



# HHS Public Access

Author manuscript

*Emotion*. Author manuscript; available in PMC 2024 September 01.

Published in final edited form as:

*Emotion*. 2023 September ; 23(6): 1670–1686. doi:10.1037/emo0001192.

## The Representation of Emotional Experience from Imagined Scenarios

Leonard Faul,

Mary G. Baumann,

Kevin S. LaBar

Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, USA

### Abstract

One of the key unresolved issues in affective science is understanding how the subjective experience of emotion is structured. Semantic space theory has shed new light on this debate by applying computational methods to high-dimensional datasets containing self-report ratings of emotional responses to visual and auditory stimuli. We extend this approach here to the emotional experience induced by imagined scenarios. Participants chose at least one emotion category label among 34 options or provided ratings on 14 affective dimensions while imagining two-sentence hypothetical scenarios. A total of 883 scenarios were rated by at least 11 different raters on categorical or dimensional qualities, with a total of 796 participants contributing to the final normed stimulus set. Principal component analysis reduced the categorical data to 24 distinct varieties of reported experience, while cluster visualization indicated a blended, rather than discrete, distribution of the corresponding emotion space. Canonical correlation analysis between the categorical and dimensional data further indicated that category endorsement accounted for more variance in dimensional ratings than vice versa, with 10 canonical variates unifying change in category loadings with affective dimensions such as valence, arousal, safety, and commitment. These findings indicate that self-reported emotional responses to imaginative experiences exhibit a clustered structure, although clusters are separated by fuzzy boundaries, and variable dimensional properties associate with smooth gradients of change in categorical judgements. The resultant structure supports the tenets of semantic space theory and demonstrates some consistency with prior work using different emotional stimuli.

### Keywords

emotional experience; imagination; narratives; semantic space theory; affect; discrete emotions

---

The experiential structure of emotions is commonly represented as either a finite number of discrete entities (i.e., categories) or as a collection of affective properties that combine

---

Correspondence concerning this article should be addressed to Kevin S. LaBar, Center for Cognitive Neuroscience, Box 90999, Duke University, Durham, North Carolina 27708-0999. [klabar@duke.edu](mailto:klabar@duke.edu).

Leonard Faul, Center for Cognitive Neuroscience, Duke University; Mary G. Baumann, Center for Cognitive Neuroscience, Duke University; Kevin S. LaBar, Center for Cognitive Neuroscience, Duke University.

We provide all 883 scenarios and their corresponding ratings at the following Open Science Framework website: <https://osf.io/y8rw7/>. For an interactive visualization of the data and analyses, visit <https://labarlab.shinyapps.io/scenarios/>.

to construct the emotions we feel (i.e., dimensional/appraisal attributes). The former, categorical structure is a key feature of basic emotion theory, which posits that a core set of emotional states are uniquely represented in terms of expression, recognition, and underlying neural circuitry (Ekman, 1992; Keltner, Sauter, et al., 2019). For instance, facial expressions associated with anger, disgust, fear, happiness, and sadness have been shown to be discretely expressed and recognized across different cultures, suggesting that these categories are universally represented among humans (Ekman, 1992; Ekman et al., 1969; Matsumoto et al., 2008; but see Barrett et al., 2019). This conserved representation may have relevant evolutionary origins, as many emotions have parallel forms and functions in non-human primates and other mammals, as evidenced by both behavioral and neurophysiological evaluations (Cowen & Keltner, 2021). A growing number of investigations have also shown that multivariate analytical techniques can successfully classify different emotional states based on separable representations of neural and psychophysiological responses to emotion eliciting stimuli (Horikawa et al., 2020; Kassam et al., 2013; Kragel & LaBar, 2013, 2016; Nummenmaa & Saarimäki, 2019; Saarimäki et al., 2016), revealing potential biomarkers of discrete emotions (but see Barrett & Satpute, 2019; Clark-Polner et al., 2017). While earlier studies in this area identified one-to-one mappings between emotional states and activation in specific brain regions—such as fear with the amygdala and disgust with the insula (Murphy et al., 2003; Phan et al., 2002)—more recent approaches demonstrate that emotions are best represented as complex configurations of activity across the brain, but primarily centered in transmodal regions within the default mode network (Horikawa et al., 2020). Importantly, these patterns are distinct for different emotions but conserved across individuals. While such findings support the presence of discrete emotions, basic emotion theorists nonetheless acknowledge that affective states can exhibit some variation in expression, with the universality of an emotion depending on the modality assessed (Keltner, Tracy, et al., 2019). Some emotions, for instance, may be better expressed via facial behavior while others may primarily involve auditory components such as in speech prosody (Keltner, Sauter, et al., 2019). Accordingly, emotions may be best thought of as multimodal patterns of behavior and physiology that cluster into distinct families of emotional states.

Alternatively, psychological construction theories argue that emotions should be understood as a combination of underlying dimensional features, such as the degree of pleasantness (valence) and activation (arousal) evoked by an emotional event (Barrett, 2006; Barrett & Bliss-Moreau, 2009; Russell, 2003; Wilson-Mendenhall et al., 2013). Although an emotion may be ultimately labeled as sad or joyful, these labels are thought to emerge from sociocultural norms or past experiences with similar events and are highly contextualized, rather than derived from discrete, evolutionarily conserved systems. Emotions may be best conceptualized as emergent phenomena arising from neurocognitive systems that evaluate an event based on its properties and then implement conceptual knowledge to organize those events in a meaningful way. As such, constructivist theories argue against an inherent geometric or categorical structure to emotional space. Despite psychophysiological and neural evaluations supporting the presence of discrete signatures for different emotional states, the structure of these signatures are variable across studies (Siegel et al., 2018). Neural patterns that correspond with discrete emotion categories are therefore considered

by constructivist theorists to reflect abstract statistical summaries of the instances of an emotion category, but those representations are not thought to exist in nature. Rather, domain general neural systems, such as networks associated with approach or avoidance, integrate to produce varied instances of feelings that we label as emotions (Lindquist et al., 2012). Thus, according to constructivist theories, variation is the norm, and basic emotion models that impose a categorical structure on emotions risk undervaluing this variation (Barrett, 2016).

