Philosophical Psychology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cphp20

Emotions in conceptual spaces

Michał Sikorski & Ohan Hominis

To cite this article: Michał Sikorski & Ohan Hominis (24 Mar 2024): Emotions in conceptual spaces, Philosophical Psychology, DOI: 10.1080/09515089.2024.2330477

To link to this article: https://doi.org/10.1080/09515089.2024.2330477

Published online: 24 Mar 2024.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=cphp20

PHILOSOPHICAL PSYCHOLOGY
https://doi.org/10.1080/09515089.2024.2330477

Emotions in conceptual spaces

Michał Sikorski [a] and Ohan Hominis [b]

[a]Center for Philosophy, Science, and Policy, Department of Biomedical Sciences and Public Health, Marche Polytechnic University, Ancona, Italy; [b]Department of Cognitive Science, Central European University, Vienna, Austria

ABSTRACT
The overreliance on verbal models and theories in psychology has been criticized for hindering the development of reliable research programs (Harris, 1976; Yarkoni, 2020). We demonstrate how the conceptual space framework can be used to formalize verbal theories and improve their precision and testability. In the framework, scientific concepts are represented by means of geometric objects. As a case study, we present a formalization of an existing three-dimensional theory of emotion which was developed with a spatial metaphor in mind. Wundt posits that the range of human emotion can be represented along three axes of basic emotions (pleasure/displeasure, excitement/inhibition, tension/relaxation), just as color vision can be represented using three basic colors. We use dimensions to represent basic emotions, points to represent emotional states, and regions to represent broader emotional concepts. We then compare our formalization to an existing structuralist formalization of Wundt's theory. Further, we discuss the empirical predictions that our formalization generates, such as comparisons of similarity and intensity. We conclude by demonstrating how the tools developed in the conceptual space framework can be used to formulate a theory of emotion based on empirical observation.

ARTICLE HISTORY
Received 29 June 2023
Accepted 29 February 2024

KEYWORDS
Conceptual spaces framework; formalization; dimensional theory of emotions

1. Introduction

The aim of the paper is to argue that the conceptual space framework can be used to formalize psychological theories. We will understand formalization throughout the paper as the process of translating given sentences from a natural language to a more precise formal language.

We contend that theories conceptualized in the framework are more precise and therefore easier to test than verbal theories. In this section, we discuss the current state of psychology and introduce the conceptual space framework.

CONTACT Michał Sikorski michalpsikorski@gmail.com Center for Philosophy, Science, and Policy, Department of Biomedical Sciences and Public Health, Marche Polytechnic University, Via Lodovico Menicucci, 6, Ancona, Italy

© 2024 Informa UK Limited, trading as Taylor & Francis Group

1.1. The state of psychological results

Since the beginning of the twenty-first century, the way scientists, and especially psychologists, view the reliability of their results has changed dramatically. In the words of Pashler and Wagenmakers (2012):

> "Is there currently a crisis of confidence in psychological science reflecting an unprecedented level of doubt among practitioners about the reliability of research findings in the field? It would certainly appear that there is."

This general skepticism (see also Baker, 2016) is a reaction to a methodological crisis that plagues psychology, often referred to as *The Replication Crisis*. The crisis lies in the fact that many psychological (or more generally, scientific) results are not replicable. Replicability is believed to be a defining feature of scientific practice (see e.g., National Academies of Sciences, E., & Medicine, 2019). Generally, scientific results are trusted because they are generated by procedures and practices that ought to produce similar results when repeated. In light of that, results of recent large-scale replication projects have cast doubts on the general replicability of psychological results (see e.g., Collaboration, 2015).

Not only is low replicability epistemically problematic (see e.g., National Academies of Sciences, & Medicine, 2019), but also potentially responsible for both a significant waste of resources (Freedman et al., 2015) and a decline in the confidence that society at large places in scientific findings (see e.g., Anvari & Lakens, 2018).

Growing awareness of the cited concerns has sparked a methodological debate among researchers. Among the commonly discussed factors contributing to questions of replicability are questionable research practices (QRP), problems with the incentive structure of science (such as publication bias), and an over-reliance on underpowered statistical inferences. Much of the focus has been on the QRP in particular, which are unreported, ad hoc modifications of experimental designs (e.g., Murphy & Aguinis, 2019). Evidence suggests that QRP are prevalent. They significantly reduce the reliability of psychological results (e.g., Fraser et al., 2018).

Various proposals to resolve the crisis have been put forward. Some of them restrict the use of QRP. For example, Nosek and Lakens (2014) proposed preregistration, i.e., submitting a description of the hypothesis under test and details of the experimental design before an experiment is performed. This disincentives the use of QRP; if the tested hypothesis is specified before the experiment, ad hoc modifications of hypotheses after data collection are no longer possible.

Much of the discussion concerning the causes of the replicability crisis focuses on the later stages of the scientific process like testing the hypothesis or publishing results. Such criticisms are valid and addressing the identified shortcomings is beneficial. However, there is a persistent gap in the literature. The shortcomings of psychological theories and possible improvements in how such theories are constructed are rarely discussed. Recently, a number of researchers claim that the early stages of the scientific process, like theory construction, are crucial for the reliability of scientific results and therefore have to be addressed within the attempts to ameliorate the replicability crisis (see e.g., Bird, 2020; Smaldino, 2019; Swedberg, 2012; van Rooij & Baggio, 2021; Witte & Zenker, 2017).

Here we intend to take a step toward closing this gap by addressing one of the well-documented deficiencies of psychological hypotheses and theories, namely their tendencies to be informal and ambiguous (Harris, 1976; Sennet, 2021). Psychological hypotheses are typically presented as claims in natural language (e.g., "Prepayment led to more positive product-related emotions than repayment." from Hahn et al., 2013). Due to its ambiguity and imprecision natural language is not a perfect tool to express a precise relation between fine-grained objects (see e.g., Sennet, 2021). These deficiencies are carried forward to hypotheses and theories formulated in natural languages, causing a number of long-standing problems (see e.g., Harris, 1976) that have been recently reasserted (see Yarkoni, 2020 or Linden & Hönekopp, 2021). For example, Linden and Hönekopp (2021) shows that the ambiguity of verbal concepts may be one of the causes of the vast heterogeneity of psychological results.

Such concerns were also raised in the psychology of emotions (see e.g., Weidman et al., 2017 or Fiske, 2020). Fiske (2020) have argued that psychologists and philosophers focus too much on the concepts of emotions from natural language (most notably English). According to the author, in most cases, it is not clear if such verbal concepts correspond to any real psychological phenomena or even if they constitute natural concepts. The verbal concepts of the same emotions from different languages have different features which strongly suggests that not all of them correspond to real psychological phenomena (see e.g., Wierzbicka, 2013). Fiske advises that instead of relying on the received verbal concepts researchers should construct new concepts of emotions suitable to be used in science. The formal character is one of the features making the concept suitable to be used in science (see e.g., Guest & Martin, 2020; Oberauer & Lewandowsky, 2019 or Haslbeck et al., 2021). In the following sections, we will try to demonstrate that conceptual space framework can be used in the construction of formal, precise concepts suitable to be used in science and we will use emotions as an example.

1.2. Formalization and conceptual space framework

Formalization offers a means of addressing some of the problems of verbal theories.[1] It amounts to translating a verbal theory from a natural language to a more precise, formal one. The most popular strategy of formalization, called *structuralism*, utilizes a language of set theory. Set theory is a branch of mathematical logic that studies objects called *sets*. Sets are axiomatically defined collections of elements. All mathematical objects, such as numbers or functions, can be reconstructed in terms of sets (see Bagaria, 2020). The precision and expressiveness of set theory make it an ideal candidate for a framework for formalizing scientific theories. Such formalization typically involves representing elements of the theory, such as models or intended applications, in terms of sets (see Andreas & Zenker, 2013 for an introduction). The strategy has been most successfully applied in natural sciences like physics (e.g., Balzer et al., 2000), but several psychological theories have also been formalized in this framework (see e.g., Westmeyer, 1989). Despite this, structuralism has yet to become a popular tool for psychologists who generally find it too demanding. Even proponents of the introduction of structuralism into psychology agree that this hurdle may be insurmountable.

