# IMPROVING CONCEPTUAL ENGINEERING BY DIFFERENTIATING THE FUNCTIONS OF CONCEPTS

Rogelio MIRANDA VILCHIS

ABSTRACT: The leading assumption of this paper is that we can improve the methodology of conceptual engineering if we differentiate between the different functions of our concepts. There is a growing body of research that emphasizes the revisionist virtues of conceptual engineering against the descriptive task of conceptual analysis. Yet, it also has faced severe critiques. Among the difficulties raised are the problems of conceptual identification and continuity. That is why several philosophers are trying to resolve these problems and improve the methodology by calling attention, for example, to the functions that concepts can play. I follow this line of argument and argue that we can increase the chances of success if we also clarify and differentiate them. Identifying and assessing the relationship between functions will help us avoid confusion, inconsistencies, and possible verbal disputes. Doing this not only serves our theoretical and practical purposes but helps us reconsider the potentialities and limits of the conceptual engineering program.

KEYWORDS: conceptual engineering, concepts, functions, emotion concepts, metaphilosophy

## 1. Introduction

Despite some difficulties found along the way, conceptual engineering has been gaining more and more attention from the philosophical community (Burgess & Plunkett 2013a; 2013b; Cappelen 2018; 2020; Plunkett & Cappelen 2020; Floridi 2011). Frequent worries like the discontinuity problem and the challenge posed by externalism are receiving a lot of different and interesting responses (Bach 2016; 2019; Brigandt 2010; Koch 2021; Riggs 2020; Sawyer 2020a; 2020b). Among them, the functionalist response has gotten considerable attention (Brigandt 2010; Nado 2019; Prinzing 2017; Kelp & Simion 2019), but I believe that this solution brings in another set of difficulties. I present these new problems and argue that we can resolve, or at least clarify, them by differentiating between the functions of our concepts. This differentiation will shed new light on familiar and avoidable confusions, as well as on the more general problems that threaten the methodology

of conceptual engineering. We will get a clearer picture of the place that this methodology has in philosophical theorizing.

In the second section of this paper, I distinguish between conceptual engineering and what Nado calls "functional conceptual engineering." I show that the latter avoids some of the difficulties that the first faces. It will also be evident that functional conceptual engineering must face a new set of problems. In the third section, I suggest that the best strategy to answer this difficulty is to distinguish the varied functions of our concepts. The most significant distinction is between the representational function and other roles. Finally, the goal of the fourth section is to highlight the implications of this distinction, taking the "theory of constructed emotion" as a case study. I argue that if we do not carry out any distinction between our emotion concepts, we must deal with a host of unnecessary confusions, inconsistencies, and perhaps verbal disputes. On the other hand, making a distinction can prevent these problems, and consequently, improve the conditions to make conceptual engineering a more fruitful enterprise.

## 2. From Conceptual Engineering to Functional Conceptual Engineering

Conceptual engineering is an increasingly popular way to make sense of part of what philosophers are doing when they philosophize. Unlike the essentially more descriptive goals of conceptual analysis, conceptual engineering is a method that enables us to improve our concepts and, therefore, to make philosophical progress that would be otherwise impossible. Scientists have improved concepts like MASS, GENE, and MENTAL DISORDER; philosophers are trying to do the same with concepts like TRUTH (Chihara 1979; Eklund 2002; Scharp 2013; 2014). Yet, the Strawsonian objection against Carnap's explication (Strawson 1963) still carries weight. One of the main worries is that revising our concepts seems to imply a change of subject. There is also the problem of defining what a concept is.

There is a lively debate on what concepts are (Margolis & Laurence 2014) and, surprisingly, on whether they even exist (Machery 2009). Given that conceptual engineering is in the business of improving, eliminating, or developing new concepts, the methodology of conceptual engineering seems to be subject to similar worries. However, these criticisms are typically directed to the adequacy of specific characterizations of concepts: whether it is better to characterize concepts as exemplars, prototypes, theories, or a combination of them. A general characterization would not be subject to the same worries. We can avoid them if we employ a general characterization of concepts like "multiple realizable functional kinds" (Isaac 2020).

We can consider concepts as cognitive tools with diverse functions like the representation of the world, the improvement of social practices, and as a way to stimulate our cognitive capabilities. Although there are influential non-representational views about the concepts employed in science, philosophy, and ordinary life (Blackburn et al. 2013; Rorty 1979; 1990), the predominant view is that we can think about the world because concepts play a representative role (I think that the arguments that I present below could apply to a non-representational take on concepts. But I lack the space to defend that view here). We ordinarily conceive of concepts as representational devices (Plunkett & Cappelen 2020) that track specific features of the world. They do this job, no matter whether they provide an accurate representation or not. When they track real worldly traits, I call them "accurate-representational concepts" (for simplicity, I usually refer to them only with the adverb "accurately" and kindred modifiers). The concepts CHAIR, HEART, and PERSON, for instance, represent certain traits of reality: chairs, hearts, and persons. But concepts can also represent inexistent things like centaurs, story characters (e.g., Sherlock Holmes), and chemical compounds made of exotic elements like XYZ. I call these concepts "merely referential" or "representational" (of course, both kinds of concepts are representational, but only one accurately represents).

Concepts also play a practical, interactive role. We use them to interact with the world and, if successful and with their assistance, we can modify reality. Concepts have the power to change the attitudes and thoughts of individuals and, therefore, their behavior. All that is required is that a substantive portion of the population entertains certain concepts to produce sociopolitical and even economic changes (Hacking 1995, 1999). Certain concepts bring about the social and political roles described in those very concepts. Indeed, believing that there is only one way of economic arrangement will cause that arrangement to exist.

