassed in section 2, labels, according to CAT, bind together the highly variable instances of emotional categories, chunking them into discrete folders of embodied information. Sensorimotor reenactments are not sufficient for emotional recognition because language is needed to partition fuzzy emotional concepts into discrete concepts.

3.5 Explaining Moebius Patient Performance

CAT is an embodied, constructionist view of emotions. It says that emotions emerge in consciousness when a subject categorizes various bodily responses as a particular emotional category. Categorization, on this view, involves sensorimotor reenactments. But as we saw, sensorimotor reenactments will only get one so far. Because emotional categories are perceptually fuzzy and complex, language is needed to bind together the heterogeneous and

[33] Paul Pietroski (2005) makes a similar point about lexicalization (and its relation to concepts) in general saying that “lexicalization is a process in which diverse mental representations can be linked via the language system. Perhaps without lexicalization, representations that are different in kind cannot be combined to form a complex concept that is usable in human thought, but (luckily for us) the language system provides resources for creating certain ‘‘common denominators’’, which make it possible to create endlessly many complex mental representations with constituents that are typologically disparate” (pg. 271).

highly variable bodily responses into perceptually discrete categories. Once the emotional categories are partitioned, labels further serve to organize the conceptual landscape into file folders containing simulations of each emotional category. Emotion labels then serve as direct routes to emotional conceptual knowledge by activating embodied simulations. Let us now turn to how CAT explains the mixed results of the Moebius syndrome studies.

I'll start by explaining Moebius' patient performance when labels are provided. The reader will recall that in these studies, Moebius patients performed comparably to healthy controls. According to CAT, when labels are provided on a task they serve as direct routes to emotional conceptual knowledge by activating embodied simulations. Of course, the patients cannot simulate facial expressions themselves; however, they have seen people make facial expressions in various contexts. So they might project themselves into a context (e.g., episodic simulation) wherein that facial expression was made.[34] Moreover, because labels bind together various bodily responses to discrete categories, labels don't just activate facial muscle simulations. They activate simulations for various bodily responses typical of the target emotional category. So perhaps label-activated simulations provide richer information than an emotional face in isolation. Even though Moebius patients cannot make or simulate facial expressions, this does not preclude them from successful recognition. Embodied emotional concepts are not exhausted by facial muscle mimicry.

What does CAT say about Moebius patients' performance when labels are not provided? Recall that Moebius patients are severely deficient in these studies, so much so that the patients could

[34] For those skeptical that episodic simulation is embodied, Lindquist (2013, pg.362) explicitly mentions episodic memory as a form of sensorimotor reenactment. See also De Brigard (2014) and Schacter et al. (2015).

not respond at all to the task. The patient in Giannini et al.'s (1984) study described the experience as like trying to understand a foreign language. Firstly, without labels provided, subjects do not know where in the conceptual landscape to look for the target emotion. This is true for both Moebius patients and healthy controls. But Moebius patients have an additional handicap which is that facial muscle mimicry is not part of the content of their emotional concepts. So the information that they are given (the presented picture of an emotional facial expression) is not a guide (in isolation) to the target category. Contrastingly, for healthy controls, even though they are not provided labels, they *can* mimic the facial expressions. So rather than having no guide as to what emotion is being depicted in the presented picture, they have a rough guide. This explanation of Moebius patients' performance vindicates the report from Giannini et al.'s patient because having no information to use in reading a facial expression is, plausibly, very much like trying to understand a foreign language.

4 Limitations and Shortcomings

Before concluding the paper, I want to address some limitations of the analysis provided here. Firstly, I showed that there is a pattern in the empirical literature on Moebius patient performance on face-based emotion recognition tasks that is attributable to differences in experimental design. The pattern was that when patients are provided labels in the task, they perform comparably to controls, but when they are not provided labels, the patients are impaired. There is an exception to this trend that needs to be addressed.