Appraisal theories propose a slightly different approach to emotion representation by suggesting that some minimal cognition is required to trigger an emotional response (Moors et al., 2013). That is, physiological changes corresponding to an emotional experience are the result of an initial appraisal as well as the object of further appraisal. By comparison, constructivist theories do not assume the same causal relationship (Brosch, 2013). Central to appraisal theories is the idea that emotional experiences are not states, but rather processes involving a sequence of appraisals and adaptive responses to the surrounding environment. These appraisals can include goal relevance, certainty, novelty, agency, and many more factors that interact with action tendencies and physiological responses to produce an emotional experience (Ellsworth, 2013). Neuroimaging studies have demonstrated, for instance, separable representations of intention and benefit appraisals that distinguish between the emotions joy and gratitude (G. Liu et al., 2020).

Despite these differing perspectives, we note that the categorical, dimensional, and appraisal perspectives outlined here are not necessarily mutually exclusive, and indeed recent work has suggested joint representations of categorical and dimensional emotional information. M. Liu et al. (2022), for instance, demonstrated that a latent set of facial movement signals can convey both categorical information (such as disgust) and dimensional information (such as negative valence and low arousal) within a facial expression. These multiplexed signals ultimately allow for the communication of broad-plus-specific emotional messages (M. Liu et al., 2022). Whether these multiplex signals are processed asynchronously (with one informing the other) or synchronously (with separable systems in parallel) remains to be further explored. Relatedly, neural evidence suggests that category information is processed earlier in frontotemporal regions, whereas dimensional representations are more evident in later limbic and temporal activation (Giordano et al., 2021), although more work is needed to determine whether such effects generalize across different modalities. Indeed, recent reviews of the literature suggest that empirical evidence is quite inconclusive and varied regarding the representation of emotional experience and cannot provide conclusive support for the models outlined above. Future research can provide more clarity to these issues by adopting methods that prevent biases in stimulus selection, sampling a broader range of emotional experiences than commonly examined, conducting cross-cultural investigations, and using multimodal techniques to characterize emotional episodes (Barrett et al., 2019).

Central to the ongoing discussion of emotion representation is the subjective experience of emotion -- that is, the way in which individuals report their emotional response to affective stimuli, and whether such reports reflect an emotional space that is clustered, multidimensional, or both. Self-reports are considered the gold standard for assessing the conscious, subjective experience of feelings, as they provide the most direct measurement of emotional appraisal (Keltner, Tracy, et al., 2019; LeDoux & Hofmann, 2018). Yet if the

full range of such experiences is not properly appreciated among empirical evaluations, then both categorical and dimensional emotion theories will be ill-equipped to make conclusive claims on emotion representation. Emotion researchers usually examine only a small set of basic emotions such as amusement, anger, disgust, fear, joy, sadness, and surprise, while dimensional appraisals are often reduced to only valence and arousal. Although specialized literatures have emerged to assess atypical categories of experience such as craving (Giuliani & Berkman, 2015), pride (Tracy & Robins, 2007), and nostalgia (Sedikides & Wildschut, 2018), these emotional states are rarely considered in combination with the typical core emotions proposed in basic emotion theories. Likewise, appraisals of certainty, commitment, dominance, safety, and other dimensions may combine in novel ways to construct emotions, but are less often measured than valence and arousal (Cowen & Keltner, 2017). Consequentially, some of the variation in emotional experience across empirical studies may simply result from researchers evaluating differing subsamples of the fully available emotional space. Although emotional self-reports cannot verify the neurobiological representation of affective states, they nonetheless provide an indication of how people characterize their emotions with semantic terms and, relatedly, how this semantic space is geometrically organized.

## Semantic Space Theory

Recent computational approaches to examining emotional self-report data have illustrated novel analytical methods that can be applied to capture systematic variation in the experience of emotions when evaluated across a broader range of categorical labels and dimensional appraisals than previously used in the literature (Cowen & Keltner, 2021). These methods have now been applied to multiple modalities of emotional stimuli in the past few years, including videos (Cowen & Keltner, 2017), vocalizations (Cowen, Elfenbein, et al., 2019), speech prosody (Cowen, Laukka, et al., 2019), faces (Cowen & Keltner, 2019), music (Cowen et al., 2020), and ancient art (Cowen & Keltner, 2020b). Findings from these investigations have collectively developed a new approach to emotion representation known as semantic space theory, which seeks to model emotion-related responses in terms of three core properties: *dimensionality*, *distribution*, and *conceptualization* (Cowen & Keltner, 2021).

*Dimensionality* refers to the number of distinct varieties of emotion that exist among emotional self-reports. Converging evidence suggests that people reliably distinguish a larger number of emotional experiences than the core emotions typically examined in basic emotion theory. Video elicitations, for instance, are organized among at least 27 varieties of affective states (Cowen & Keltner, 2017), whereas nonverbal vocalizations produce at least 24 distinct emotions (Cowen, Elfenbein, et al., 2019). These estimates indicate that the semantic space of emotional experience is high-dimensional and more complex than current models appreciate, with participants reliably dissociating emotional states less commonly examined in the literature such as awkwardness and confusion.