In the second section, we discuss an existing example of a structuralist formalization of a psychological theory. We then present an alternative method of formalization, the conceptual space framework introduced by Gärdenfors (2000) (see Figure 1. for example), and how it can be implemented to formalize verbal psychological theories. We argue that conceptual spaces as a methodology tend to be less demanding and therefore more useful for psychologists.

To demonstrate the potential of the conceptual space framework, we offer a formalization of the theory of emotions presented in Wundt (1897). The theory was proposed with a spatial metaphor in mind and therefore is well suited for formalization in the conceptual space framework. Additionally, the theory is a good example of a wider class of dimensional theories of emotions. Finally, there is some evidence suggesting that the problems described above are present in the literature dedicated to emotions (see e.g., Weidman et al., 2017).

Section 2 provides a brief overview of Wundt's theory and briefly discusses its existing structuralist formalization (Reisenzein, 1992). In Section 3, we present our formalization in the conceptual space framework and discuss what is gained. In Section 4, we will argue for the broader usefulness of the conceptual space framework within psychology and discuss how the tools developed in the framework can be used to construct a new conceptual space of emotion on the basis of experimental results. Finally, Section 5 is devoted to general lessons concerning conceptualizing psychological concepts that can be drawn from our formalization and possible future work.

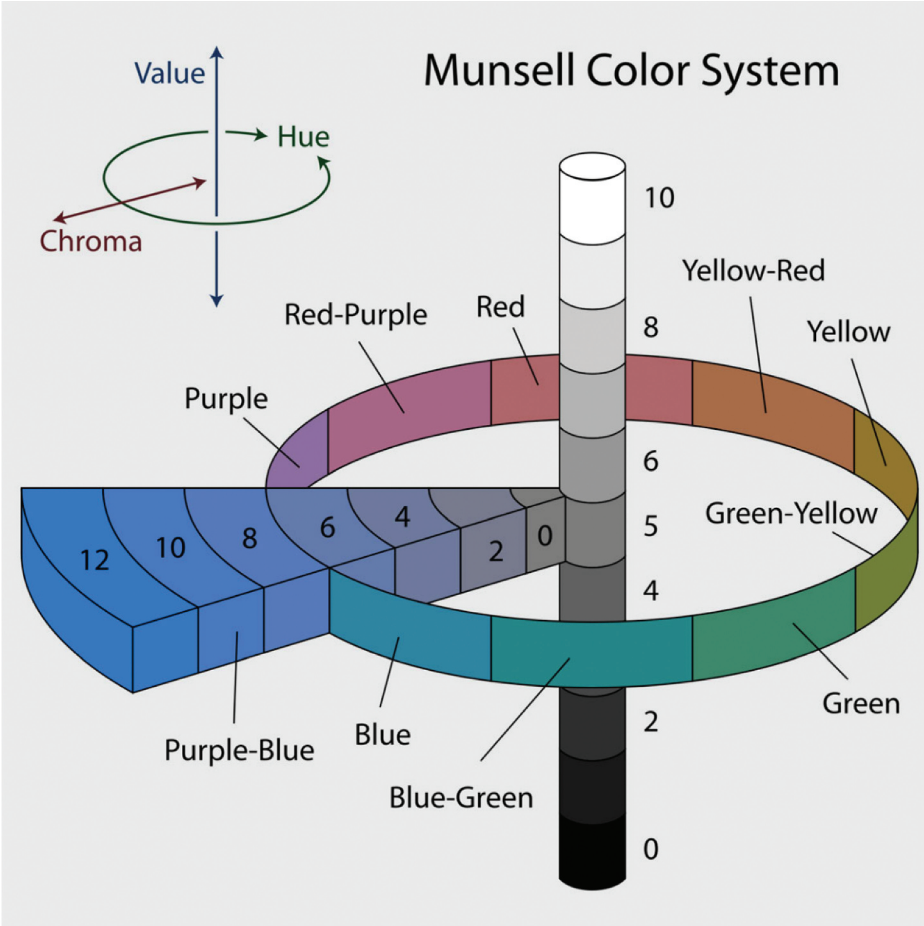

Figure 1. The conceptual space of colors (source of the picture is: Rus (2007)).

2. Wundt's theory of emotions

Wundt (1897) presents an early, quantitative theory of emotional experience that seeks to reduce the multitude of emotions to a set of basic mental states. This account of the underlying dynamics from which emotions emerge is built on a three-dimensional model of basic states, making it an ideal candidate for formalization in the conceptual space framework. Through a dimensional model, Wundt attempts to provide a causal, reductive account of emotion that is embedded in his broader view of psychology and mental phenomena. For practical purposes the theoretical superstructure which supports Wundt's account of emotions will be largely ignored as providing a viable theory of emotion is not an aim of this paper. Instead, we utilize the theory as a case study for the formalization of psychological theories within the conceptual spaces paradigm.

Following a brief presentation of the relevant elements of Wundt's theory of emotion, we provide a summary of the structuralist formalization of the theory offered by (Reisenzein, 1992). While Reisenzein's formalization offers a precise and robust account of Wundt's theory, it is complex demanding. Our aim is to demonstrate that the conceptual spaces paradigm improves both the accessibility and explanatory power of formalization without sacrificing accuracy.

2.1. Informal presentation of Wundt's theory

Wundt targets the underlying structure of emotions and the relationships that define their dynamics, by presenting emotions as emergent from the interactions between simpler states. While reductive strategies among theories of emotion are common, most propose a subset of emotions as ontological primaries from which others emerge (Izard, 1977; Tomkins, 1962). Wundt, however, focuses solely on the emergence of emotional states as psychological phenomena. Wundt's introduction of basic mental states, or what he refers to as *feelings*, is distinct in that they act as dimensions that serve to quantitatively describe emotions.

To Wundt, the psychological analysis provides insight into two kinds of *basic psychic elements* from which *complex conscious experiences* emerge: sensory elements, which produce ideas, and affective elements, which produce emotions. He further breaks down affective elements into three components: feelings, affects, and volitions. His theory does not impose strict compositionality of individual emotions and maintains that, through introspection, emotions can be reduced to jointly experienced basic states.

Wundt's theory hinges upon the dynamics of feelings that compose emotions. While Wundt maintains that the number of emotional states is innumerable, he posits that they can be described by three sets of bipolar measures which he refers to as quality dimensions: pleasure and displeasure, excitement and inhibition, and tension and relaxation.[2] While these quality dimensions (hereafter referred to as *feelings*) are used to explain the emergence of complex affective states, they serve to describe their phenomenological qualities rather than to define their composition. Unlike the emotions they help describe, feelings are irreducible components of the theory, and any number of them, even just one can be used to describe an emotion. Of course, whether the term "emotion" or the term "feeling" is used to describe the fundamental component is an issue of semantic consistency with Wundt rather than ontological assertion.[3]

Wundt introduces a detailed hierarchy of states from which a variety of emotions emerge. We will focus on feelings and how they are combined to generate emotions. He held that the feelings themselves were not rigidly defined and that in any perceivable amount of time, an individual can experience any number of feelings. Additionally, feelings must be defined in relation to a particular object. Any fusion of feelings must be defined by the predominant feelings in relation to a particular object. Throughout his descriptions of emotion, Wundt is ambiguous regarding the fusion of feelings and how more complex states emerge from them. His fluid use of terminology, lack of formal definitions, and generally rather poetic descriptions of psychology make it difficult to discern a clear idea of fusion (Reisenzein's interpretation of Wundt's fusion of feelings is described in further detail in the next section).