Bate et al. (2013) did a face-based emotion recognition study with Moebius patients (n=6). In this study, 5 out of the 6 patients were impaired *even when emotion labels were provided*. This finding is in tension with not only the trend I highlighted, but also with the larger theoretical point I want to make: that labels allow the patients to perform comparably to controls by activating embodied conceptual knowledge (sans facial muscle mimicry).[35]

However, Bate et al. (2013) deployed an experimental design that involves time restriction on exposure to the target emotional face. Specifically, "stimuli are presented in a random order for five seconds per face, followed by a blank screen" (pg.e62656). After the blank screen, subjects must choose between the six basic emotions. The selection portion of the task is not timed. Here, subjects hold the presented face in working memory, and then select the correct emotion label. This difference in experimental design adds computational costs to the task. For one, it is well known that storage in working memory requires attentional resources (Cowan 2005). And secondly, representations in working memory are disposed to decay due to issues in encoding (Baddeley and Hitch 1974). If, in the absence of labels, Moebius patients must rely on facial expressions to identify the target emotion, and they cannot simulate facial expressions, then it's not clear that patients would be able to hold an embodied representation of an emotional face in working memory-- especially not the details of the facial expression. This is conjecture, but I suspect that encoding of the facial expression among Moebius patients is degraded due to their facial paralysis.

Some support for this conjecture actually comes from the Warrington memory task in Calder et al 's (2000). In this task, subjects are shown fifty faces individually for approximately three

[35] Thanks to an anonymous reviewer for this!

seconds each. During those three seconds, subjects are to report whether the face looks pleasant or unpleasant. Then, they are given fifty pairs of faces (in each pair, one is from the training set and one is *not* from the training set). Subjects are asked to identify the face from the training set. The patient mean score for this task was 36.67/50.00 whereas the control mean score was 43.60/50.00. The significant impairment on this task suggests that Moebius patients might have encoding impairments for emotional faces.

More straightforward evidence for the claim that Moebius patients have difficulty encoding representations of emotional faces in visual working memory comes from Gambarota et al. (2020). These authors issued a delayed estimation task with no verbal component. Here, subjects were presented with a to-be-memorized emotional face of a certain intensity for approximately 1 second. They were then shown an array of emotional faces and instructed to select the face they were initially presented with. Gambarota et al. reported a significant difference between Moebius subjects and controls, with Moebius subjects deficient. The authors conclude that Moebius "participants built lower quality representations of the intensity of emotional expressions when compared to healthy participants. These findings support the role of sensorimotor simulation in improving the quality of emotional representations of facial expressions during early stages of processing" (pg. 2).

One final concern I'd like to address came from an anonymous reviewer. The concern is that it's possible that labels are *not* cues to embodied simulations at all. Rather, it could be that labels allow the subjects to guess among the choices provided. For example, if provided no labels the task may be more difficult than if 4 labels were provided simply because in the latter case one has a 25% chance of getting it right by guessing. If one is given 2 labels, it's even easier. In other words, the objection agrees that emotions are perceptually fuzzy, but disagrees that emotions are embodied. So all that labels do is reduce the search space for the relevant emotional category.

There seems to be two subtly different objections here that I'm going to separate out. The first is that Moebius subjects are guessing when labels are provided. The second is that there is a label superiority effect, but that the explanation for the superior performance is not an embodied one. The former objection is consistent with embodiment because guessing when labels are provided does not presuppose nor entail that embodiment (in the sense I have defined it here) is false. For example, labels might activate embodied simulations, but *only* facial muscle simulations. In this case, Moebius patients would not be helped by labels being provided, so they guess among the options provided by the labels. The latter objection outright denies that labels activate *any* embodied simulations, but grants my arguments for the perceptual fuzziness of emotional categories. The implication for Moebius' patient performance is that provided labels reduce the search space of emotional categories, facilitating categorization, but not via embodied simulations. I'll respond to the guessing objection first, then handle the reducing the search space objection.

If it's possible to guess when labels are given, it's also possible to guess when labels are not given. Afterall, Moebius patients have an emotional vocabulary just like controls do. However, as reviewed, when labels are *not* provided Moebius patients fail to respond to the task. This suggests that guessing is not at play here.