*Distribution* refers to the geometric structure of semantic space, such as whether discrete or fuzzy boundaries separate clusters of varied affective states. Visualizations of self-report data suggest the presence of fuzzy boundaries, whereby categorical judgements are

typically bridged by continuous gradients of emotional change (Cowen & Keltner, 2021). In other words, emotions seem to transition across “conceptually related chains of reported experiences” (Cowen & Keltner, 2017, p. E7903). The categories fear and surprise, for example, are bridged by composite experiences that have intermediate meanings. Thus, affective responses are not purely discrete from one another but rather systematically blended in meaningful ways.

Finally, *conceptualization* refers to the concepts that are best able to capture variance in emotional reports across stimuli, such as categorical descriptors, dimensional/appraisal properties, or some combination of them. One approach to evaluating this property of semantic space is to determine factors that unify meaning across these classes of ratings. Canonical correlation analysis (CCA), for instance, finds linear combinations of variables from two different sets of data that maximizes the correlation between the sets (Hotelling, 1935; Thompson, 1984). Applying this approach to self-report data has indicated that emotional appraisals such as valence, commitment, and safety are associated with concomitant variation in categorical representation (Cowen & Keltner, 2017). Importantly, though, categorical descriptors consistently account for a greater share of variance in reported emotional experience than dimensional appraisals (Cowen & Keltner, 2021). Thus, emotional responses are high-dimensional and best organized by categorical descriptors, although boundaries between category clusters are not discrete and still contain meaningful variation in dimensional appraisals.

Although these conclusions have already been demonstrated across a diverse array of auditory and visual stimuli (e.g., video, faces, music, speech, and vocalization elicitations), much remains to be explored regarding the generalizability of semantic space models to other modalities of emotional experience, such as to stimuli that contain more self-relevant information. Personal emotional encounters, for instance, are characteristically different from the passive observation of emotional situations occurring to or depicted by others. Numerous studies have shown that self-reference moderates the processing of emotional stimuli (Fields & Kuperberg, 2012; Herbert et al., 2011; Northoff et al., 2006; Zhou et al., 2017), and that self-referential material is particularly influential in facilitating affective biases in emotional memory, such as in the case of mood-congruent memory (Blaney, 1986; Gaddy & Ingram, 2014; Matt et al., 1992). Self-referential content is also more susceptible to varied interpretation, as two individuals may respond to a situation in different ways depending on past experiences with similar events. Moreover, certain emotions such as pride and guilt are expressed self-referentially (Zinck, 2008), and therefore difficult to capture with the auditory or visual stimuli previously used in semantic space studies.

Although collecting self-report data from situational encounters as they occur in real-time would be ideal, employing such an approach is challenging and too uncontrolled for the computational methods applied in semantic space studies, which necessitate sampling from many participants responding to similar sets of stimuli (Cowen & Keltner, 2021). Evaluating a stimulus modality such as imaginative scenarios can circumvent this methodological hurdle by constraining reported experience to a controlled set of hypothetical stimuli, while also exposing participants to self-referential content depicting real-life events. As a form of constructive simulation, imagination is supported by the same neurocognitive system that

underlies episodic memory more broadly, but places greater demands on this system by flexibly recombining event details into novel situations (Schacter & Addis, 2007). Indeed, emotional responses to imaginative scenarios converge with reactions to real-world events or stimuli (Joseph et al., 2020; Robinson & Clore, 2001). Studying semantic space theory in the context of imaginative experiences is therefore a logical next step in uncovering whether the categorical structure of emotions is preserved when participants play a more active role in constructing emotional events that are relevant to their personal lives.

### Emotional Mental Imagery

Mental imagery has a powerful relationship with emotion and memory. Compared to simply listening to verbal descriptions of events, generating mental images amplifies both positive and negative emotional reactions (Holmes et al., 2006, 2009; Holmes & Mathews, 2005) and are more easily confused with real memories (Mathews et al., 2013). Several mechanisms have been proposed to account for why mental imagery has such an influence on emotion, including the close link between imagery and perceptual processes (Ganis et al., 2004), as well as the priming of autobiographical memories that further extend and enrich an imagined emotional experience (Holmes & Mathews, 2010). Mental imagery is able to activate physiological and behavioral response systems that mirror real-life experiences, and thus can facilitate meaningful behavioral change (Ji et al., 2016; Lang, 1979). Indeed, simulating hypothetical events emulates the same mental processes that would operate if those events actually occurred (Moulton & Kosslyn, 2009), perhaps explaining the convergence between imagined and real reactions to emotional stimuli (Robinson & Clore, 2001).

Various methods have been used to evoke emotions via mental imagery, including cueing participants to recall previously experienced emotional events and asking them to mentally manipulate the contextual or perceptual features of the memories (e.g., Faul et al., 2020), simulating emotional situations that could occur in the future (e.g., D'Argembeau et al., 2008; Rasmussen & Berntsen, 2013), or reading/listening to descriptive narratives that portray hypothetical scenarios (e.g., Fields & Kuperberg, 2012; Holmes et al., 2006). The former two methods, in which participants are cued to retrieve or generate their own emotional scenarios, allows for imaginative experiences uniquely targeted to the individual. Such investigations have revealed, for instance, that imagining alternate versions of the past or constructing possible futures can evoke feelings of regret, relief, or optimism depending on the type of mental simulation that is performed (Peters et al., 2010; Roese & Epstude, 2017). Yet, despite the added sensitivity afforded by having participants generate these personalized mental images, the lack of a consistent stimulus set limits the feasibility of item-specific analyses to examine conserved emotional experience across individuals.