But, even if a broad functional definition along these lines can avoid the problem of the existence of concepts, conceptual engineering still faces several challenges. Among them, we find Strawson´s challenge: Conceptual amelioration seems to imply a change of subject. We face a discontinuity problem. But this difficulty can also be surmounted by a functional characterization of concepts. Besides Isaac (2020), several philosophers have proposed to shift our attention from referents to functions (Brigandt 2010; Nado 2019; Prinzing 2017; Thomasson 2020). As I indicated above, one well-known function of concepts is to accurately represent distinct aspects of reality. Other functions are the encouragement of particular patterns of social behavior, the stimulation of creativity, aesthetic pleasure, etc. Although the reference of concepts usually varies for several reasons

(like semantic drift and conceptual engineering), they preserve their function. The function of a concept can remain the same even if its reference and meaning have changed. It is, then, more advisable to focus our attention on functional conceptual engineering instead of the more usual form of conceptual engineering that is more closely concerned with referents and meanings.[1] The concept FISH, for example, now excludes whales from its reference but its function is the same: to provide a taxonomy of animals (indeed, we can say that the new concept FISH performs this function better than the old one).

This functional approach seems to be a promissory way to meet Strawson's challenge. The problem is that if we do not discriminate between the different functions that a concept could have, we could fall prey to idle theoretical disputes or merely verbal disputes. We must get clear on questions like the following: Does the concept *c* have the function of representing some aspect of reality? Does it stimulate a particular kind of behavior? Is it intended just for entertainment?

Consider the concepts that appear in fairy tales and myths. Surely all, or most of them, lack a referent, so they cannot be employed to represent some aspect of reality. Concepts such as MINOTAUR or CYCLOPS continue to play a recreational function. These concepts are also useful to stimulate our imagination, and consequently, to foster our creativity. They could be extremely helpful for storytellers, moviemakers, and fiction writers. Other concepts like FREE MARKET, SUPPLY, and DEMAND have produced a new set of economic and social practices.

From these examples, we can see that concepts can play many different functions, but we must ask ourselves if this is an advantage or a disguised obstacle for our engineering goals. This situation is potentially dangerous because we can mistake one function for another. This is not a problem peculiar to philosophy but can arise in every theoretical domain. If we employ, for example, the term *minotaur* with the clear intention to describe something in the world, we would not achieve our goal. There is not anything in the world that can be referred to by this expression. It cannot play an accurate-representational function. Analogously, other concepts can play a practical role, but they might not perform an accurate-representational role. We can employ the concept MONEY and still hold genuine worries about the existence of the reference of this concept (Barrett 2017, 133-4;

[1] According to Prinzing (2017), concepts are functional kinds that are preserved through conceptual revision of, for example, extension. Nado (2019) rightly points out that it is very troublesome to define what a function is. I remain neutral on what exactly it is. After all, we do not need a precise definition to insist, as I do in the next sections, that there are different kinds of functions.

Goldstein 2020). We can mistake the accurate-representational function of a concept with its aesthetical, recreational, or social role (undoubtedly, we can entertain concepts with the single purpose of playing a game: a concept-guessing game).

We can see now that, despite all the virtues that functional conceptual analysis may bring about, we need to address some problems. In section fourth, I analyze these and other difficulties in more detail and show how we can avoid and lessen some of the potential harmful effects. But first, we need to distinguish the different kinds of conceptual functions.

## 3. Differentiating the Diverse Functions of Concepts

In the previous section, we have seen that conceptual engineers can overcome the discontinuity problem if they focus on the functions of concepts. This shift of focus is very promising, but we have seen that we face a new challenge: the potential confusion and inconsistency brought about by distinct and conflicting functions or by the failure to recognize the proper function of a given concept. In this section, I suggest that the obvious way to overcome these problems is by differentiating as far as possible the different kinds of functions that concepts play.[2]

Among other goals, philosophy and science are theoretical enterprises that try to describe reality accurately. Their sentences must be *true* or *approximately true* to achieve this goal. They must *accurately represent the world*. Like scientists, many philosophers seek to accurately describe what the world is like (Eklund 2014, 295). However, we saw that we can make one valuable distinction between the following two functions: representation and accurate-representation. Many concepts that do not represent the world accurately (like MINOTAUR) are, nonetheless, accomplishing their representational function (representing a minotaur. It does not matter that it is a fictional entity that exists only in people's minds). The accurate-representational function can be part of the representational function but not vice versa. In any case, as we will see, the primary function of concepts is to represent (I leave aside the problem of whether a representing function variant can work in non-correspondence theories of truth). They may represent non-existent things, but they owe their existence as concepts to this representational function. All other functions derive from this primordial role.[3]

[2] The differentiation may not be completely clear, but we must try to do the best we can even if "it is unlikely that description and prescription can be clearly separated" (Griffiths 2002, 908).

[3] I think that this emphasis on the representational function – instead of the accurate-representational one – is valuable because it allows us to avoid skepticism about truth. Insisting on the representational role is weaker (and hence more useful for the purposes of this paper)

At this point, someone (e.g., Nado 2019) may object that – concerning the accurate-representational function – only sentences can be true or false. But this is not problematic. Although isolated concepts cannot have a truth value, they can represent a feature of the world. We say that they are adequate concepts if they accurately represent what is in the world and, because of that property, they help to make sentences true. And that is all we need to maintain that concepts have an accurate-representational function that is not only very important but that, most of the time, grounds the possibility of many other functions. But what are those other functions?

I do not attempt to offer a sharp and complete taxonomy of concepts' functions because there are many, and some might not have been invented yet. A rough characterization is sufficient for the present purposes. Based on the mere representational function, we can find the following ones: 1) the accurate-representational; 2) the epistemic; 3) the sociopolitical; 4) the moral; 5) the emotionally-and-cognitively-stimulating. This list is not exhaustive, but it is sufficient for clarifying the distinct functions of concepts and their relationships. We have already discussed the merely representational function and the accurate-representational one. Let us examine the remaining four.