While Wundt explicitly defined his three categories of affective states (feelings, affects, volitions) as being distinct from emotions themselves, for the sake of simplicity we will reduce all references to his affective states to one category: feelings. Similarly, we will also consolidate what Wundt alternatively considers complex feelings, affects, and affective states into one category: emotions. In our formalization, similar to other dimensional theories, feelings are represented by the three pairs of bipolar dimensions from which emotions emerge.[4] This paper is not concerned with questions of the ontological status of emotions or whether the states from which emotions emerge are emotions themselves. While the formalization does give us a sense of how to account for common complex emotions like jealousy and guilt, the critical component of the theory with regard to conceptual spaces is its dimensionality. It allows for quantifiable predictions of qualitative phenomena and a versatile composition space for the construction of a taxonomy of emotions. Incidentally, Wundt himself drew analogies between his theory and the study of three-dimensional representations of the color space, effectively anticipating a formalization of his theory in the conceptual space framework.

2.2. Existing formalization of the theory

Reisenzein (1992) presents a structuralist reconstruction of Wundt's theory of emotion, leveraging its dimensionality to build a dynamic, set-theoretic formalization of the theory. Because Wundt's theory of emotion was published in the larger context of his general view of psychology as a field, Reisenzein simplified various aspects of the theory to ensure it was both compatible with structuralist framework and able to stand on its own as a theory outside of the context of Wundt's work. Notably, he circumvents the broader theoretical superstructure of Wundt's psychological paradigm. This allows him to ignore Wundt's account of sensory elements and the causal role they play on affective elements. Reisenzein then provides a simplified account of affective elements, ignoring volitions entirely and treating affects as compounds of feelings. He begins his formalization by centralizing the three bipolar dimensions of feeling, the objects of those feelings, and the affects that emerge from them. The resulting formalization provides a faithful representation of the core of Wundt's theory, namely a reduction of emotions to more basic elements, without being entrenched in his other commitments.

Following the general structuralist set-theoretic strategy, Reisenzein defines the theory, TE(WUNDT), as containing the theory core, K, and the intended applications, I, with the theory core containing the core models, the constraints, and links to other compatible theories. In our presentation of Reisenzein's formalization, we focus on the basic definitions and axioms. The formal definition of the central terms is presented as follows (Reisenzein, 1992, pp. 150–151):

"x is a potential model of the theory *WUNDT*
($x \in M_p(WUNDT)$) iff there are *OBJECTS*, $\oplus$, *BASICS*, *AFFECTS*, q such that

(1) $x = \; < OBJECTS, \oplus, BASICS, AFFECTS, q >$;
(2) *OBJECTS*, *BASICS*, and *AFFECTS* are finite, nonempty, and pairwise disjoint sets;
(3) $\oplus : OBJECTS \times OBJECTS \rightarrow OBJECTS$ is an associative, commutative and idempotent [i.e., $\oplus\,(0,0) = 0$] operation;
(4) $BASICS = \{P, D, E, I, T, R\}$, with all $B \in BASICS$ being numerical functions $B \in OBJECTS \rightarrow R^{0+}$;
(5) $AFFECTS = \{\ldots A_i \ldots\}$, with all $A \in AFFECTS$ being numerical functions $A \in OBJECTS \rightarrow R^{0+}$;
(6) $q : AFFECTS \rightarrow V$; with
  (a) q is a unique (i.e., one-one) function and
  (b) $V = \{\langle v_1, v_2, v_3, v_4, v_5, v_6 \rangle \in R_t^6 : \sum_j^6 v_j = 1\}$."

Objects, basics, and affects serve as the object-propositions of emotions, basic emotions, and nonbasic emotions respectively. The object-propositions are represented in the formalization by non-zero positive real numbers that can be arranged and fused in any order without altering the result. The six basic feelings (or 3 bipolar pairs) described in the previous section (referred to as basics in the above formalization) can only be positive real numbers. Feelings can only arise as functions of specific objects, and objects themselves can be combined in any order to create new objects of feelings. A distinction is also made between the quality of an affect and its intensity. The quality of an affect is its composition, in which each of the feelings have values between 0 and 1 which sum to 1. The type of affect generated is therefore defined as the sum of the proportions of the six feelings that comprise it. The intensity values describe how prominent the emotion is in the individual's experience, with the total value being contingent upon the maximum value of the scale being used.

What follows in this section is a description of the three axioms Reisenzein presents as anchors of his formalization. He provides a structuralist account of the constraints, links, and specializations of the theory net, but we will address the axioms as they are the most pertinent to our account.

Axiom 1 regards the combination of six basic feelings and imposes a restriction on how we combine them to obtain emotions. Wundt maintains that while one cannot experience conflicting feelings regarding a single object, that one can experience conflicting feelings regarding its components upon evaluations of complex objects. Critically, conflicting feelings are not summed to obtain the total value of emotion (6 units of pleasure and 4 units of displeasure do not equal 2 units of pleasure). For example, one can simultaneously experience 6 units of pleasure from being gifted a cup of coffee but 4 units of displeasure from it being with milk instead of black. Wundt suggests that speaking in terms of net emotional states is inaccurate in that these experienced states, and therefore their objects, are distinct.

Axioms 2 and 3 address the fusion of feelings (i.e., basic emotions) of the same quality and of different qualities respectively. The fusion of feelings results in what Reisenzein refers to as affects (i.e., emotions). Regarding axiom 2, Reisenzein argues that Wundt is unclear about how to fuse basic feelings and goes forward with the assumption that when calculating the intensity of an affect the basic feelings regarding the same simple objects are summed (again, on the assumption that a single object cannot have a value for both feelings in a bipolar pair). For example, if one were to experience 4 units of pleasure from having a cup of coffee and 2 units of pleasure from that coffee having been free, one would experience 6 units of pleasure. However, as affects are fused (meaning sets of nonbasic feelings), it is important not to double-count any basic feelings regarding simple objects that exist in both affects. Simply, if affect A contains pleasure of the taste of coffee (4 units) and the pleasure of caffeine (1 unit), and affect B contains pleasure of the caffeine (1 unit) and that the coffee was free (2 units), the resulting complex affect would only count the pleasure of the caffeine once (taste + caffeine + free = 3 + 1 + 2).

Axiom 3 arguably requires the most consideration when converting the theory to a conceptual spaces account. Reisenzein interprets Wundt's ambiguity surrounding the combination of a variety of basic feelings, and Wundt's position on retaining the particular formulation of a given affect, to mean that the quality (its composition as opposed to intensity) of a given affect is represented by the sequential presentation of values of the six basic feelings. For example, if fear is described as an equal proportion of displeasure and excitement it would be defined as [Pleasure(0), Displeasure(0.5), Excitement(0.5), Inhibition(0), Tension(0), Relaxation(0)]. The particular combinations of feelings are combined in a way that preserves the qualities of the resulting affects. This allows one to define affects as proportional combinations of basic feelings. Reisenzein then goes on to formalize the intended applications of each of the defined terms and their functions, as well as how they interact to form concepts.

While presenting Wundt's theory in this way is unambiguous, it is also demanding. Complex relationships can be stated efficiently and with precision, but this comes at the cost of accessibility. As Reisenzein himself admits:

> "A problem that arises in this context, however, is that within psychology, formalizations still seem to constitute a communication barrier: Many psychologists seem to feel uncomfortable with formalized descriptions of theories or are at least uncertain about whether or not getting familiar with a formalization is worth the effort." (Reisenzein, 2000, p. 248)

The conceptual spaces account presented in the following section simplifies the relevant components by representing them as points and regions in three-dimensional space.

3. Wundt's theory in the conceptual space framework

Because Wundt's theory was proposed with a spatial metaphor in mind, the formalization of the theory in the conceptual space framework is rather straightforward. The three bipolar pairs of basic feelings constitute the three dimensions of our space, with pleasure–displeasure represented by the x axis, excitement–inhibition by the y axis, and tension–relaxation by the z axis (see Figure 2). Each dimension is supplemented with metrics. Those metrics can be related to the measurements of the corresponding qualities. Emotional states are therefore represented in the geometric space by points, with each point representing a unique state. For example, unique feelings of pure pleasure can be represented by any point with x value greater than 0, while both y and z values equal zero: such as (1,0,0) or (2,0,0). The same can be done with any of the other basic feelings like displeasure (−1,0,0), excitement (0,1,0), or tension (0,0,1). Representing non-basic emotional states is equally straightforward. If we assume that states of anger are constituted by the qualities of excitement and displeasure, they can be represented in the framework by any points whose x is less than 0 and y is greater than 0 for example, (−1,1,0). Similarly, if joy is constituted by pleasure, excitement, and tension it can be represented by a point such as (2,2,2).