Thus, researchers can also provide participants with textual vignettes that describe the same content to each participant. This approach has been used extensively in research on moral psychology, where brief vignettes portray challenging and aversive moral situations that often evoke negative (e.g., disgust and anger) emotional reactions (Ugazio et al., 2012). Research on self-reflective emotions—such as pride and guilt—and their neural correlates have also benefited from using narrative scenarios, as these allow participants to mentally place themselves within events where they are responsible for the committed

actions (Morey et al., 2012; Takahashi et al., 2004, 2008). Many emotions, including anger, happiness, disgust, anger, and fear, can be effectively induced with mental imagery techniques (Siedlecka & Denson, 2019). Moreover, textual scenarios have also been used in emotion production studies to generate a wide range of emotional stimuli in the form of vocalizations (Simon-Thomas et al., 2009) or facial expressions (Cordaro et al., 2018), which can then be further analyzed with semantic space approaches (i.e., having participants rate the emotions portrayed in the vocalizations or facial expressions).

As mentioned, although the core tenets of semantic space theory have been tested on multiple modalities of emotional stimuli, to our knowledge no study has applied these same principles to emotional reactions directly arising from exposure to imaginative scenarios. Whether categorical ratings are still conserved across participants when they are asked to imagine themselves in hypothetical events remains to be explored but would provide an important test of the generalizability of previous findings using visual and auditory elicitations. Eliciting emotions via mental imagery requires more active engagement (since participants are responsible for generating their own mental images for each scenario), and consequently may result in more varied expression across stimuli. Imaginative scenarios also provide a better opportunity to test the semantic space of self-reflective emotions, as well as examine whether a similar distribution of emotion categories is observed when participants are imagining themselves as the subjects in episodic events (as opposed to viewing videos of others). However, the availability of text-based stimuli suitable for such an investigation is relatively scarce in the emotion research literature, as most studies use rather lengthy vignettes that often encompass only a few select categories of emotion. By comparison, the 2,185 stimuli used by Cowen & Keltner (2017) to investigate the semantic space of emotions elicited by videos were sourced from video aggregation websites. These videos consisted of short clips that on average lasted for only about five seconds, which allowed for examining a wide variety of emotion-eliciting content, as well as obtaining self-report measures from the same participant for multiple stimuli. Extending this approach to imaginative experiences requires many brief and descriptive scenarios that span a variety of emotions, although such an endeavor lacks an appropriate public database. The construction of such a stimulus set would not only allow for examining semantic space theory in the context of mental imagery, but also provide normed scenarios suitable for a wide range of studies that seek to induce emotional responses via imagination.

### The Current Study

Accordingly, here we collected self-report ratings for hundreds of hypothetical scenarios, and then subsequently tested the semantic space of these responses. To this end, we first searched the literature to obtain scenarios that were appropriate for such an evaluation. To our knowledge, the largest collection of suitable scenarios currently available was developed by Fields and Kuperberg (2012) to evaluate the interaction of self-relevance and emotion on the neurocognitive processes underlying discourse comprehension. The authors created two-sentence scenarios, each with two self-relevance conditions (self and other) crossed with three emotion conditions (pleasant, neutral, and unpleasant). Each scenario was written in the present tense, with the first sentence introducing a situation that was neutral or ambiguous in valence (from either the self or other perspective – e.g., *A man knocks*

on Sandra's/your hotel room door), while the second sentence continued the scenario but contained a single critical word that shifted the outcome of the event to be pleasant, neutral, or unpleasant (e.g., *She/You see(s) that he has a gift/tray/gun in his hand*). These scenarios have been used in multiple studies to examine how neurophysiological signals represent unique interactions of self-relevance and valence, and imagining these events consistently evoke a range of emotional responses among participants (Fields et al., 2019; Fields & Kuperberg, 2012, 2014, 2016). However, the scenarios have yet to be rated on the full set of categorical and dimensional qualities that are necessary to evaluate their corresponding semantic space. The situations depicted are also constrained to the sentence structure of the multi-variant format, which may limit the degree to which these scenarios represent the full emotional space that can be expressed via imagination.

In the current study, we extended the computational methods from semantic space theory to emotional responses evoked by self-relevant hypothetical scenarios, not only by evaluating the self-perspective scenarios provided in Fields and Kuperberg (2012), but also by norming hundreds of new scenarios that we developed in-house. We examined whether previous findings regarding the dimensionality, distribution, and conceptualization of emotions from visual and auditory elicitations can be replicated when instead evoked via imagination. Following the methods and analytical approach of Cowen and Keltner (2017), we recruited a sample of online participants and asked them to report their emotional experience when imagining various hypothetical scenarios, either by selecting among discrete category labels (e.g., disgust, excitement, joy, pride, etc.) or by rating dimensional properties of their response to the imagined scenario (e.g., arousal, dominance, effort, valence, etc.). With these data in hand, we then evaluated the following: (i) the agreement among participants when choosing category labels and whether these concordance rates could be reduced to a smaller number of components, (ii) if category judgements exhibit a clustered structure and, if so, whether clusters are arranged with distinct or fuzzy boundaries, and (iii) how concordance rates compare with dimensional ratings of the same stimuli. In doing so, we contribute novel insights to semantic space theory while also providing a shared resource for emotion researchers. This resource contains a large library of hypothetical scenarios normed on a diverse set of emotional properties that are suitable for experimental designs requiring brief yet descriptive text stimuli, along with the normative ratings and data visualizations of the semantic structure.

## Method

### Survey Design

All participants were recruited via Amazon's Mechanical Turk and ratings surveys were hosted on Qualtrics. The experimental procedures were approved by the Institutional Review Board at Duke University, and all participants gave informed consent. We only recruited participants that were located in the United States and in the age range of 18-68 years old. Participants were required to have an approval rating greater than or equal to 90% (in later surveys this was raised to 95%) and at least 50 approved Human Intelligence Tasks (HITs; in later surveys this was increased to 100 approved HITs). The survey consisted of consent,



instructions, the ratings task, and demographic questions (see Figure S1 for an overview of the task design and example trials).