Still, the objector could grant this point and suggest that a forced-choice paradigm makes it *more likely* that subjects would guess. This is certainly possible; however, Giannini et al. (1984) used labels for amounts of winnings that the subject was instructed to select on the basis of video clips of people's emotional reactions at a slot machine. As reported, the subject failed to respond to the task. Granted, this is a sample size of 1, so it's hard to generalize. But the point is that since the subject did not guess here, why think they are guessing in the other experiments?

However, I grant that it is certainly possible that subjects are guessing in these tasks. I can't rule that out, but I think it would be surprising. Here's why: In the studies in which controls and Moebius subjects were provided labels to match to emotional faces, both groups performed comparably. Suppose that both groups were guessing in these tasks. It would be incredibly coincidental that they ended up performing comparably to each other. Further, suppose only the Moebius subjects were guessing-- the controls were not. Still, it would be incredibly coincidental that patient performance was comparable to the controls. On the other hand, if you buy my explanation, it is not coincidental at all. Emotion labels serve as direct routes to emotional concept knowledge through activating embodied simulations in both groups.

Now for the second objection. Is it possible that the label superiority effect, in the face-based emotion recognition tasks, is best explained by a non-embodied reduction of the search space of emotional categories? At the end of section 2, I considered this explanation for the label superiority effect. There, I said that it's not clear how the label superiority effect helps the embodied theorist because labels, *prima facie*, activate non-embodied, amodal representations.

Thus, it is not clear how this shows that Adams' premise 1 is false. Indeed, you might think that the label superiority effect supports Adams' premise 1. In responding to this, I showed how the CAT accommodates the label superiority effect. However, you might be unsatisfied with this response because this only shows that we have two competing explanations of Moebius patient performance– one that is embodied and one that is not. Therefore, we need positive reasons for thinking that language, and in particular labels, activate embodied simulations.

The best evidence for my case would likely be imaging data on Moebius patients and controls during an emotional recognition task showing sensorimotor activation when labels are provided. At the time of writing this, I know of two imaging studies with Moebius patients (Japee et al. 2022; Sessa et al. 2022), but neither used emotion labels. Additionally, both of these studies deployed a match-to-sample paradigm. While interesting, this paradigm does not probe the target phenomenon-- emotion labels'(in)ability to induce embodied simulations. The lack of evidence however is not evidence of a lack. This is an area where further work should be done.

However, there is positive evidence for embodied simulations in language comprehension more generally. As briefly mentioned in the introduction, the action-sentence compatibility effect is a well established finding (Bergen 2012). There is also evidence of embodied simulation in metaphor comprehension (Carston 2010; Gibbs, Jr. 2006; Green 2017; Lakoff and Johnson 2008), dialogue processing (Pickering and Garrod 2004; Zwaan and Radvansky 1998), and grammar processing (Bergen 2012).

I think most relevant to present purposes is recent work on polysemy resolution. Polysemy can be understood as an instance of lexical ambiguity in which “words have a single meaning that can be modulated to fix distinct denotations depending on context” (Quilty-Dunn 2021; pg.164). For example, “Jack chugged the can, and crushed it on his head.” has two occurrences of “can". In the first instance, “can" denotes the liquid that Jack chugged, but in the second instance “can" denotes the smashed aluminum object. While this sentence seems felicitous, some sentences involving polysemes are infelicitous: “This chicken is scrumptious and chirpy!”. Here, the first occurrence of “chicken” denotes the food item while the second denotes the farm animal.

Michelle Liu (forthcoming) argues that polysemy resolution depends upon embodied simulations. In particular, she argues that infelicitous polysemy-involving sentences entail conflicting simulations whereas felicitous ones do not. The chicken sentence entails conflicting simulations because it involves the comprehender simulating chicken-as-meat and chicken-as-animal, but our everyday experiences with chicken meat do not involve it being chirpy. Contrastingly, in the Jack sentence, can-as-liquid and can-as-vessel overlap in our everyday experiences. This sentence does not involve conflicting simulations because it does not demand the comprehender to simulate a fundamentally different entity (as in the chicken sentence).