The conceptual space framework automatically provides a measure of the dissimilarity of different emotional states in the form of the distance

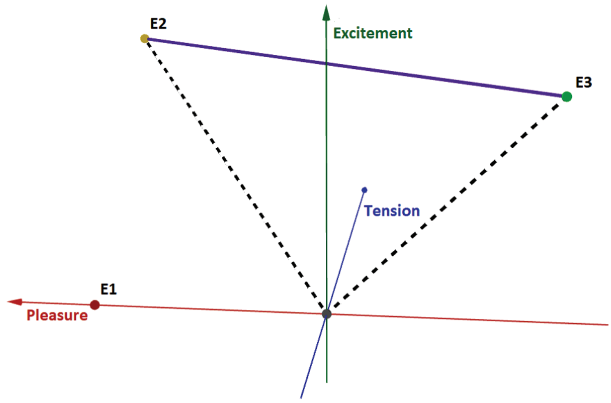

Figure 2. The conceptual space of emotion with axes representing basic feelings, with the neutral point at the intersection of the axes. Points E1, E2, and E3 represent emotion states. E1 (1,0,0) represents a state of pleasure, E2 (1,1,1) is a state of joy, and E3 (−1,1,0) represents a state of anger. The purple line between E2 and E3 represents the dissimilarity and equals 2.24. Dashed black lines represent the intensity of E2 and E3. They equal 1.73 and 1.41 respectively.

between the points representing the emotional states. The greater the distance between two points, the less similar the two emotional states represented by them. The dissimilarity of emotions was not defined in Reisenzein's formalization but is a natural feature of conceptual spaces and, as we will discuss below, can be used to generate a testable prediction.

Another quality easy to formalize in the framework is the intensity of an emotional state. Let us start from the point where the three axes intersect, the point (0,0,0). In our space, we can interpret this point as representing an emotionally neutral state in which no emotions are present (regardless of whether or not that is possible).[5] Plausibly, the intensity of such a state equals 0 as there is no emotion of any quality felt. Consequently, it is natural to expect that the farther, and therefore less similar, a given point is from this neutral point, the more intense the corresponding emotional state is. In light of that, the distance from the neutral point can be used as a measure of the intensity of a given emotional state.

This way of conceptualizing intensity is not only natural but also parsimonious. We do not need to introduce additional dimensions or a new conceptual space just to represent intensity. At the same time, as discussed in Reisenzein (1994), theories of emotion which include measures of both the quality and intensity of emotions are infrequent in the literature. This makes our formalization particularly attractive.

Further, a crucial element of Wundt's theory is the fusion of emotional states. It can be represented in many different ways in the current framework. We will discuss here simple vector addition. To use this operation to represent fusion, we can start by creating suitable vectors. For example, for each point representing an emotional state, we can easily construct a vector from the neutral point (0,0,0) to the given point. Any two such vectors can be added. The operation of adding two vectors starting from the same point is mathematically trivial, we just need to add the coordinates of the ending points of both vectors:

$$\vec{a} = (x^a, y^a, z^a), \vec{b} = (x^b, y^b, z^b) \;\; \vec{a} + \vec{b} = (x^a + x^b, y^a + y^b, z^a + z^b)$$

The resulting vector will lead from the neutral point to the point representing the new emotional state which is an effect of the fusion. For example, if we fuse an emotional state (1,0,-1) with an emotional state (1,0,1) we obtain a (2,0,0) state. This way of representing the fusion is just one of many possible options. A different approach would be to reduce the intensity of the result of the fusion by adding a parameter to the operation, for example:

$$\vec{a} +^* \vec{b} = 0.75(x^a + x^b), 0.75(y^a + y^b), 0.75(z^a + z^b)$$

This way of formalizing the fusion of emotions differs significantly from what was presented by Reisenzein. A formalization of the fusion operation analogous to his requires a six-dimensional conceptual space. This is discussed below.

Now that the pieces are in place, we can easily define categories of emotion – like pleasure, joy, or anger – as regions within the space. Concepts, both in general and those describing emotions, are sets of objects that can be defined by their similarity along particular dimensions. Hence, concepts are represented in the conceptual space framework by regions of the space. One of the existing methods used to determine geometric regions that correspond to concepts is the Voronoi Tessellation. This method starts by identifying paradigmatic examples of given concepts which are represented by points in our space, called generator points. For example, a generator point (1,0,0) can be considered a paradigmatic example of pleasure.

When our space contains one paradigmatic example for each emotional concept we wish to represent, we can divide the space into non-overlapping regions using the following rule: a region (e.g., a region representing the concept of pleasure) around a given center point (e.g., point (1,0,0)) contains only those points which are closer to this generator point than to any other generator point. In our case, the identified points will be those defined as paradigmatic examples of a particular emotion. This means that all of the points in the space that are closer to the paradigmatic state of joy than they are to any other paradigmatic emotions will be considered an example of joy.

In the present form, the formalization describes a quality space of emotional states directed toward simple objects. According to Wundt, no simple object can cause opposing feelings (e.g., pleasure and displeasure). Yet, this is possible in the case of complex objects as each feeling would be directed toward a different object. The emotional states containing opposing feelings (e.g., a bittersweet emotion containing both happiness and sadness) can be represented in the conceptual space framework. For example, it can be done in a six-dimensional conceptual space in which three dimensions are divided into two separate dimensions.[6] For example, the pleasure and displeasure dimension would be split into the dimension of pleasure and the dimension of displeasure. In such a space we can represent emotional states composed of opposite feelings and a fusion operation that can generate such states. This would be analogous to how it was done by Reisenzein. If two emotional states with opposite components are fused, both feelings can be combined in a new emotional state.

This section aimed to demonstrate that the central elements of Wundt's theory can be represented in the conceptual space framework straightforwardly. In the presented framework, emotional states and concepts are precisely defined, as are dissimilarity and intensity. The precision with which the elements of the theory are defined not only generates empirical predictions but also makes new hypotheses formulated in the context of the theory easier to test. Examples of testable predictions generated by the formalization are judgments concerning the similarity of given emotional states. For instance, if we represent emotional states of joy, pleasure, and anger by points (1,1,1), (1,0,0), and (−1,1,0), respectively, the formalized theory predicts that the state of joy is more similar to the state of pleasure than to the state of anger (see Figure 2). In the next section, we discuss an experiment presented in Douven (2016) in which analogous similarity judgments are tested. Another type of new empirical prediction of a theory is claims concerning the intensity of different emotional states. Results presented in Junge and Reisenzein (2013) suggest not only that testing of intensity claims can be used to indirectly apprise a theory of emotion but also that such appraisals are superior to alternative tests.

Once the emotional concepts are defined, it is clear how given states might be categorized (as pleasure, fear, etc). As such, a hypothesis generated in formalized theory is no longer merely verbal. When emotion-related concepts are described by well-defined regions of conceptual space, they can be related to measurements. For example, an individual can be asked to assess how pleasurable a given emotional state is along a Likert scale. Similarly, when defined in the framework, complex emotions can also be connected to (or alternatively defined by) the results of such tests. Moreover, the formalization is simple and driven by geometrical intuitions and therefore easier to understand than the structuralist formalization. In light of that, both our formalization and other possible formalizations of other psychological theories in the conceptual space framework seem to have more potential to be useful in psychology.[7]

Note that the construction we present is intentionally rather reductive to display the process in a simple manner. There are other components of emotions, particularly cognitive components such as evaluation, which can be added to the formalization proposed here. This can be done either in the same conceptual space, which we only limit to three dimensions for the sake of maintaining a visual presentation, or can be included in a complementary n-dimensional space within the formalization. As long as the components of a theory are quantitative, the relationships between their measurements can be mapped in conceptual space. The added benefit is the ability to ensure that testing hypotheses drawn from the theory address elements of the theory as explicitly as possible.