Each participant completed one of two ratings tasks. In the categorical survey, participants selected one or more categories among 34 options that best represented the emotion felt while imagining the two-sentence scenario shown in a banner at the top of the screen. These category labels were the same as those used by Cowen & Keltner (2017), except for *Boredom* which was replaced with *Neutral* (see Table S1 for the full list of category labels provided). As an exploratory measure, we also asked participants to rate the degree to which they experienced the selected emotion(s) on a scale of 1 (*not at all*) to 7 (*very strongly*), although these ratings were not incorporated into the present analyses (see Figure S2 for the distribution of these ratings). In the dimensional survey, participants indicated the extent to which they felt 14 dimensions of emotional experience (Table S1), including approach, arousal, focus, certainty, commitment, control, dominance, effort, fairness, identity, obstruction, safety, upswing, and valence (Cowen and Keltner, 2017). In both surveys, the order of the categories or dimensions were randomized for each participant, and each participant rated a total of 30 scenarios randomly selected from varied sets of available stimuli. Participants were paid \$5 for completing the survey. Data collection continued until all scenarios were rated by at least 11 different raters for both survey types, in order to match the number of raters collected for the video stimuli in Cowen & Keltner (2017).

We initially normed the two-sentence, second-person perspective scenarios created by Fields & Kuperberg (2012). These scenarios were designed to contain a single, critical word in the second sentence that shaped the outcome of the event to be either pleasant, neutral, or unpleasant. We normed 180 of these scenarios, each of which contained three variants (540 total stimuli). To ensure that the same participant did not rate variants from the same scenario, we organized the variants into separate sets that were assigned to different surveys on Qualtrics, and each participant was only allowed to complete the survey once. After norming these initial sets of scenarios, we then created 343 additional scenarios that were designed to capture emotional categories less represented among those created by Fields & Kuperberg (2012). These stimuli maintained a two-sentence structure with second-person perspective but did not adhere to the variant format (i.e., did not have a critical word that shaped the outcome of the scenario to be either positive, neutral, or negative).

### Inclusion and Exclusion

We included several attention checks throughout the survey to identify and remove inattentive participants, including free-response questions to ensure comprehension of general task instructions, an attention-check scenario that prompted specific ratings from the participant, asking participants to indicate whether they took the survey seriously, and checking for straight-lined responses (see supplemental materials for more details). In total, 1,219 submissions were accepted and paid for completing the survey, although we only analyzed data from individuals who passed all the required attention checks. After applying these exclusion parameters, the final data set containing both categorical and dimensional ratings consisted of 796 participants ( $M_{\text{Age}} = 36.8$ ,  $SD = 10.2$ , 342 females, 452 males, 2

non-binary), each of whom passed all attention checks and provided ratings for 30 different scenarios. See supplemental materials (Table S2) for an overview of all demographic information.

### Statistical Analysis

All statistical analyses were completed in R (R Core Team, 2021). We followed the same methods as Cowen and Keltner (2017) by calculating the concordance rates for each category as the proportion of times that the category was chosen for each scenario. We then performed a principal component analysis (PCA) of these concordance rates, which aims to reduce the dimensionality of the data, thereby retaining a smaller number of variables with minimal loss of information. By calculating concordance rates and then submitting those rates to a PCA, we can determine the general agreement among raters across stimuli, as well as the number of categories that reliably represent distinct emotional meanings. For instance, categories that are consistently co-endorsed across stimuli will load onto the same component in the PCA, thereby resulting in a smaller number of components than the initial number of categories. This analysis provides an estimate of the *dimensionality* of semantic space associated with imaginative scenarios. PCA was then followed by cluster visualization with t-distributed stochastic neighbor embedding (t-SNE). This technique provides a qualitative, visual overview of the clustering of the stimuli. Emotional scenarios with shared affective meaning (across all components) will cluster together, whereas those with different meanings will repel. By using this method, we can discern whether stimuli group into meaningful clusters, as well as how clusters transition among one another (e.g., via discrete or fuzzy boundaries). This analysis allows us to discern the *distribution* of self-report ratings to emotional scenarios. Finally, to compare the categorical and dimensional ratings with one another, we then performed CCA across the two datasets. CCA measures the association between two sets of variables by determining pairs of canonical variates that maximize the correlation between the two. In doing so, we can determine which dimensional attributes account for the most variation in categorical judgements (and vice versa). From this analysis, we can also determine whether the dimensional attributes best capture variance in the categorical judgements, or whether the categorical judgements best capture variance in the dimensional attributes. CCA thus allows us to identify the best *conceptualization* of the variance in self-reported emotional responses. Each of these analyses and their technical details are described below.

PCA was performed with the *stats* package using the ‘prcomp’ function. Data were centered, but not scaled, for this analysis since the same scale was used across all categories (i.e., proportion rates, ranging from 0 to 1). Varimax rotation was subsequently performed with the ‘varimax’ function, using Kaiser normalization. Following the methods of Cowen and Keltner (2017), all analyses that employed t-SNE and CCA used the loadings of each stimulus onto the varimax-rotated PCA factors, rather than the original concordance values.

Our primary method of component selection for varimax rotation was based on the cumulative proportion of variance explained by the components derived via PCA. However, to examine the reliability of the categorical ratings across different sets of raters, we also applied the same split-half CCA (SH-CCA) approach used by Cowen & Keltner (2017).

First, we randomly allocated half of the raters for each stimulus to Set A and the other half to Set B. For each set, we then calculated the concordance rates for all categories across all stimuli. Next, CCA was performed to compare the concordance rates from Set A to those of Set B, with a leave-one-out cross-validation procedure. That is, one stimulus was held out at a time and CCA was performed on the remaining stimuli between Sets A and B, resulting in a loading of Set A and Set B categories on Set A and Set B canonical variates. Given that both sets came from the same underlying population, canonical variate loadings and cross-loadings (i.e., structure correlations) were averaged across the pairs from each set. That is, the loading (structure correlations) of concordance rates from Set A on canonical variates from Set A were averaged with the loading of concordance rates from Set A on canonical variates from Set B (and vice versa). These loadings therefore reflect the extent to which category labels reliably associate with canonical variates derived from each set.