For "epistemic function," I mean the function that involves cognitive processes like abstraction (see, e.g., Cartwright 1989, chapter 5), idealization (see, e.g., McMullin 1985), and in general, the result of the modification of concepts that offer us a more simple and manageable representation of reality (they still play the mere representational function). The resulting concepts now have the epistemic functions of representing in more general, simple, and unifying ways. These entail some departure from the accurate-representational function, but the result is highly beneficial for beings with limited computational capacities like us.[4] Nonetheless, most epistemic functions overlap with the accurate-representational function. We can appreciate this in our scientific and philosophical concepts. But these functions need not overlap. We can have simple and general concepts that lack a referent in the world (at least in the actual one). These concepts could be describing an alternative reality with a set of wholly different laws of nature. Perhaps they refer to fictional entities and characters.

---

than demanding true sentences and accurate concepts. For a defense of the latter view, see Simion (2018).

[4] Distinguishing the epistemic function from the accurate-representational function has proved to be extremely difficult (Brigandt (2010) seems to conflate both. See Giere (2006) for a helpful perspective on these issues). But there is, intuitively, a worth distinction to make here.

We do not have to go very far to appreciate the potential lack of overlapping between the epistemic and the accurate-representational functions. Many scientific and philosophical concepts do not accurately represent the world because their epistemic function will be compromised if they do. A general and simple concept must idealize and abstract away many real worldly features. Consider the concepts FRICTIONLESS PLANE, POINT-PARTICLE, ISOLATED SYSTEMS, or MARKETS IN PERFECT EQUILIBRIUM. Yet, these concepts are very useful for understanding the world.

Concepts play sociopolitical functions too. Most of our concepts about what kinds of behavior, attitudes, and thoughts a person must have, shape sociopolitical reality. Having a concept representing a specific political structure reinforces a particular political behavior. The same is true of legal, social, and economic interactions. If a person believes that some juridical law is right, he will obey it because he recognizes it as true. When a person thinks that his concept of, for example, LAW is appropriate, then he will exhibit a certain kind of behavior that reflects his understanding of this concept towards other human beings.

The sociopolitical function can overlap with accurate-representational and epistemic ones (the diverse functions of concepts can overlap but need not). Instances of the concept DISEASE represent real sets of disorders in organic structures and faculties. These concepts do not merely represent the world but serve to shape it and shape the social and cultural relationships between individuals. DISEASE also plays the epistemic function of being an abstract summary of countless disease instances that allows theory construction and smooth communication.

The status of moral concepts is different. There is an intense debate about the accurate-representational or the mere-representational functions of concepts like GENEROUS. Cognitivists claim that moral concepts play an accurate-representational function, non-cognitivists deny it. One can argue that, although moral concepts do not play an accurate-representational function, they play a mere representational one because people bring certain images of properties to their mind when they think about generous persons. But it need not concern us whether non-cognitivists are right. The important point now is that it is possible to have a concept with a moral function without having an epistemic or accurate-representational function.

In the same way, a moral concept need not have any emotionally-and-cognitively-stimulating function as does the concept INFINITY. A concept can promote peace and respect without having to arouse curiosity, feelings of joy, sadness, surprise, or foster the incubation of new ideas (of course, this concept can

play all these emotionally-and-cognitively-stimulant functions but need not). Furthermore, a concept can have some of these functions without trying to accurately describe reality (accurate-representational function) to be epistemically useful (epistemic function) or to play any social and moral function.

INFINITY plays a cognitively stimulating role, promoting in some individuals the construction of new philosophical theories that are relevant to assess the status of other concepts like CAUSALITY and NUMBER. SADNESS or ANXIETY can have an emotionally therapeutic function in people if we re-engineer them as concepts that refer to emotive states that do not necessarily arise in situations of adversity and pain (see section 4).

After this rough characterization of the functions of concepts, I want to point out some of the most important relations between them and concepts. They can relate at least in the following ways: a) a function lacks a concept; b) a concept lacks a function; c) one function has two or more concepts; d) one concept has two or more functions, and e) a concept has only one function.

In the first relation, a function can lack a concept. In that case, we can devise new concepts that fulfill that function. If we need a concept that can describe a feature of reality, we can design it to provide us with a better description of the latter. Perhaps we would not require a concept that accurately represents reality but one that plays an epistemic role and can help us handle the large amount of information that reaches our senses. Concepts like FRICTIONLESS PLANE, POINT-PARTICLE, and ISOLATED SYSTEMS are doing this job. They abstract from the specific characteristics of frictionless planes, point-particles, and isolated systems. Functions like these offer us a perfect opportunity to engineer new concepts from scratch.

Some concepts lack a function (b). This situation may mean the inexistence of any function whatsoever or the lack of a second one. But there are no concepts without a function because they (as I stated above) at least merely represent (it is hard to think of a concept that lacks any representational property whatsoever). Concepts can lack a second function when, for instance, a concept has an accurate-representational function but lacks a sociopolitical, moral, or epistemic one.

A function can have two or more concepts (c), therefore producing a redundancy. It may not be problematic, but in some cases, we may also want this functions to be simpler and then it is better to eliminate one or more of the concepts. It can also happen that a concept can have two or more functions (d). DISEASE has the function of accurately describing physiological reality, and at the same time, it may have a social role.

Finally, although (e) is very controversial, there may be concepts that only have one function: They, accurately or not, merely represent. These concepts elicit some representation in our minds. Even those concepts that depict things that do not exist have a representational function.[5]

The distinctions just made are of great importance to avoid confusion, inconsistencies, and other difficulties when we engineer concepts. Making distinctions can be very useful to design concepts purposefully. Very likely, concepts designed with a clear purpose can more easily fulfill it than those designed to achieve vague and tangled goals. In the same way, a clear distinction between functions can pave the way to understand and revise existing concepts. We will cover this topic in the next section.