4. Beyond Wundt's theory of emotions

In this section, we show that the usefulness of the conceptual space framework we strive to demonstrate does not depend on the plausibility of Wundt's theory. Firstly we argue that other theories of emotions can be formalized in an analogous way, then show that tools developed in the conceptual space framework can be used to construct a new conceptual space of emotions based on experimental results.

4.1. Other theories of emotions

For present purposes, we can divide the theories of emotions into dimensional and nondimensional theories. The first group includes Wundt's theory and other similar theories of emotion that seek to reduce complex emotion to other simpler measurable elements which then can be presented in the dimensional space. Dimensional theories differ in the number of proposed dimensions. Watson and Tellegen (1985) or Russell (2003) proposed a two-dimensional theory, Osgood (1966) or Latinjak (2012) proposed a three-dimensional theory while Fontaine et al. (2007) argued that a four-dimensional space is necessary to "satisfactorily represent similarities and differences in the meaning of emotion words"(p. 1050). Pleasure-displeasure and arousal – relaxation dimensions are frequently included in such theories but there is significant variation in additional dimensions. Of course, neither the number of dimensions nor their naming conventions impact the process of formalization.

Dimensional theories remain popular in psychology. While it remains controversial which dimensions should be included, dimensional theories prove useful. For example, recent work in building computational models and neural networks that can recognize and classify various emotions has relied on dimensional theories because of their analytical approach (Chen & Shizhe, 2015; Lee et al., 2012).

That said, numerous researchers have argued against reductive theories of emotions (see, for example, Deonna & Teroni, 2012, de Sousa, 2006, or; Majeed, 2022). These critiques, along with similar concerns, have motivated non-reductive proposals. Cognitive theories of emotions, for instance, aim to highlight the distinction between automatic and controlled (or reflective) processes in their approach (e.g., Arnold & ARNOLD, 1960 or Lazarus, 1982). These cognitive theories often emphasize appraisal as either the core mechanism generating the emergence of affective states or as one of two central mechanisms, as seen in dual-process theories (the other being physiological arousal). Many of these theories rely on a sequence of processes or successive appraisals of emotional states relative to particular contexts (see, for instance, Smith & Kirby, 2009).

We do not intend to assert that non-reductive theories are inferior to dimensional, reductive theories, nor do we question the validity of arguments presented against them. Our primary objective is to demonstrate that the conceptual space framework can be effectively used to formalize psychological theories, offering certain advantages over the standard structuralist approach. In this regard, we have utilized Wundt's theory as an example, and it is not our intention to make claims about its adequacy.

The flexibility of the conceptual space framework makes it likely applicable to other theories of emotions and various psychological theories. For instance, it can plausibly be used to formalize appraisal theories. Due to their reliance on subjective functions like appraisal and a range of structures and processes, formalizing such theories, while certainly plausible, becomes a bit more complex. Factors such as the duration of emotional experiences (see, for example, Verduyn et al., 2011) or stress (see, for example, Folkman et al., 1986) could serve as dimensions in constructing corresponding conceptual spaces.

Another limitation of our case study is the static nature of Wundt's theory. Psychological theories typically include representations of dynamic phenomena such as evolution, correlations, and causal connections. We acknowledge that such elements may be more challenging than representing the static conceptualization of emotions constituting Wundt's theory. On the other hand, representing a dynamic aspect of theories in conceptual space the framework was discussed in the literature (see e.g., the sixth chapter of Gärdenfors, 2000). In light of that, we are confident that formalizing more dynamic theories in a conceptual space framework is possible. The formalization of such a theory goes beyond the scope of this paper. At the same time, the concepts of emotions formalized along the lines developed here can be used as building blocks, for example, in the conceptualization of theories describing some of the dynamic aspects of emotions.

4.2. A new conceptual space of emotions

A more principled approach to constructing conceptual spaces was developed in the conceptual space framework. It was used, for example, by Douven (2016) to construct a conceptual space describing vase and bowl shapes. Douven starts by preparing a chart composed of pictures of representative shapes of containers. He then conducted an empirical study in which participants were presented with two pictures and asked to assess the similarity of the two shapes. On the basis of the result, he was able to estimate how similar each of the shapes is to any other. Then the dimensions that best explain the detected differences were generated by multidimensional scaling. In a second experiment, participants were asked to assess which of the shapes were typical examples of cups and bowls. Based on those results, Douven was able to approximate regions of the conceptual space that correspond to the concept of bowl and cup.[8] The conceptual space constructed in this way is based solely on the results of empirical experiments testing the intuitions of the participants. Because of that, it is more generalizable than a similar space based only on the intuitions of its author.

As we have seen, it was argued that psychologists should develop new emotional concepts instead of relying on concepts inherited from natural language (e.g., Fiske, 2020). The strategy sketched above can be used to construct a conceptual space for emotions populated by such concepts. The structure of the study would be analogous to Douven's one. The construction of conceptual spaces of emotions may present more challenges than the reconstruction of bowl and vase shapes. It may be difficult to create a representative set of stimuli that elicit the range of emotions necessary for obtaining a comprehensive measure of similarity in the first stage of the study. There is no agreed exhaustive list of emotions, nor is it clear what kind of stimuli should be used. For example, the two experiments presented in Junge and Reisenzein (2013) used different stimuli to test different emotions. Stories about lotteries served as stimuli to elicit emotions of relief and disappointment of different intensities and pictures were used to elicit disgust. Additionally, the interpretation of dimensions generated by multidimensional scaling may not be straightforward. At the same time, the second challenge presents an opportunity. While the natural interpretation of the generated dimensions may overlap with the traditionally described dimensions (pleasure-displeasure etc.), examining newly generated dimensions may lead to discovering unknown factors that drive participants' intuitions concerning the similarity of emotions. Neither of those challenges seems to be impassable, nor do they make the Douven-like approach any less attractive.

5. Convexity hypothesis and future work

As we have seen, it is possible to reconstruct Wundt's theory of emotions in a conceptual space framework. The new formalization is less complex and demanding than a structuralist formalization and therefore more useful for psychologists. Additionally, we have shown how we can construct a new conceptual space of emotions on the basis of empirical results. In this section, we first describe how the convexity hypothesis, developed in the conceptual space framework, can inform the construction of psychological concepts. Finally, we will discuss possible directions for future work.

5.1. Convexity hypothesis and construction of psychological concepts

The convexity hypothesis can be used to verify the soundness of scientific concepts. The hypothesis claims that natural concepts, when reconstructed in the framework, will be represented by convex regions of the space. Natural concepts are an important class of concepts extensively studied in philosophy (e.g., Bird & Tobin, 2018; Jantzen, 2015). For our purposes the most important feature of natural concepts is that they are projectable, that is they can be used in inductive reasoning. A region is convex if, for any two points in the region, all points between these points also belong to the region. For example, a region shaped like a circle or a square would be convex, whereas a region in the shape of a star or a cross would not. The hypothesis was first presented in Gärdenfors (1990) and it is in part motivated by the notorious grue problem (Goodman, 1955). Grue is defined by Goodman as: "green before the year 2100 and blue afterward". The problem he posits is to explain why the hypothesis "All the emeralds are green" is supported by all previous observations of green emeralds, but the hypothesis "All the emeralds are grue" seems not to be supported by those observations. Goodman's question is what is the difference between natural, projectable concepts like green and non-natural and not projectable concepts like grue?

According to Gärdenfors (1990) this difference can be explained by the notion of convexity. Green, blue, and other natural projectable concepts, when reconstructed in the conceptual space, are represented by convex regions. In contrast, non-natural concepts such as grue or non-raven were shown by Gärdenfors (1990) to correspond to non-convex regions of a conceptual space. Therefore, convexity is seen as a necessary, though not sufficient, condition for natural concepts. This supposition is sometimes called *convexity hypothesis*.

A number of other arguments for the hypothesis are summarized by (Strößner & Schurz, 2020). For example, Jäger (2007) used a signaling game model to argue that convexity is an expected consequence of language evolution. In line with this argument, it was shown that many concepts from natural language are convex. For example, results of statistical analysis of typological data concerning color concepts collected by the World Color Survey (Jäger, 2010) strongly suggest that color concepts are convex. Results of experiments presented in Douven (2016) suggest that vessel concepts like vase or bowl are also convex. Similarly, Gärdenfors (2004) discusses evidence suggesting that concepts describing vowels in English are convex.