I have argued that emotional categories are perceptually fuzzy and complex because emotional experience and perception are highly variable inter- and intra-personally through time. Emotion labels are needed, according to CAT, to bind together those heterogenous representations into unified wholes. This means that “angry” can be used to denote, for example, a red-faced person shouting and throwing furniture, but also a person quietly standing in a corner with their arms crossed. Emotion labels can be used to denote different objects depending on context. If the best explanation of polysemy resolution involves embodied simulations, then we should expect that comprehending emotional labels also involve embodied simulations. Again, though, more work should be done on this front.

5. Conclusion

Fred Adam's argument against embodied concepts has been shown to be unsound. Premise 2, which states that Moebius patients succeed on emotional recognition tasks has been shown to be misleading. These patients sometimes succeed on such tasks, and when they do they are provided emotional labels to match to the facial expression. When these patients are not provided labels, they fail to respond to the task. In explaining these performance differences by reference to psychological constructionism, I have shown that Adam's premise 1 is false. It is not the case that if Moebius patients succeed on face-based emotional recognition tasks, then embodiment is false. Indeed, CAT, as I showed, is an embodied view of emotions. Because CAT can explain Moebius patients' successful performance, embodiment is not threatened. In fact, embodied views, and in particular CAT, gains motivation.

Besides defending the embodiment of emotion concepts, we have learned some things about the nature of emotion and the language-cognition-perception interface. I discussed overwhelming evidence that shows that, contra basic theories, emotions are constructed out of various more basic psychological components. For this reason, emotions are perceptually fuzzy and complex. I also discussed how perceptually fuzzy categories are (at least) easier to learn with the help of labels. Moreover, once those concepts are learned, labels facilitate categorization by simplifying the environment and organizing our conceptual landscape. Despite common sense dictating that emotions are innate, pre-linguistic, and subserved by distinct mechanisms, emotions are actively (albeit implicitly) constructed, via conceptual acts, for the context that the subject finds themself in.

REFERENCES

Adams, F. (2010). Embodied cognition. *Phenomenology and the Cognitive Sciences*, *9*(4), 619-628.

Allport, F. H. (1924). *Social psychology*. Boston, Houghton.

Baddeley, A., & Hitch, G. (1974). Working Memory: In Bower GA. The psychology of learning and cognition.

Barrett, L. F. (2006). Are emotions natural kinds?. *Perspectives on psychological science*, *1*(1), 28-58.

Barrett, L. F., & Lindquist, K. A. (2008). The embodiment of emotion. *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*, 237-262.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, *22*(4), 577-660.

Bate, S., Cook, S. J., Mole, J., & Cole, J. (2013). First report of generalized face processing difficulties in möbius sequence. *PloS one*, *8*(4), e62656.

Bell, C., Nevitt, S., McKay, V. H., & Fattah, A. Y. (2019). Will the real Moebius syndrome please stand up? A systematic review of the literature and statistical cluster analysis of clinical features. *American Journal of Medical Genetics Part A*, *179*(2), 257-265.

Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. Basic Books.

Bergen, B., & Feldman, J. (2008). Embodied concept learning. In *Handbook of Cognitive Science* (pp. 313-331). Elsevier.

Bertoux, M., Duclos, H., Caillaud, M., Segobin, S., Merck, C., de La Sayette, V., ... & Laisney, M. (2020). When affect overlaps with concept: emotion recognition in semantic variant of primary progressive aphasia. *Brain*, *143*(12), 3850-3864.

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2016). Concepts as semantic pointers: A framework and computational model. *Cognitive science*, *40*(5), 1128-1162.

Bogart, K., & Matsumoto, D. (2010). Facial mimicry is not necessary to recognize emotion: Facial expression recognition by people with Moebius syndrome. *Social neuroscience*, *5*(2), 241-251.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology.

Buckner, C. (2018). Empiricism without magic: Transformational abstraction in deep convolutional neural networks. *Synthese*, *195*(12), 5339-5372.

Calder, A. J., Keane, J., Cole, J., Campbell, R., & Young, A. W. (2000). Facial expression recognition by people with Möbius syndrome. *Cognitive neuropsychology*, *17*(1-3), 73-87.