We then subtracted the concordance rates of the held-out stimulus from the average of the left-in stimuli and multiplied this difference with the averaged canonical variate loadings derived from each set, thereby computing a set-specific loading of the held-out stimulus on each canonical variate. Loadings were computed for each stimulus and concatenated together, resulting in a loading of each stimulus on all the canonical variates for Set A and for Set B. These loadings were then compared between Set A and Set B via linear regression. Since variate loadings for the held-out stimuli are nonorthogonal to one another, the linear regression of each pair of canonical variate loadings from Set A and Set B controlled for all previous canonical variates. These steps were repeated 50 times, with each step randomly separating the data into two sets and conducting SH-CCA between the two sets. The resulting 50  $p$ -values were averaged for each of the canonical correlations, thus providing an estimate of the number of canonical dimensions with reliable variance across different sets of raters.

Visualizing the clustering of component loadings in a lower dimensional space was achieved with t-SNE, implemented in the *Rtsne* package (Krijthe, 2015). Parameters were specified such that the output dimensionality consisted of 2 dimensions and the initial normalization and PCA steps were skipped (since varimax-rotated component loadings were provided as input). The perplexity value was set at 30, with a theta value of 0 (standard t-SNE implementation), and the analysis consisted of 5000 iterations. To assign colors to the resulting map, we first assigned a unique color to each of the components derived via PCA. Then, we determined all the components on which each stimulus loaded positively and constructed a weighted blend of the associated colors using the ‘mixcolor’ function from the *colorspace* package. Weights of each color were based on component loadings, rescaled to a range of 0-1 (i.e., the maximum loading possible across all scenarios and all components was transformed to a value of 1).

CCA was conducted between the raw dimensional ratings and the varimax-rotated PCA loadings, using the ‘cancor’ function in *candisc* (Friendly & Fox, 2021). The significance of each canonical variate was tested using the *CCP* package with Wilk’s test statistic. Canonical weights for each category and dimension variable are derived from the resulting coefficient scores, which represent the unique contribution of variables to a given canonical variate. Redundancy indices were calculated to determine the proportion of variance in

each set (categorical or dimensional) accounted for by variables in the other set (through each canonical variate) by using the ‘redundancy’ function from *candisc*. Redundancy is computed by multiplying the fraction of variance explained by a canonical variate in Set A or B by the fraction of shared variance between the pair of canonical variates from Sets A and B (i.e., the canonical correlation coefficient). In this way, a redundancy index indicates the proportion of variance in Set A accounted for by the variance in Set B (or vice versa) through a given canonical variate.

### Transparency and Openness

We report how we determined our sample size, all data exclusions, all measures, and the software used for analyses. The study’s design and its analysis were not pre-registered, but closely follow prior work on semantic space theory (Cowen & Keltner, 2017, 2021). We provide all 883 scenarios and their corresponding ratings at the following Open Science Framework website: <https://osf.io/y8rw7/>. For an interactive visualization of the data and analyses, visit <https://labarlab.shinyapps.io/scenarios/>.

## Results

### Categorical and Dimensional Ratings

We first investigated concordance rates among the category judgements, defined as the proportion of raters that agreed on a given category (Cowen & Keltner, 2017). For each scenario, we calculated concordance rates across all 34 categories. We analyzed category judgements from a total of 431 participants (see Table S2 for demographic information), which resulted in an average of 14.6 raters per scenario (minimum 11,  $SD = 2$ ). In total, 619 of the 883 scenarios (70%) were affiliated with at least one category that was agreed upon by 50% or more of raters. Except for envy, all other categories contained at least one scenario that a majority of raters selected as eliciting that particular category of emotion (Figure 1). When compared to the video concordance rates analyzed by Cowen & Keltner (2017), self-reports of emotional response to hypothetical scenarios exhibit a greater representation of anger, awkwardness, excitement, disappointment, satisfaction, and pride.

The other 365 participants from our final dataset (see Table S2 for demographic information) provided dimensional ratings on 14 affective dimensions, with an average of 12.4 raters per scenario (minimum 11,  $SD = 1.6$ ). Dimensional ratings produced varied distributions across the imagined scenarios (Figure 1). Effort ratings, for instance, show a unimodal, normal distribution, with most scenarios requiring a neutral level of effort to imagine, while relatively equal amounts demanded either more or less effort. Alternatively, valence was bimodally distributed, encompassing both negatively and positively valenced stimuli. Similar bimodal distributions were observed for approach, control, fairness, and upswing ratings. However, we note that in all these instances, the density distributions were skewed towards values above the neutral midpoint on the scale. Most dimensions were also strongly correlated with one another, as more pleasantness was associated with a greater desire to approach, more stimulation, more focus, more certainty, more commitment, more control, more dominance, more fairness, stronger group identity, less obstruction, more safety, and more upswing (see Figure S3 for an overview of all correlations). This

correlational pattern generally matches with what we observed in the data from Cowen & Keltner (2017), with the notable exception of effort, which was negatively correlated with most dimensions among the video stimuli but was uncorrelated or lowly correlated with all other dimensions among the scenarios (Figure S3). We note, however, that the distributions of affective dimensions are constrained by the stimuli that were made available to participants (e.g., the ratio of unpleasant to pleasant stimuli) and many of these dimensions can interact non-linearly to produce unique emotional experiences (Figure S4). Thus, examining the dimensional data is better suited to a multivariate statistical analysis of canonical correlations (see below), which appreciates this multidimensionality by differentially weighting each dimensional appraisal in relation to corresponding weights from categorical labels.