In light of these considerations, it seems that the hypothesis is well-supported by available evidence and therefore at least plausible. If so, it makes sense to consider are consequences of the convexity hypothesis for theory construction. The main consequence seems to be that the non-convex concepts should not be used in the formulation of hypotheses.[9] If a non-convex and therefore non-natural concept is present in a hypothesis, then testing it is just like testing a hypothesis claiming that all emeralds are grue. Plausibly the results of such testing would give us very little information concerning the truth of such a poorly formed hypothesis.

The convexity hypothesis may also be used in a post-factum examination of published experimental results. If a psychological hypothesis contains non-convex concepts the experimental result in question should be, at the very least, treated with caution.

Are such non-convex concepts present in psychology? The assessment of the convexity of the concepts used in published papers is not an easy task. In most cases, it is necessary to reconstruct the conceptual spaces of the given concepts. Yet, in some cases, it is reasonably clear that the concepts in question are non-convex. For example, considered the result of Annett and Kilshaw (1984):

> "The lateral preferences and L–R [Left–Right] skill of 109 male and 20 female dyslexics were as expected if the distribution of lateral asymmetry *is shifted less far to the right in dyslexics than in controls*. Several aspects of the data were consistent with Annett's hypothesis that some dyslexics lack the left hemisphere speech-organizing factor postulated by the right shift theory of handedness and that this would be sufficient to account for the proportion of affected relatives. *Some dyslexics were strongly dextral and these differed from the less dextral cases in several ways which resembled the distinction between backward and retarded readers*." (Annett & Kilshaw, 1984, p. 376)[emphasis added]

It seems that the results (in italics) can be paraphrased as follows: dyslexia is associated with both left-handedness and strong right-handedness (see Bishop, 1990). The concept of "left-handedness and strong right-handedness" is clearly non-convex. If we reconstruct the one-dimensional conceptual space of handedness we can see that the region denoted by the concept is not connected and therefore not convex (see regions marked by blue lines in Figure 3). In her later work on the relation between dyslexia and handedness, Annett (2011) distinguishes two kinds of dyslexia: "phonological" dyslexia and "surface" or "dyseidetic" dyslexia. Results of experiments presented there suggest that "phonological" dyslexia correlates with left-handedness while "surface" or "dyseidetic" dyslexia correlates with right-handedness. The original hypothesis was thus split into two separate hypotheses by splitting both concepts used in it. Dyslexia was split into phonological dyslexia and dyseidetic dyslexia, while left-handedness and strong right-handedness was split into left-handedness and right-handedness. The new concepts correspond to convex regions of the conceptual space (see Figure 3).

In light of the results of Annett (2011), it seems plausible that the original concept of dyslexia was also non-convex. It was composed of two related phenomena which differ significantly in some of their properties, for example, they likely correlate with opposite deviations from the norm in

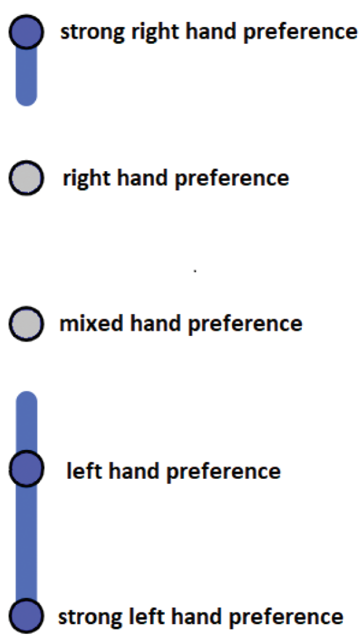

Figure 3. A one-dimensional conceptual space of handedness and a representation of the concept of left-handedness and strong right-handedness from Annett and Kilshaw (1984). Dots represent prototypical examples of the listed categories and blue lines represent regions corresponding to the concept of left-handedness and strong right-handedness.

handedness. If two types of dyslexia, combined in the original concept of dyslexia, are different enough, then the concept represented in a reasonable conceptual space will thus not be convex.

5.2. Future work

The conceptual space formalization is simple and driven by geometrical intuitions which make it simultaneously easier to understand than the structuralist formalization and better defined than a verbal theory. While in this paper we focus on emotions, we find the results encouraging for the possible formalization of other psychological theories.

The most obvious continuation of this work is to develop a full-fledged formalization of a dimensional theory of emotion that is currently in use. To this end, one must (empirically) establish which points correspond to paradigmatic examples of different emotions, and which regions correspond to commonly used emotional concepts like joy or fear.[10] Alternatively, one can construct a new conceptual space of emotion along the lines described in section 4.2. Once such a theory space is established, any further empirical tests can be used to evaluate and refine the conceptual categories.[11]

It is not clear what shape the space containing all emotional states might look like. This can be tested in a series of experiments analogous to the first experiment conducted in Douven (2016), which evaluates the accuracy of given concepts against the robustness of the proposed dimensions. The only difference is that the stimuli need to be suitable to elicit as comprehensive a range of emotions as possible. With such stimuli, one can obtain similarity judgments and on this basis construct a populated space. If the stimuli were close to being complete, the space of the generated region will approximate the populated region of the space. An analogous comprehensive stimulus is present in the case of colors in the form of 1625 Munsell color chips. The chips can be used to visualize a CIELAB color space (see Douven & Gärdenfors, 2019). The shape of the populated region of the CIELAB color space resembles the shape of the spindle and gives us an approximation of the populated region of color space.

The regions of the conceptual space of emotions can be empirically defined, meaning their correspondence to emotion-related concepts can be mapped. This can be done for example in an experiment analogous to one conducted in Douven (2016). Once the data is plotted, we can test both if the shapes of those regions are supported by incoming data as well as whether the shapes adhere to the convexity hypothesis. In such experiments, participants could be asked to categorize cases of emotions and points in conceptual space corresponding to the concepts in question. Each emotional concept whose shape-space corresponds to the data and retains convexity is likely a sound and natural emotion concept.

Additionally, a conceptual space framework can be used to unify the evidence in the field of emotions from different sources. First, self-reporting of emotional experience is the most popular way of testing emotions. As we have seen, self-reports can be used to construct a theory of emotions and to test the predictions of the existing theories today. Second, one can also use behavioral emotion indicators like facial expressions as used in the experiment by Fontaine et al. (2007), which compiled a list of 144 features of emotions. Participants were asked to rate how likely it is that a case of a given emotion is accompanied by each of these features. On the basis of obtained likelihoods, the authors argue that four dimensions (evaluation-pleasantness, potency-control, activation-arousal, and unpredictability) are required to account for differences in the meanings of emotion-related words. Third, it's possible to record patterns of brain activity associated with emotional states. For example, Lindquist et al. (2012) conduct a meta-analysis on neuroimaging studies of emotion to test the hypothesis that the emergence of emotions is the result of a distributed process across a set of interacting brain regions. They find that a psychological constructionist account of emotions, which claims that emotions emerge from more basic psychological processes (similar suggestions were present in Wundt, 1897), offers the most plausible account of the emergence of emotions. All this evidence can be combined in a conceptual space framework. A way to represent complex concepts that need to be characterized by multiple regions in many different conceptual spaces is studied in a conceptual space framework. A typical example of such a concept is an apple. Apples possess typical colors, shapes, tastes, etc., and all those characteristics correspond to regions in different conceptual spaces (conceptual space of colors, shapes, tastes, etc.). Perhaps a similar multi-domain conceptual framework can be constructed for some of the emotional concepts. Such a unified approach would make it possible to take advantage of all ways of testing emotions.