Caramazza, A., & Mahon, B. Z. (2006). The organisation of conceptual knowledge in the brain: The future's past and some future directions. *Cognitive Neuropsychology*, *23*(1), 13-38.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and brain sciences*, *25*(6), 657-674.
Carston, R. (2010, October). XIII—Metaphor: Ad hoc concepts, literal meaning and mental images. In *Proceedings of the Aristotelian society* (Vol. 110, No. 3_pt_3, pp. 295-321). Oxford, UK: Oxford University Press.
Casasanto, D., & Lupyan, G. (2011). Ad hoc cognition. In *The 33rd Annual Conference of the Cognitive Science Society [CogSci 2011)* (p. 826). Cognitive Science Society.
Chemero, A. (2011). *Radical Embodied Cognitive Science*. MIT Press.
Clark, A. (1998). Magic words: how language augments. *Language and Thought: Interdisciplinary Themes*, 162.
Clark, A., & Chalmers, D. (1998). The extended mind. *analysis*, *58*(1), 7-19.
Colombetti, G. (2014). *The feeling body: Affective science meets the enactive mind*. MIT press.
Cowan, N. (2005). Working memory capacity limits in a theoretical context. In *Human learning and memory: Advances in theory and application. The 4th Tsukuba international conference on memory* (pp. 155-175). Mahwah, NJ: Erlbaum.
Darwin, C. (1872). *The Expression of Emotions in Man and Animals.* Murray.
De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, *191*(2), 155-185.
De Stefani, E., Ardizzi, M., Nicolini, Y., Belluardo, M., Barbot, A., Bertolini, C., ... & Ferrari, P. F. (2019). Children with facial paralysis due to Moebius syndrome exhibit reduced autonomic modulation during emotion processing. *Journal of Neurodevelopmental Disorders, 11*(1), 1-16.
De Stefani, E., Nicolini, Y., Belluardo, M., & Ferrari, P. F. (2019). Congenital facial palsy and emotion processing: The case of Moebius syndrome. *Genes, Brain and Behavior*, *18*(1), e12548.
Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain* (No. 202). WW Norton & Company.
Dennett, D. C. (1984). Cognitive wheels: The frame problem of AI. *Minds, machines and evolution*, 129-151.
Dennett, D. C. (2015). *Elbow Room, new edition: The Varieties of Free Will Worth Wanting*. mit Press.
Deonna, J., & Teroni, F. (2012). *The emotions: A philosophical introduction*. Routledge.
DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in cognitive sciences*, *11*(8), 333-341.
Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*(3), 412-431.
Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, *1*, 242.
Edelman, S., & Intrator, N. (1997). Learning as formation of low-dimensional representation spaces. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society, Erlbaum, Mahwah, NJ* (pp. 199-204).
Edelman, S., & Intrator, N. (1997). Learning as extraction of low-dimensional representations. *Psychology of learning and motivation*, *36*, 353-380.
Ekman, P. (1992). Facial expressions of emotion: New findings, new questions.
Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion review*, *3*(4), 364-370.
Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental psychology and nonverbal behavior*, *1*(1), 56-75.

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive science*, *33*(4), 547-582.

Farina, M. (2021). Embodied cognition: dimensions, domains and applications. *Adaptive Behavior*, *29*(1), 73-88.

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard university press.

Gallagher, S. (2011). Interpretations of embodied cognition. In W. Tschacher and C. Bergomi (eds.), *The Implications of Embodiment: Cognition and Communication*. Exeter: Imprint Academic.

Gambarota, F., Luria, R., Pastore, M., Ferrari, P. F., & Sessa, P. (2020). Deficits of visual working memory representations of emotional facial expressions in patients with congenital facial palsy. *psyarXiv* 2020.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*(2), 314.

Giannini, A. J., Tamulonis, D., Giannini, M. C., Loiselle, R. H., & Spirtos, G. (1984). Defective response to social cues in Möbius' syndrome. *Journal of Nervous and Mental Disease*.

Gibbs Jr, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind & Language*, *21*(3), 434-458.