### Distinct Varieties of Reported Experience

Prior to conducting CCA, we first determined whether the categorical judgement data could be reduced to a smaller set of underlying components that capture distinct varieties of reported experience. To do so, we examined the reliability of the concordance rates across stimuli with SH-CCA (Cowen & Keltner, 2017). In brief, SH-CCA tests canonical correlations between two random sets of raters and computes the loading of each stimulus onto the resulting canonical covariates. Loadings are then correlated between the two sets, producing an estimate of the number of dimensions needed to explain the reliability of emotional reports across all stimuli. Repeating this process with multiple iterations estimates the number of reliable canonical variates that unify categorical judgements across varied sets of raters (see Methods for further detail). This analysis indicated 30 significant canonical variates at  $p < .05$  and 24 at a more stringent threshold of  $p < .01$ , suggesting that the emotional reports in our dataset reflected at least 24 reliable varieties of experience.

Yet, although simulation studies have supported the use of SH-CCA in applications such as these (Cowen & Keltner, 2017), this approach deviates from standard methods of component selection such as the proportion of variance explained, which more directly identifies the number of components needed to explain the variance in the original data. Thus, to apply a more validated threshold, we therefore also examined the cumulative proportion of variance explained by each component, thereby discerning the number of components needed to reduce the dimensionality of the dataset while still effectively describing the original concordance rates. Reproducing the PCA from Cowen & Keltner (2017), for instance, revealed that only 23 components were necessary to account for 95% of the total variance explained, even though 27 components were identified via their implementation of SH-CCA. Similarly, when we conducted the same analysis of explained variance on the present data, we observed that 24 components were needed to account for 95% of the total variance explained in the PCA, while 31 components were needed to account for 99% of the total variance.

To investigate how these components map onto the category judgements, we therefore extracted the 24 components that explained the most variance in the data, followed by varimax rotation to improve interpretability. For each component, we then determined the loading of the component with the original category judgements. The results reflect a similar

structure as that observed by Cowen & Keltner (2017), with most components associating with single category labels (Figure 2A). Notable exceptions include adoration (co-loaded with admiration and romance), aesthetic appreciation and entrancement (co-loaded with awe), contempt (co-loaded with anger), empathic pain and sympathy (co-loaded with sadness), horror (co-loaded with fear), sexual desire (co-loaded with romance), and triumph (co-loaded with pride). Interestingly, sadness not only loaded onto the same component as sympathy and empathic pain, but also onto a different component primarily representative of disappointment. Envy had negligible loadings on all components. The distribution of category concordance rates for each component are provided in supplemental materials (Figure S5).

Using t-SNE, we visually examined this semantic space of emotional experience by applying a clustering algorithm to the loadings of each stimulus onto the 24 principal components that were identified via PCA. This method projects the data into a lower-dimensional space consisting of two nonlinear axes that accurately preserve local structure in the data (i.e., clusters) while approximating global structure (i.e., distances between clusters). Previous applications of t-SNE on emotional self-report data have indicated the presence of fuzzy boundaries that separate clusters of emotional experience (Cowen & Keltner, 2021), and here we observed similar gradients of change in the emotional experience of imagined scenarios. Rather than grouping into distinct clusters with sharp borders, the categorical data typically organized into linked chains of affective experience, transitioning among related categories by crossing intermediate zones of mixed emotions (Figure 2B). Of note, though, the nostalgia, craving, and romance clusters did separate into particularly discrete representations, similar to the findings of Cowen & Keltner (2017) with video elicitations. In sum, consistent with previous evaluations of semantic space theory, we show that emotions elicited by hypothetical scenarios are organized in a high-dimensional space of semantic meaning, but that categories can have overlapping boundaries within this space.

### Identifying Factors that Unify Categorical and Dimensional Ratings

Semantic space theory suggests that some semantic meaning is shared between categorical judgements and dimensional attributes (Cowen & Keltner, 2021), and this shared variance is commonly evaluated by applying CCA to the categorical and dimensional data (Cowen & Keltner, 2017). CCA computes canonical variates (linear combinations of variables within each set) that maximize the correlation between sets, thereby identifying latent factors that unify variance in the data. Following the methods of Cowen & Keltner (2017), here we used the category component loadings obtained via PCA to maximize the orthogonality of reported emotional states from the concordance rates. Upon implementing CCA, we found that ten of the 14 canonical variates were significant (all  $p < .01$ ), and these variates were usually comprised of primarily one or two dimensional features (Figure 3A). Valence, for instance, negatively loaded onto the first canonical variate, indicating that category components loading positively onto this same canonical variate were generally rated as unpleasant to imagine (e.g., fear/horror, sadness, and anger). This finding matches closely with the CCA conducted by Cowen & Keltner (2017), where valence also accounted for the most variation in categorical endorsement.

Before interpreting the category and dimensional coefficients on additional canonical variates, however, we must acknowledge a few important caveats about CCA. First, the original distributions of the dimensional ratings should be considered when interpreting the weights associated with each canonical variate. Valence ratings were distributed both below (unpleasant) and above (pleasant) the neutral midpoint on the rating scale, indicating that high positive and negative loadings on the first canonical variate reflect both endpoints. Other dimensions such as attention and arousal, however, were skewed towards ratings of “focused” and “stimulated” (similarly observed in Cowen & Keltner, 2017). As such, a canonical variate consisting of a negative loading from arousal does not necessarily indicate subdued (as defined in the original rating scale), but rather *less stimulating*.

Further, we note that the loading of a category onto a canonical variate is relative to the ratings on those dimensions for all other categories, as well as loadings on previous canonical variates. That is, each canonical variate accounts for additional variation that was left unexplained by previous canonical variates. Canonical coefficients therefore reflect relative, rather than absolute, differences between variables. Finally, although we have labeled the canonical variates based on the affective dimensions that primarily load onto them, each canonical variate actually represents a combination of features, and therefore should be appreciated as a latent construct more complex than any individual dimensional or categorical variable. These points are important to consider when evaluating the semantic space of reported emotional experience via CCA, both in the present analyses and in previous applications.