6. Conclusion

We presented a novel formalization of Wundt's theory of emotions to display the applicability and relevance of the conceptual spaces framework for psychological theories. The formalization of Wundt's work was first attempted by Reisenzein, who nearly a decade after publishing his thorough structuralist account conceded that its complexity rendered it impractical for widespread adoption. In contrast to Wundt's verbal formulation and Reisenzein's structuralist one, we offer a formalization of the theory in the conceptual space framework, a paradigm first introduced by Gardenfors. We demonstrated how by building a theory's conceptual mapping in geometric space, researchers can develop a dynamic visual tool with which to both model hypotheses and define a theoretical structure. The formalization removes the ambiguity that plagues verbal theories without sacrificing comprehensibility or accessibility.

Though over a century old, Wundt's dimensional theory is not structurally dissimilar to other dimensional theories in use today. While the details of the theory itself are outdated, Wundt's theory and Reisenzein's treatment of it offered us a useful case study in our attempt to demonstrate that the conceptual space framework is one potential avenue for reducing vagueness in scientific theories. We hope that we demonstrated that the framework can be used in an iterative, ongoing effort to perfect the psychological concepts, which will lead to improvement in the predictive power of constructed theories.

Notes

1. We do not claim that the formalization by itself solves the replication crisis. Of course, improvements to statistical designs and to the reward system are necessary to successfully ameliorate the crisis (see Eronen & Bringmann, 2021, p. 785). However, in light of the problems caused by the verbal nature of psychological theories, it seems that formalization has to be a part of the solution.
2. These terms are variously translated, here we use translations from Reisenzein (1992).
3. The terms are sometimes used differently; Damasio (2004) and Hansen (2005) argue that emotions are fundamental, and the experience of the emotion is to be considered the feeling. To maintain coherence in the case study, we adopt the terminology of Wundt and Reisenzein.
4. It would be a disservice to Wundt to not mention the distinctions seeing as he took pains to draw them. Yet, it would similarly be a disservice to our readers to provide a formalization laden with outdated and needlessly confusing terminology.
5. According to Wundt, a state in which no emotions are felt is unlikely. Our formalization is not committed to the realizability of such a state. Yet, it seems to be a natural interpretation of the point (0,0,0). Alternatively, point (0,0,0) can be interpreted as a neutral emotional state.
6. Berrios et al. (2015) conduct a meta-analysis of mixed emotion. The results suggest that "mixed emotions are a robust, measurable and non-artifactual experience".
7. Measurements play a central role in quantitative theories of emotions. At the same time, such theories need to relate their explanations to measurable variables. We do not argue here that theories need to be reduced to formal models. Rather, we argue that the increase in formalization, particularly of the quantitative components, would help the appraisal of theories. As we have seen in the introduction, there have been tendencies for papers with opposing viewpoints to speak past one another instead of explicitly addressing one another's concerns. Formalization offers an avenue toward more explicitly defined realms of debate.
8. The main aim of Douven (2016) is to test the theory of graded membership developed in Kamp and Partee (1995). Because of that, he does not focus too much on the details of the conceptual space he developed. For example, he does not interpret all of the generated dimensions.

9. Arguments supporting the convexity hypothesis strongly suggest that non-convex concepts should not be used in inductive reasoning. Consequently, it seems clear that such concepts should not be employed in formulating empirical hypotheses. However, it is not clear if these concepts can play an auxiliary role in the process of theory construction, for example, during the derivation of consequences of the theory in question. Perhaps a good way to study this issue and provide additional support for the convexity hypothesis is to conduct historical studies to determine if non-convex concepts were present in successful scientific hypotheses and theories, and if so, what role they served.
10. Theories that aim to identify a set of universal, innate emotions are able to provide a plausible starting list of empirically testable affective concepts (Ortony, 2021).
11. The validity of a newly constructed conceptual space can be tested. First, one can test whether the generated dimensions correspond to quantitative and measurable attributes (see e.g., Trendler, 2009). Then, they can test whether the construct validity of the space and measurements correlated to the posited dimensions (see e.g., Embretson, 2016). Plausibly such tests and the adjustments based on their results can be used to further develop the space into a robust framework for the classification of emotions.

Acknowledgements

Frank Zenker is the author of the idea of using conceptual spaces to formalize psychological theories. We would also want to thank Müge Kuyumcuoğlu Tütüncüoğlu, Aran Arslan, Bican Polat, Seçil Aracı, Zeynep Burçe Gümüşlü, Rainer Reisenzein, Disa Sauter, Dilectiss Liu, Rafał Urbaniak, Bartosz Wcisło, and Frank Zenker for their useful comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research has been funded by the Polish National Science Centre (grant number 2019/35/B/HS1/03281) and The Scientific and Technological Research Council of Turkey (grant number 118C257).

ORCID

Michał Sikorski http://orcid.org/0000-0002-6251-0183
Ohan Hominis http://orcid.org/0000-0003-3928-4779

Author contribution

Michał Sikorski's contribution amounts to 70% of all work related to the manuscript and Ohan Hominis contributed the remaining 30% of the work.

References

Andreas, H., & Zenker, F. (2013). Basic concepts of structuralism. *Erkenntnis*, *79*(S8), 1367–1372. https://doi.org/10.1007/s10670-013-9572-y

Annett, M. (2011). Dyslexia and handedness: Developmental phonological and surface dyslexias are associated with different biases for handedness. *Perceptual and Motor Skills*, *112*(2), 417–425. https://doi.org/10.2466/10.19.24.PMS.112.2.417-425

Annett, M., & Kilshaw, D. (1984). Lateral preference and skill in dyslexics: Implications of the right shift theory. *Journal of Child Psychology and Psychiatry*, *25*(3), 357–377. https://doi.org/10.1111/j.1469-7610.1984.tb00158.x

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology*, *3*(3), 266–286. https://doi.org/10.1080/23743603.2019.1684822

Arnold, M. B. (1960). Emotion and personality. *Pediatrics*, *26*, 732–734.

Bagaria, J. (2020). Set theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (Spring 2020)*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=set-theory

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(533), 452–454. https://doi.org/10.1038/533452a

Balzer, W., Sneed, J. D., & Moulines, C. U. (Eds.). (2000). *Structuralist knowledge representation: Paradigmatic examples*. Brill.

Berrios, R., Totterdell, P., & Kellett, S. (2015). Eliciting mixed emotions: A meta-analysis comparing models, types, and measures. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00428

Bird, A. (2020). Understanding the replication crisis as a base rate fallacy. *The British Journal for the Philosophy of Science*, *72*(4), 965–993. https://doi.org/10.1093/bjps/axy051

Bird, A., & Tobin, E. (2018). Natural kinds. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (Spring 2018)*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=natural-kinds

Bishop, D. V. M. (1990). How to increase your chances of obtaining a significant association between handedness and disorder [PMID: 1701771]. *Journal of Clinical and Experimental Neuropsychology*, *12*(5), 812–816. https://doi.org/10.1080/01688639008401022

Chen, Q., & Shizhe, J. (2015). Multi-modal dimensional emotion recognition using recurrent neural networks. *ACM Press the 5th International Workshop – Brisbane, Australia* (2015.*10.26-2015.10*.26).

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943–950.

Damasio, A. R. (2004). Emotions and Feelings: A Neurobiological Perspective. In A. S. Manstead, S. R. Antony, N. F. Frijda, & Agneta (Eds.), *Feelings and Emotions: The Amsterdam Symposium* (pp. 49–57). Cambridge University Press.