## 4. The Importance of Differentiating the Diverse Functions of Concepts: Emotion Concepts, a Case Study

In this section, I introduce some of the difficulties that arise in the absence of a clear distinction between the different functions of concepts and why it is crucial that we differentiate them. I argue that clarifying and differentiating the conceptual functions can help us understand and implement them when doing conceptual engineering. Let us begin with an analysis of some of the risks that arise when we do not differentiate between functions.

Lisa F. Barrett (2016; 2017) argues that emotion concepts like ANGER and FEAR are not natural kinds but constructions of mind and culture. Her revolutionary "theory of constructed emotion" challenges the classic theory of emotion that posits a set of basic universal emotive states: anger, fear, sadness, disgust, surprise, and happiness. She denies any biological and psychological basis that determines emotions. Emotions are real as other human constructions like money but are not determined by discrete sets of firing neurons nor by specific changes in the autonomic nerve system, groups of facial expressions, or body movements. On the contrary, the social constructivist hypothesis says that we construct them on the spot and "love (or curiosity, hunger, etc.) is an emotion as long as people agree that its instances serve the functions of an emotion" (Barrett 2017, 138).

According to the theory of constructed emotion, there are not emotive natural kinds but internal sensations that we individuate as instances of specific

[5] Indeed, concepts like SQUARED-CIRCLE probably also have a representational function. A case can be made that they arouse curiosity, creativity, and distinctive emotional states. I am not going to settle the issue of the existence of concepts with only one function here. But if there is any, this seems to be a merely representational one.

emotions. In a process called "interoception," the interoceptive network (composed of body-budgeting regions and the primary interoceptive cortex) represents all the sensations from our internal organs, tissues, hormones, immune system, etc. The name for this set of sensations is "affect" and is composed of two parameters: valence and arousal. Valence is the quantity of pleasure or displeasure. Arousal is the low or high level of arousal. With allostasis in view (the regulation of our body for growth, survival, and reproduction), the visceromotor regions of the interoceptive network inform us about the diverse combinations of high and low arousal, pleasure, and displeasure. But this information does not strike us as pure internal sensation. Concepts mediate it.

The body-budgeting part of the interoceptive network employs our conceptually organized past experiences about situations, events, persons, angles, places, times, and feelings involved in anger or sadness episodes and makes multiple predictions (applies and constructs concepts) about what is happening in the world in a given situation before any sensory input reaches our senses. Outside conscious awareness, from the conceptually organized past experiences and the clues in the immediate environment, a set of firing neurons unpacks our conceptually organized experiences about, say, anger into multiple fine-grained concepts about angles, places, times, and feelings. Then the neurons of the interoceptive cortex test whether these concepts fit the actual affective sensory input. Part of the fine-grained concepts is discarded because they do not conform with anger sensations of high arousal and displeasure, in which case the brain constructs new predictions, and the process begins again. But the "confirmed" set, as it usually happens, is assembled into a concept that represents the emotion of anger (indeed, representation is essential for Barrett's theory. Concepts play the accurate-representational function regarding events and objects because concepts are "populations of representations that correspond to those events or objects" (Barrett 2016, 10)). The fine-grained concepts that compose the set are assembled into the concept ANGER for this specific situation, and we experience anger.

This explanation suggests that "emotional experiences have no objective fingerprints in the face, body, or brain that would enable us to compute an answer" (Barrett 2017, 107). Barrett has conducted many experiments that support the theory, and many other studies challenge the classical view of emotions as universal traits of human nature and stress the variability across cultures. In one study, participants had to identify emotions from three sources: reading about scenarios that usually enact a particular emotion, seeing photos of facial expressions, or both. She found that those who only read the scenario or read it and saw the face correctly predicted the emotion sixty-six percent of the time.

Those that only saw the facial expression got it right thirty-eight percent of the time. Tassinary et al. (1992) used a technique called electromyography (EMG) and did not find that participants displayed the same facial expression when they felt the same emotion. The electrodes on the surface of the skin detected the electrical signals that make the muscles responsible for the facial expressions, but people did not move the same facial muscles in the same pattern when experiencing a given emotion. There is evidence of "degeneracy" or, in philosophical terms, "indetermination" of emotive states by affective states (Wilson-Mendenhall et al. 2011; 2013) or populations of neurons (Whitacre & Bender 2010). Human emotions are compatible with different affective states. Besides, some metanalyses show that numerous studies have not found the "fingerprint" of emotion in changes in the autonomic nervous system, specific sets of neurons, facial expressions, or body movements (Lindquist et al. 2012). Other studies support the hypothesis that emotion is constrained by culture (Barrett, Mesquita, & Gendron 2011; Dixon 2008; Frijda & Mesquita 1995; Harré 1986; Mesquita & Frijda 1992; Russell 1991; Williams 1977).[6]

If the theory of constructed emotion is right, then we are the architects of our own emotional experiences. As individuals and as societies, we have the power to create our emotions because we decide what configurations of persons and situations bring about particular emotive states. If we, as individuals, look at persons and situations from different sociopolitical, moral, and emotionally-and-cognitively-stimulating perspectives, we can change our emotions. We can change them even more if we further change how our society looks at persons and situations. We can transform or construct, that is, we can re-engineer existing concepts or engineer new ones for specific purposes like treating arachnophobia and anxiety, fostering positive emotions like happiness and positive habits like doing more exercise, or using anger to perform better at sports events.

The constructivist theory fits nicely with the conceptual engineering program because emotion concepts are not restricted by their biological and psychological foundations: "Emotion concepts are goal-based concepts" (Barrett 2017, 92). There is not a biological glue that holds together sensations; our individual and cultural goals do it. Depending on the goal, we can construct a feeling of displeasure and high arousal as fear, anger, or intense sadness. We can create the sensations of pleasure and high arousal as anger or excitement. One can also engineer new emotion concepts by combining existing ones.