Green, E. J. (2020). The perception-cognition border: A case for architectural division. *Philosophical Review*, *129*(3), 323-393.

Green, M. (2017). Imagery, expression, and metaphor. *Philosophical Studies*, *174*(1), 33-46.

Gregor, K., Besse, F., Jimenez Rezende, D., Danihelka, I., & Wierstra, D. (2016). Towards conceptual compression. *Advances In Neural Information Processing Systems*, *29*.

Griffiths, P. E. (2008). *What emotions really are: The problem of psychological categories*. University of Chicago Press.

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific reports*, *8*(1), 23.

James, W. (1884). What is an emotion? *Mind, 9,* 188-205.

Japee, S., Jordan, J., Licht, J., Lokey, S., Moebius Syndrome Research Consortium, Chen, G., ... & Ungerleider, L. G. (2022). Inability to make facial expressions dampens emotion perception. *bioRxiv*, 2022-10.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*(2), 113-139.

Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. *Psychological stress: Issues in research*, 14-37.

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.

Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120392.

Lange, C. G. (1885). *Om sindsbevaegelser: et psyko-fysiologisk studie.* Kjbenhavn: Jacob Lunds. Reprinted in *The emotions*, C. G. Lange and W. James (eds.), I. A. Haupt (trans.) Baltimore, Williams & Wilkins Company 1922.

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review*, *5*(4), 356-368.

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological science*, *19*(9), 898-903.

Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception.

Emotion Review, *5*(1), 66-71.

Lindquist, K. A., Gendron, M., Barrett, L. F., & Dickerson, B. C. (2014). Emotion perception, but not affect perception, is impaired with semantic memory loss. *Emotion*, *14*(2), 375.

Liu, M. (forthcoming). Mental simulation and language comprehension: The case of copredication. *Mind & Language.* https://philarchive.org/archive/LIUMSA-2

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in psychology*, *3*, 54.

Lupyan, G., & Bergen, B. (2016). How language programs the mind. *Topics in cognitive science*, *8*(2), 408-424.

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science*, *24*(4), 279-284.

Lupyan, G., & Lewis, M. (2019). From words-as-mappings to words-as-cues: The role of language in semantic knowledge. *Language, Cognition and Neuroscience*, *34*(10), 1319-1337.

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, *141*(1), 170.

Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in cognitive sciences*, *24*(11), 930-944.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological science*, *18*(12), 1077-1083.

Mahon, B. Z. (2015). What is embodied about cognition?. *Language, cognition and neuroscience*, *30*(4), 420-429.

Matsumoto, D., & Ekman, P. (2004). The relationship among expressions, labels, and descriptions of contempt. *Journal of personality and social psychology*, *87*(4), 529.

Matsumoto, D., & Ekman, P. (2006). Multi-ethnic facial expression set. *Unpublished images*.

McClamrock, R. (1995). *Existential cognition: Computational minds in the world*. University of Chicago Press.

Michael, J., Bogart, K., Tylén, K., Krueger, J., Bech, M., Østergaard, J. R., & Fusaroli, R. (2015). Training in compensatory strategies enhances rapport in interactions involving people with Möbius syndrome. *Frontiers in Neurology*, *6*, 213.

Michel, C. (2021). Overcoming the modal/amodal dichotomy of concepts. *Phenomenology and the Cognitive Sciences*, *20*(4), 655-677.

Möbius, P. J. (1888). Ueber angeborene doppelseitige Abducens-Facialis-Lahmung. *Munch med wochenschr*, *35*, 91-94.

Newen, A., De Bruin, L., & Gallagher, S. (Eds.). (2018). *The Oxford handbook of 4E cognition*. Oxford University Press.

Nicolini, Y., Manini, B., De Stefani, E., Coudé, G., Cardone, D., Barbot, A., ... & Ferrari, P. F. (2019). Autonomic responses to emotional stimuli in children affected by facial palsy: The case of Moebius syndrome. *Neural plasticity*.

Núñez, R. (2008). Conceptual metaphor, human cognition, and the nature of mathematics. *The Cambridge handbook of metaphor and thought*, 339-362.