Deonna, J., & Teroni, F. (2012). The emotions: A philosophical introduction.

de Sousa, R. (2006). Against emotional modularity. *Canadian Journal of Philosophy*, *36*, 29–50. https://doi.org/10.1353/cjp.2007.0034

Douven, I. (2016). Vagueness, graded membership, and conceptual spaces. *Cognition*, *151*, 80–95. https://doi.org/10.1016/j.cognition.2016.03.007

Douven, I., & Gärdenfors, P. (2019). What are natural concepts? a design perspective. *Mind and Language*, *3*(3), 313–334. https://doi.org/10.1111/mila.12240

Embretson, S. (2016). An integrative framework for construct validity. *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*, 102–123. https://doi.org/10.1002/9781118956588.ch5

Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*, *16*(4), 779–788. PMID: 33513314. https://doi.org/10.1177/1745691620970586

Fiske, A. P. (2020). The lexical fallacy in emotion research: Mistaking vernacular words for psychological entities. *Psychological Review*, *127*(1), 95–113. https://doi.org/10.1037/rev0000174

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A., & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology*, *50*(5), 992–1003. https://doi.org/10.1037/0022-3514.50.5.992

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–1057. https://doi.org/10.1111/j.1467-9280.2007.02024.x

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., Fidler, F., & Wicherts, J. M. (2018). Questionable research practices in ecology and evolution. *Public Library of Science ONE*, *13*(7), 1–16. https://doi.org/10.1371/journal.pone.0200303

Freedman, L., Cockburn, I., & Simcoe, T. (2015). The economics of reproducibility in preclinical research. *PloS Biology*, *13*(6), e1002165. https://doi.org/10.1371/journal.pbio.1002165

Gärdenfors, P. (1990). Induction, conceptual spaces and ai. *Philosophy of Science*, *57*(1), 78–95. https://doi.org/10.1086/289532

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. MIT Press.

Gärdenfors, P. (2004). Conceptual spaces as a framework for knowledge representation. *Mind and Matter*, *2*(2), 9–27.

Goodman, N. (1955). *Fact, fiction, and forecast*. Harvard University Press.

Guest, O., & Martin, A. E. (2020). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Hahn, L., Hoelzl, E., & Pollai, M. (2013). The effect of payment type on product-related emotions: Evidence from an experimental study. *International Journal of Consumer Studies*, *37*(1), 21–28. https://doi.org/10.1111/j.1470-6431.2011.01072.x

Hansen, F. (2005). Distinguishing between feelings and emotions in understanding communication effects. *Journal of Business Research*, *58*(10), 1426–1436. https://doi.org/10.1016/j.jbusres.2003.10.012

Harris, R. J. (1976). The uncertain connection between verbal theories and research hypotheses in social psychology. *Journal of Experimental Social Psychology*, *12*(2), 210–219. https://doi.org/10.1016/0022-1031(76)90071-8

Haslbeck, J. M. B., Ryan, O., Robinaugh, D. J., Waldorp, L. J., & Borsboom, D. (2021). Modeling psychopathology: From data models to formal theories. *Psychological Methods*. https://doi.org/10.1037/met0000303

Izard, C. E. (1977). *Human emotions*. Plenum Press.

Jäger, G. (2007). The evolution of convex categories. *Linguistics and Philosophy*, *30*(5), 551–564. https://doi.org/10.1007/s10988-008-9024-3

Jäger, G. (2010). Natural color categories are convex sets. *Proceedings of the Logic, Language and Meaning: 17th Amsterdam Colloquium*, Amsterdam, The Netherlands, December 16-18(pp. 11–20). Springer. https://doi.org/10.1007/978-3-642-14287-1_2

Jantzen, B. C. (2015). Projection, symmetry, and natural kinds. *Synthese*, *192*(11), 3617–3646. https://doi.org/10.1007/s11229-014-0637-5

Junge, M., & Reisenzein, R. (2013). Indirect scaling methods for testing quantitative emotion theories. *Cognition & Emotion*, *27*(7), 1247–1275. https://doi.org/10.1080/02699931.2013.782267

Kamp, H., & Partee, B. H. (1995). Prototype theory and compositionality. *Cognition*, *57*(2), 129–191. https://doi.org/10.1016/0010-0277(94)00659-9

Latinjak, A. (2012). The underlying structure of emotions: A tri-dimensional model of core affect and emotion concepts for sports. *Revista Iberoamericana de Psicología Del Ejercicio y El Deporte*, *7*, 71.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*(9), 1019–1024. https://doi.org/10.1037/0003-066X.37.9.1019

Lee, P., Teng, Y., & Hsiao, T. C. (2012, July). XCSF for prediction on emotion induced by image based on dimensional theory of emotion. *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation* (pp. 375–382 https://doi.org/10.1145/2330784.2330842).

Linden, A., & Hönekopp, J. (2021). Heterogeneity of research results: A new perspective from which to assess and promote progress in psychological science. *Perspectives on Psychological Science*, *16*(2), 358–376. https://doi.org/10.1177/1745691620964193

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143. https://doi.org/10.1017/S0140525X11000446

Majeed, R. (2022). Modularity and the politics of emotion categorisation.

Murphy, K. R., & Aguinis, H. (2019). Harking: How badly can cherry-picking and question trolling produce bias in published results? *Journal of Business and Psychology*, *34*(1), 1–17. https://doi.org/10.1007/s10869-017-9524-7

National Academies of Sciences, E., & Medicine. (2019). *Reproducibility and replicability in science*. The National Academies Press.

Nosek, B., & Lakens, D. (2014). Registered reports a method to increase the credibility of published results. *Social Psychology*, *45*(3), 137. https://doi.org/10.1027/1864-9335/a000192

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, *26*(5), 1596–1618. https://doi.org/10.3758/s13423-019-01645-2

Ortony, A. (2021). Are all "basic emotions" emotions? a problem for the (basic) emotions construct [PMID: 34264141]. *Perspectives on Psychological Science*, 1745691620985415.

Osgood, C. E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, *7*(1), 1–30. https://doi.org/10.1111/j.1467-9450.1966.tb01334.x

Pashler, H., & Wagenmakers, E.-J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528–530. https://doi.org/10.1177/1745691612465253

Reisenzein, R. (1992). A structuralist reconstruction of wundt's three-dimensional theory of emotion 1. In H. Westmeyer (Ed.), *The structuralist program in psychology: Foundations and applications*.

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, *67*(3), 525–539. https://doi.org/10.1037/0022-3514.67.3.525

Reisenzein, R. (2000). Wundt's three-dimensional theory of emotion. In W. Balzer, J. D. Sneed, & C. U. Moulines (Eds.), *Structuralist knowledge representation: Paradigmatic examples (poznan studies in the philosophy of the sciences and the humanities* (Vol. 75, pp. 75–219). Brill.

Rus, J. (2007). The munsell color system. https://commons.wikimedia.org/wiki/File:Munsell-system.svg
Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145
Sennet, A. (2021). Ambiguity. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy *(Fall 2021)*. Metaphysics Research Lab, Stanford University.
Smaldino, P. (2019). Better methods can't make up for mediocre theory. *Nature*, *575*(7781), 9. https://doi.org/10.1038/d41586-019-03350-5
Smith, C. A., & Kirby, L. D. (2009). Putting appraisal in context: Toward a relational model of appraisal and emotion. *Cognition and Emotion*, *23*(7), 1352–1372. https://doi.org/10.1080/02699930902860386
Strößner, C., & Schurz, G. (2020). Criteria of naturalness in conceptual spaces(??). *Cognitive Science*, *44*(2). https://doi.org/10.1111/cogs.12815
Swedberg, R. (2012). Theorizing in sociology and social science: Turning to the context of discovery. *Theory and Society*, *41*(1), 1–40. https://doi.org/10.1007/s11186-011-9161-5
Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1. The positive affects*. Springer.
Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory & Psychology*, *19*(5), 579–599. https://doi.org/10.1177/0959354309341926
van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high- verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science*, *16*(4), 682–697. PMID: 33404356. https://doi.org/10.1177/1745691620970604
Verduyn, P., van Mechelen, I., & Tuerlinckx, F. (2011). The relation between event processing and the duration of emotional experience. *Emotion*, *11*(1), 20–28. https://doi.org/10.1037/a0021239
Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. https://doi.org/10.1037/0033-2909.98.2.219
Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2017). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion*, *17*(2), 267–295. https://doi.org/10.1037/emo0000226
Westmeyer, H. (Ed.). (1989). *Psychological theories from a structuralist point of view*. Springer-Verlag Publishing.
Wierzbicka, A. M. (2013). *Imprisoned in English: The hazards of English as a default language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199321490.001.0001.
Witte, E., & Zenker, F. (2017). From discovery to justification: Outline of an ideal research program in empirical psychology. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01847
Wundt, W. M. (1897). *Outlines of psychology*. W. Engelmann.
Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, 1–37. https://doi.org/10.1017/S0140525X20001685