[6] Japanese use the word "itoshii" for a feeling of longing for an absent loved one and Bengali use "obhiman" for sorrow caused by the insensibility of a loved one (Russell 1991, 426), for example.

According to Barrett (2017), we can stop categorizing pain as exhaustion and then exercise more or deconstruct the concept ANXIETY and re-engineer it as the concept EXCITEMENT to treat anxiety. She argues that "people who recategorize anxiety as excitement show similar effects with better performance and fewer classic symptoms of anxiety when speaking in public and even when singing karaoke" (2017, 189). Categorization allows an affective sensation "to become an emotional experience such as happiness or fear, giving it additional meaning and functions understood within your culture. Categorization bestows new functions on biological signals" (2017, 126). With practice, Barrett says,

> You can dissolve anxiety into a fast-beating heart. Once you can deconstruct into physical sensations, then you can recategorize them in some other way, using your rich set of concepts. Perhaps that pounding in your chest is not anxiety but anticipation, or even excitement. (2017, 188)

The theory implies that we can engineer and re-engineer emotion concepts from our basic affective states. Engineering "bestows new functions on biological signals, not by virtue of their physical nature but by virtue of your knowledge and the context around you in the world" (Barrett 2017, 126). But can we really engineer emotion concepts at will? This theory faces some difficulties.

First, the theory of constructed emotion is not uncontroversially confirmed by the data. In the same way that many studies support it, there are also multiple studies and metanalyses confirming the existence of emotion fingertips in the autonomic nervous system, facial expressions, and neural patterns (Elfenbein & Ambady 2002; Norenzayan & Heine 2005; Phan et al. 2002; Vytal & Hamman 2010)

Second, if emotions are constructed at will, then the logical consequence is that it is highly probable that there are human communities that cannot communicate their emotive states because there is no need for two dissimilar cultures to share any emotive concept. But there is no evidence of any culture whose emotion concepts are not partially shared with the researcher or with other cultures. Translation would be impossible in such a relativistic framework, but we translate emotion words from different cultures. Communication is not total and translation is not exact, but there must be something to emotions that explains why intercultural communication and translation exceed chance. Barrett (2017, 38) suggests that "emotions are not inborn, and if they are universal, it's due to shared concepts," but the theory of constructed emotion cannot explain why we share them. On the other hand, the classical theory that posits a universal emotional biological basis across cultures can easily explain why we tend to share the same categories.

Finally, the theory of constructed emotions posits the existence of four fundamental affective states or dimensions that are not exclusive of humans but occur in all species: pleasure, displeasure, high and low arousal (Barrett 2016, 17).[7] Indeed, arousal and valence are practically impossible to separate because if one changes, the other also changes (Barrett & Bliss-Moreau 2009) and scientists have found that valence is common to all cultures (Farroni 2007; Wierzbicka 1992). We can accept the idea that these elementary states do not determine emotion concepts, but it is compatible with affective states *partially* determining them.[8] It is compatible with the fact that we share a set of *structural ranges* that underlie the possibility of translations and "they occur over a very wide range of cultures, and probably is universal, even if no words exist for certain of the structures in a particular language" (Frijda et al. 1995, 124).

"Negative" emotions like anger, fear, and sadness are unpleasing, and "positive" emotions like happiness and love are pleasing. Unless we allow certain twists of meaning in our usual understanding of the basic emotions, we cannot say that fear is an emotion concept that can be constructed from pleasing and low arousal affective states. Anger cannot be an emotion concept that arises from low arousal and pleasing states, and sadness is not a category that rests on a high arousal and pleasing affective basis because they share the displeasing valence.[9] In other words, the concept FEAR tracks an affective feeling that is displeasing and provokes high arousal. The properties of displeasure and high arousal do not determine the concept FEAR, and there is room for re-engineering it. But the partial determination partially restricts the engineering activity. On pain of inconsistency, confusion, and possible verbal disputes, the accurate-representational function of FEAR represents, though imperfectly, a region in the affective state that is characterized by high arousal and displeasure. It does not matter that we do not perfectly identify the corresponding area. Perhaps no region in nature has precise boundaries, but it does nothing to deter us from identifying gold, quarks, and atoms.

It is true that the neurological and physiological facts by themselves cannot determine the emotional facts. Emotive states are indetermined by the most,

[7] Barrett et al. (2011) even describe how a patient with semantic dementia categorizes emotions as "positive" and "negative."

[8] Griffiths (2002) advances a similar view. According to Griffiths, emotion words "partially refer," for example, to affect programs and to socially sustained practices.

[9] As the Ilongot word "liget" demonstrates in Russell's paper (1991), there can be overlapping between, for example, anger and sadness because they share a displeasing valence, but that does not mean that they are the same.

ontologically speaking, basic neurological and physiological ones, but this is perfectly compatible with partial determination. Emotion concepts fulfill their accurate-representational function when partially referring to specific affective regions even when they are blurry because of the varying set of individual and cultural practices.[10]

The accurate representational function does not depend on the individual or on culture. The world determines it. Yet, we can engineer what is not determined by the world to achieve other functions besides the accurate-representational. As long as these functions do not conflict with each other, all is fine. Besides playing the accurate-representational function of a region of high arousal and low pleasure in the affective state, redefining FEAR so that certain situations like speaking in public and meeting new people are culturally understood as non-threatening events plays the social function of diminishing individuals' fear states. A given concept can unproblematically play the accurate-representational and social, recreative, or moral functions at the same time. But is it always like that? I do not think so.