Panksepp, J. (2011). The basic emotional circuits of mammalian brains: do animals have affective lives?. *Neuroscience & Biobehavioral Reviews*, *35*(9), 1791-1804.

Perry, L. K., & Lupyan, G. (2014). The role of language in multi-dimensional categorization: Evidence

from transcranial direct current stimulation and exposure to verbal labels. *Brain and Language*, *135*, 66-72.
Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169-190.
Pietroski, P. (2005). Meaning before truth. *Contextualism in philosophy: Knowledge, meaning, and truth*, *255*, 302.
Pons, F., Harris, P. L., & Doudin, P. A. (2002). Teaching emotion understanding. *European Journal of Psychology of Education*, *17*(3), 293-304.
Prinz, J. J. (2004a). *Gut reactions: A perceptual theory of emotion*. Oxford University Press.
Prinz, J. J. (2004b). Emotions embodied. *Thinking about feeling: Contemporary philosophers on emotions*, 44-59.
Prinz, J. J. (2004c). *Furnishing the mind: Concepts and their perceptual basis*. MIT press.
Prinz, J. J. (2005). The return of concept empiricism. In *Handbook of categorization in cognitive science* (pp. 679-695). Elsevier Science Ltd.
Quilty‐Dunn, J. (2021). Polysemy and thought: Toward a generative theory of concepts. *Mind & Language*, *36*(1), 158-185.
Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature‐label‐order and their implications for symbolic learning. *Cognitive science*, *34*(6), 909-957.
Rasmussen, L. K., Rian, O., Korshoej, A. R., & Christensen, S. (2015). Fatal complications during anaesthesia in Moebius syndrome: a case report and brief discussion of relevant precautions and preoperative assessments. *International Journal of Anesthesiology & Research*, *3*(6), 116-118.
Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145.
Russell, J. A., & Widen, S. C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*(1), 30-52.
Scarantino, A., & Griffiths, P. (2011). Don't give up on basic emotions. *Emotion Review*, *3*(4), 444-454.
Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of learning and memory*, *117*, 14-21.
Sessa, P., Schiano Lomoriello, A., Duma, G. M., Mento, G., De Stefani, E., & Ferrari, P. F. (2022). Degenerate pathway for processing smile and other emotional expressions in congenital facial palsy: an hdEEG investigation. *Philosophical Transactions of the Royal Society B*, *377*(1863), 20210190.
Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual review of psychology*, *54*(1), 145-173.
Shapiro, L. (2019). *Embodied cognition*. Routledge.
Shariff, A. F., & Tracy, J. L. (2011). Emotion expressions: On signals, symbols, and spandrels—A response to Barrett (2011). *Current Directions in Psychological Science*, *20*(6), 407-408.
Shweder, R. A. (1994). You're not sick, you're just in love": Emotion as an interpretive system. *The nature of emotion: Fundamental questions*, 32-44.
Shweder, R. A., Haidt, J., Horton, R., & Joseph, C. (1993). The cultural psychology of the emotions. *Handbook of emotions*, 417-431.
Slaby, J. (2014). Emotions and the extended mind. *Collective emotions*, 32-46.
Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In *Advances in*

psychology (Vol. 93, pp. 233-272). North-Holland.

Stewart, T. C., Bekolay, T., & Eliasmith, C. (2011). Neural representations of compositional structures: Representing and manipulating vector spaces with spiking neurons. *Connection Science*, *23*(2), 145-153.

Valeriani, L., Doimo, D., Cuturello, F., Laio, A., Ansuini, A., & Cazzaniga, A. (2023). The geometry of hidden representations of large transformer models. *arXiv preprint arXiv:2302.00294*.

Vannuscorps, G., Andres, M., & Caramazza, A. (2020). Efficient recognition of facial expressions does not require motor simulation. *Elife*, *9*, e54687.

von Graefe A. In: von Graefe A, Saemisch T, eds. Handbuch Der Gesamten Augen-heilkunde. Vol 6. Leipzig, Germany: Engelmann; 1880:60-67.

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72-77.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, *123*(2), 162.