Concepts like CHAIR, GOLD, and ANIMAL purport to refer, primarily, to some features of reality: chairs, gold, and animals. This function is independent of its potential sociopolitical, aesthetical, and recreational uses. We can use GOLD to promote a specific social behavior. If, for some strange reason, we want people to behave towards chairs and animals as they do with gold, we can re-engineer the concept of GOLD to refer not only to gold but to chairs and animals. But we would have completely destroyed its accurate-representational function of representing objects that are yellow, shining, and with atomic number 79. In the same way, if some robots acquire enough external and behavioral similarity with animals, we may also want to preserve the meaning of ANIMAL as referring to organisms composed of cells and re-engineer (expand its extension) it to refer to robots. But then we would end up with a defective inconsistent concept because its accurate-representational function still includes referring to organisms composed of cells but now also includes robots in its extension. The deviation from the original concepts would be so great that the concept ANIMAL would be a new concept competing with its previous version, and it will lose. If we want to re-engineer or engineer

[10] It is true that Barrett (2012) acknowledges the contribution of biology to emotion construction. The problem is that a concession like that means that we cannot construct concepts at will and on the spot. The theory of constructed emotion must be amended to accommodate the partial determination of the biological substrate.

concepts to fulfill other functions, we must differentiate the diverse roles that they can consistently play.[11]

Thus, we can distinguish two functions in Barrett´s ameliorated concept of ANXIETY as excitement. One function is the accurate-representational of referring to a high arousal and displeasing affective state, and the other is a therapeutic and performance improving one (that are instances of the emotionally-and-cognitively-stimulating function). If this re-engineered concept is successful, individuals and society alike will improve their mental health and performance in various ways. But ANXIETY faces the same problems that GOLD and ANIMAL.

The re-engineering of ANXIETY as excitement conflates two distinct but conflicting functions. If ANXIETY means that one feels excitement, then one experiences an affective state of high arousal and pleasure, but the accurate-representational function of ANXIETY depends on referring to an affective state of high arousal and displeasure. They are incompatible.[12] Perhaps ANXIETY refers to a sensation of pleasure and displeasure at the same time, but this new concept also conflicts with our original English concepts of ANXIETY and EXCITEMENT that pick up unpleasing and pleasing states, correspondingly. And unless we defend the existence of inconsistent concepts, we must fix the conflict.[13] We can hold back our re-engineering pretenses, add a new compatible therapeutic function to ANXIETY, or engineer a different concept.

The lack of attention to the functions we assign to concepts engenders inconsistencies and confusion. If a patient uses EXCITEMENT as an instance of the feeling of anxiety can confuse psychologists and vice versa. If a group of psychologists that adhere to the theory of constructed emotion employ ANXIETY as a concept that picks up pleasing affective states, it will engender a verbal dispute with psychologists fond of the classical theory. One of the main sources of these difficulties is the over-optimistic idea that we can construe emotions and that "we can impose functions that would not otherwise exist, thereby inventing reality" (Barrett 2017, 135). This idea obscures the accurate-representational function that many concepts possess and how it can conflict with other "imposed functions." If ANXIETY accurately describes a feature of reality, the high arousal and displeasing

[11] Wishful thinking, for example, can promote better health; "nevertheless, practical reasons are not good reasons for belief: wishful thinking is bad believing" (Simion 2018, 95).

[12] We can also think of future incompatibility. Whether a compatible relationship between two or more functions continues or not is an empirical question. If, for example, future scientific research reveals that the accurate-representational function of a concept conflicts with others, then we would need to fix the problem.

[13] We do conceptual revision with amelioration in view, not inconsistency or "perversion" (Marques 2020).

affective state, then we cannot invent new anxiety concepts but only modify them to the extent that the slack between the world and concepts allows. Understanding this can aid us to design emotion concepts and concepts overall with clearer functions, and then more easily fulfill goals like treating anxiety, avoiding negative emotions, increasing positive ones, and using them to improve performance in sports and life. We can construe our concepts, but we lack complete engineering control over them and what we want them to be.

In summary, when we do conceptual engineering, it will pay off to differentiate the distinct functions of our concepts as far as possible. This distinction will be useful to improve the understanding of our current concepts, foresee the possible ways to ameliorate them, avoid inconsistencies, confusion, and possible verbal disputes.

## 6. Conclusion

We have seen that functional-conceptual-engineering is a good candidate for overcoming the discontinuity problem. We have also noted that this way to understand conceptual engineering must face other difficulties. I pointed out that one way to avoid these obstacles is to differentiate between the different kinds of functions of concepts, as well as the multiple relationships between them.

We can engineer or re-engineer most of our concepts, but there are limits to this activity. If we do not examine the functions that we are adding to our concepts, we can end up with conceptual and functional inconsistency. We must be aware that adding sociopolitical, moral, or therapeutic functions to concepts that fulfill an accurate-representational or epistemic function, or vice versa, may be detrimental. One way to improve the activity of conceptual engineering is to differentiate between the different functions that our concepts can play.

One point that has been overlooked is that, although there is some slack between the world and our concepts, we must acknowledge that the world partially determines concepts – one exception is perhaps fictional concepts. We also need to be aware that our assessment of this relation between the world and concepts can change with the advancement of science. Our evaluation of the independence of sociopolitical, epistemic, entertaining, and therapeutic practices from their neurological, physiological and psychological substrate will change with our scientific knowledge. Surely, there is necessary much more work in clarifying and differentiating the numerous functions played by our current concepts – and

perhaps future ones. But, I hope, this rough exploration can shed some light on the practice and improvement of conceptual engineering.[14]

## References

Bach, Theodore. 2016. "Social Categories Are Natural Kinds, Not Objective Types (and Why It Matters Politically)." *Journal of Social Ontology* 2 (2): 177-201. https://doi.org/10.1515/jso-2015-0039.

Bach, Theodore. 2019. "Real Kinds in Real Time: On Responsible Social Modeling." *The Monist* 102 (2): 236-58. https://doi.org/10.1093/MONIST/ONZ008.

Barrett, F. Lisa, and Eliza, Bliss-Moreau. 2009. "Affect as a Psychological Primitive." *Advances in Experimental Social Psychology* 41, 167-218. https://doi.org/ 10.1016/S0065-2601(08)00404-8.

Barrett, F. Lisa, Batja Mesquita, and Maria Gendron. 2011. "Context in Emotion Perception." *Current Directions in Psychological Science* 20 (5): 286-90. https://doi.org/10.1177/0963721411422522.

Barrett, F. Lisa. 2012. "Emotions Are Real." *American Psychological Association* 12 (3): 413-29. https://doi.org/10.1037/a0027555.

Barrett, F. Lisa. 2016. "The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization." *Social Cognitive and Affective Neuroscience* 12 (1): 1-23. https://doi.org/10.1093/scan/nsw154.

Barrett, F. Lisa. 2017. *How Emotions Are Made: The Secret Life of the Brain.* Boston: Houghton Mifflin Harcourt.

Brigandt, Ingo. 2010. "The Epistemic Goal of a Concept: Accounting for the Rationality of Semantic Change and Variation." *Synthese* 177 (1): 19-40. https://doi.org/10.1007/s11229-009-9623-8.

Burgess, Alexis, and David Plunkett. 2013a. "Conceptual Ethics I." *Philosophy Compass* 8 (12): 1091-101. https://doi.org/10.1111/phc3.12086.

Burgess, Alexis, and David Plunkett. 2013b. "Conceptual ethics II." *Philosophy Compass* 8 (12): 1102-10. https://doi.org/10.1111/phc3.12085.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering.* Oxford: Oxford University Press.

—. 2020. "Conceptual Engineering: The Master Argument." In *Conceptual Engineering and Conceptual Ethics*, edited by Herman Cappelen, David Plunkett, and Alexis Burgess, 1-34. Oxford: Oxford University Press.

Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement.* Oxford: Oxford University Press.

[14] **Acknowledgments:** I thank Herman Cappelen for his helpful suggestions.

Chihara, Charles. 1979. "The Semantic Paradoxes: A Diagnostic Investigation." *The Philosophical Review* 88 (4): 590-618.

Dixon, Thomas. 2008. *The Invention of Altruism: Making Moral Meanings in Victorian Britain*. Oxford: Oxford University Press.

Eklund, Matti. 2002. "Inconsistent Languages." *Philosophy and Phenomenological Research* 64 (2): 251-75. https://doi.org/10.1111/j.1933-1592.2002.tb00001.x

Eklund, Matti. 2014. "Replacing Truth." In *Metasemantics: New Essays on the Foundations of Meaning*, edited by Alexis Burgess and Brett Sherman, 293-310. Oxford: Oxford University Press.

Elfenbein, A. Hilary, and Nalini Ambady. 2002. "On the Universality and Cultural Specificity of Emotion Recognition: A Meta-analysis." *Psychological Bulletin* 128 (2): 203-35. https://doi.org/10.1037/0033-2909.128.2.203.

Farroni, Teresa, Enrica Menon, Silvia Rigato, and Mark H. Johnson. 2007. "The Perception of Facial Expressions in Newborns." *European Journal of Development Psychology* 4 (1): 2-13. https://doi.org/10.1080/17405620601046832.

Floridi, Luciano. 2011. "A Defence of Constructionism." *Metaphilosophy* 42: 282-304. https://doi.org/10.1111/j.1467-9973.2011.01693.x.

Frijda, H. Nico, Suprapti S. Markam, Kaori Sato, and Reinout Wiers. 1995. "Emotions and Emotion Words." In *Everyday Conceptions of Emotion: An Introduction to Psychology, Anthropology and Linguistics of Emotion*, edited by James A. Russell, José M. Fernández-Dols, Anthony S. R. Manstead, and Jane C. Wellenkamp, 121-43. Madrid: Springer.

Frijda, H. Nico, and Batja Mesquita. 1995. "The Social Roles and Functions of Emotions." In *Emotion and Culture: Empirical Studies of Mutual Influence*, edited by Shinobu Kitayama, and Rose R. Hazel, 51-87. American Psychological Association.

Giere, N. Ronald. 2006. *Scientific Perspectivism.* Chicago: The University of Chicago Press.

Goldstein, Jacob. 2020. *Money: The True Story of a Made-Up Thing*. New York: Hachette Books.

Griffiths, E. Paul. 2002. "Emotions as Natural and Normative Kinds." *Philosophy of Science* 71 (5): 901-11. https://doi.org/10.1086/425944.

Hacking, Ian. 1995. "The Looping Effects of Human Kinds." In *Causal Cognition: A Multidisciplinary Debate*, edited by Dan Sperber, David Premack, and James Ann Premack, 351-94. Oxford: Clarendon Press.

—. 1999. *The Social Construction of What?* Harvard: Harvard Univesity Press.

Harré, Rom. 1986. *The Social Construction of Emotion.* Oxford: Blackwell.

Isaac Gustavo, Manuel. 2020. "How to Conceptually Engineer Conceptual Engineering?" *Inquiry: An Interdisciplinary Journal of Philosophy.* https://doi.org/10.1080/0020174X.2020.1719881.

Koch, Steffen. 2021. "The Externalist Challenge to Conceptual Engineering." *Synthese* 198 (1): 327-48. http://doi.org/10.1007/s11229-018-02007-6.

Lindquist, A. Kristen, Tor D. Wager, Hedy Kober, Eliza Bliss-Moreau, and Lisa F. Barrett. 2012. "The Brain of Emotion: A Meta-analytic Review." *Behavioural and Brain Sciences* 35 (3): 121-202. https://doi.org/10.1017/S0140525X11000446.

Machery, Edouard. 2009. *Doing Without Concepts.* Oxford: Oxford University Press.

Margolis, Eric, and Stephen Laurence. 2014. *"Concepts" The Stanford Encyclopedia of Philosophy.* Edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2014/entries/concepts/.

Marques, Teresa. 2020. "Amelioration vs Perversion." In *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*, edited by Teresa Marques and Åsa Wikforss, 260-84. Oxford: Oxford University Press.

Mesquita, Batja, and Nico H. Frijda. 1992. "Cultural Variations in Emotions: A Review." *Psychological Bulletin* 112 (2): 179-204. https://doi.org/10.1037/0033-2909.112.2.179.

McMullin, Ernan. 1985. "Galilean Idealization." *Studies in the History and Philosophy of Science* 16: 247-73. https://doi.org/10.1016/0039-3681(85)90003-2.

Nado, Jennifer. 2019. "Conceptual Engineering, Truth and Efficacy." *Synthese* 198: 1507-27. https://doi.org/10.1007/s11229-019-02096-x.

Norenzayan, Ara, and Steve J. Heine. 2005. "Psychological Universals: What Are They and How Can We Know?" *Psychological Bulletin* 131 (5): 763-84. https://doi.org/10.1037/0033-2909.131.5.763.

Phan, K. Luan, Tor Wager, Taylor F. Stephan, and Israel Liberzon. 2002. "Functional Neuroanatomy of Emotion: A Meta-analysis of Emotion Activation Studies in PET and fMRI." *Neuroimage* 16 (2): 331-48. http://doi.org/10.1006/ nimg.2002.1087.

Plunkett, David, and Herman Cappelen. 2020. "A Guided Tour of Conceptual Analysis and Conceptual Ethics." In *Conceptual Engineering and Conceptual Ethics*, edited by Herman Cappelen, David Plunkett, and Alexis Burgess, 1-34. Oxford: Oxford University Press.

Prinzing, Michael. 2017. "The Revisionist Rubric: Conceptual Engineering and the Discontinuity Objection." *Inquiry: An Interdisciplinary Journal of Philosophy* 61 (8): 854-80. https://doi.org/10.1080/0020174X.2017.1385522.

Riggs, Jared. 2020. "Conceptual Engineers Shouldn't Worry about Semantic Externalism." *Inquiry.* https://doi.org/10.1080/0020174X.2019.1675534.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature.* Princeton: Princeton University Press.

Rorty, Richard. 1990. *Objectivity, Relativism, and Truth*. Cambridge. Cambridge University Press.

Russell, A. James. 1991. "Culture and the Categorization of Emotions." *Psychological Bulletin* 110 (3): 426-50.

Sawyer, Sarah. 2020a. "The Role of Concepts in Fixing Language." *Canadian Journal of Philosophy* 50 (5): 555-65. https://doi.org/10.1017/can.2020.5.

Sawyer, Sarah. 2020b. "Truth and Objectivity in Conceptual Engineering." *Inquiry* 69 (9-10): 1001-22. https://doi.org/10.1080/0020174X.2020.1805708.

Scharp, Kevin. 2013. "Truth, the Liar, and Relativism." *Philosophical Review* 122 (3): 427-510.

Scharp, Kevin. 2014. "Replacing Truth." *The Philosophical Quarterly* 64 (256): 535-37. https://doi.org/10.1093/pq/pqu019.

Simion, Mona. 2018. "Epistemic Trouble for Engineering 'Woman'." *Logos & Episteme* 9 (1): 91-8. https://doi.org/10.5840/logos-episteme2018916.

Simion, Mona, and Christopher Kelp. 2019. "Conceptual Innovation, Function First." *Noûs* 54 (4): 1-18. doi:10.1111/nous.12302.

Strawson, F. Peter. 1963. "Carnap's Views on Constructed Systems Versus Natural Languages in Analytic Philosophy." In *The philosophy of Rudolf Carnap*, edited by Paul A. Schilpp, 503-18. La Salle: Open Court.

Tassinary, G. Louis, and John T. Cacioppo. 1992. "Unobservable Facial Actions and Emotion." *Psychological Science* 3 (1): 28-33. https://doi.org/10.1111/j.1467-9280.1992.tb00252.x.

Thomasson, Amie. 2020. "A Pragmatic Method for Conceptual Ethics." In *Conceptual Ethics and Conceptual Engineering*, edited by Herman Cappelen, David Plunkett, and Alexis Burgess, 435-58. Oxford: Oxford University Press.

Vytal, Katherine, and Stephan Hamman. 2010. Neuroimaging Support for Discrete Neural Correlates of Basic Emotions: A Voxel-based Meta-analysis. *Journal of Cognitive Neuroscience* 22 (12): 2864-85. https://doi.org/10.1162/jocn.20b09.21366.

Whitacre, James, and Axel Bender. 2010. "Degeneracy: A Design Principle for Achieving Robustness and Evolvability." *Journal of Theoretical Biology* 263 (1): 143-53. https://doi.org/10.1016/j.jtbi.2009.11.008.

Wierzbicka, Anna. 1992. *Semantics, Culture, and Cognition: Universal Human Concepts in Culture-Specific Configurations*. Oxford: Oxford University Press.

Williams, Raymond. 1977. "Keywords: A Vocabulary of Culture and Society." *Science and Society* 41 (2): 221-24.

Wilson-Mendenhall, D. Christine, Lisa F. Barrett, Kyle W. Simmons, and Lawrence W. Barsalou. 2011. "Grounding Emotion in Situated Conceptualization." *Neuropsychologia* 49 (5): 1105-127. https://doi.org/10.1016/j.neuropsychologia.2010.12.032.

Wilson-Mendenhall, D. Christine, Lisa F. Barrett, and Lawrence W. Barsalou. 2013. Neural Evidence that Human Emotions Share Core Affective Properties. *Psychological Science* 24 (6): 947-56. https://doi.org/10.1177/0956797612464242.