By acknowledging these points, we can more precisely interpret the meaning of the other canonical variates. For example, the second canonical variate consisted of a positive loading from safety *in combination with* a negative loading from arousal. When evaluating the corresponding loadings from the categorical judgements, we interpret this canonical variate as suggesting that higher ratings of safety combined with lower ratings of arousal account for variation in reports of neutral emotional experiences, while lower ratings of safety combined with higher ratings of arousal account for variation in fearful and horrific emotional experiences. Importantly, these associations are necessary to explain additional variation in neutral experiences that was unaccounted for by the first canonical variate (see Figure S4 for an example of how these effects emerge as non-linear associations among dimensions).

The third canonical variate primarily consisted of a positive loading from commitment in combination with a negative loading from fairness. Regarding the categorical loadings, sadness loaded positively on this same canonical variate. We thus interpret the third canonical variate as suggesting that higher ratings of commitment combined with lower ratings of fairness was associated with variation in reports of sad emotional experiences compared to other categorical endorsements. That is, commitment and fairness ratings were identified in the CCA as higher and lower, respectively, than what would be expected from the variance already explained by the first two canonical variates. Additional canonical variates further indicate that dominance, certainty, approach, identity, and obstruction all accounted for extra bits of variation in the categorical loadings (Figure 4). Again, each variate reflects relative differences after accounting for loadings on all previous canonical

variates, although we provide the absolute distributions of each dimensional appraisal (split by principal component) in supplemental materials (Figure S6).

CCA attempts to maximize the correlation between two sets of data, but the resulting canonical variates are often asymmetrical regarding the amount of variance explained in the original ratings. A canonical variate may, for instance, account for more variance in the dimensional ratings than the categorical ratings. To evaluate this proposal, we next conducted a redundancy analysis, which first determines the amount of variance accounted for in each dataset by their respective canonical variates, and then multiplies this value by the canonical correlation coefficient. This calculation yields a *redundancy index* for each set, which represents the proportion of variance in one set explained by the other through a given canonical variate. Summing the redundancy indices across all canonical variates provides an estimate of the total proportion of variance accounted for between the two sets.

By applying this method to the categorical and dimensional ratings of hypothetical scenarios, we found that dimensional ratings accounted for a total of 27% of the variance in the categorical ratings, whereas the categorical ratings accounted for a total of 64% of the variance in the dimensional ratings. Replicating the findings of previous evaluations of cross-judgement explained variance (Cowen & Keltner, 2021), this result indicates that the categorical judgements were better able to predict dimensional properties than vice versa. In other words, categorical judgements provided a richer evaluation of emotional experience resulting from imagined scenarios. In addition to comparing explained variance between datasets, the redundancy analysis also provides more context to the importance of each canonical variate. The explained variance in both directions was primarily captured by the first canonical variate (59% for categories to dimensions and 18% for dimensions to categories), and we observed a similar bias towards the first canonical variate in our re-analysis of the video data from Cowen & Keltner (2017). Therefore, despite 10 canonical variates significantly correlating the two sets of ratings, we must acknowledge that the first canonical variate is, by and large, most useful in accounting for the variance in the original data. The other variates capture additional bits of shared variance among relatively fewer categories and dimensions and should always be compared with respect to the variance explained by the first canonical variate (valence).

To facilitate further inspection of these unifying latent factors that were extracted via CCA, we then mapped the first six canonical variates onto the original t-SNE map produced from the PCA loadings (Figure 3B and 3C). Here we coded the red, blue, and green color channels of each depicted stimulus to reflect loadings on each of three different canonical variates. These maps show how several of the affective dimensions are able to capture differing gradients of change in categorical loadings. Similar to Cowen & Keltner (2017), for instance, gradients of change among unpleasant categories are primarily organized around shifts in safety, arousal, certainty, and commitment. Neutral and pleasant categories also shift among loadings on safety and arousal, as well as dominance.



## Discussion

The present study investigated the structure of self-reported emotional responses to imagined scenarios by applying the same computational methods as previous evaluations of semantic space theory (Cowen, Elfenbein, et al., 2019; Cowen, Laukka, et al., 2019; Cowen et al., 2020; Cowen & Keltner, 2017, 2020b). Our findings are consistent with several key observations that have now been shown across several modalities of emotional stimuli. First, at least 24 varieties of emotional experience are represented among the self-report data we collected, since reducing the dimensionality of the data any further would lose valuable information provided by the original categorical judgements. We therefore conclude that, like video, music, speech, and vocalization elicitation (Cowen & Keltner, 2021), the emotional space of imagined scenarios is relatively high-dimensional and reliably captures a broad range of emotional states. Second, while the categorical judgements did exhibit some discrete clustering of stimuli (most clearly distinguished for craving, fear, neutral, nostalgia, sadness, and romance), in most instances these clusters were separated by fuzzy boundaries with intermediary zones that transitioned one emotion into another along a smooth gradient of change. This finding suggests that the structure of the emotional space we examined is neither purely clustered nor uniformly distributed, but rather organized among families of emotional experiences that can blend into each other (Cowen & Keltner, 2021). Third, using CCA, we examined the relation between the categorical and dimensional ratings, finding that valence ratings (relative to other dimensional appraisals) accounted for the most variance in the categorical judgements, while ratings of safety, arousal, commitment, fairness, dominance, and certainty accounted for extra variance that could not be explained by valence alone. Fourth, categorical judgements accounted for more variation in the dimensional appraisals than vice versa, indicating that people more precisely conceptualize their emotional experience with category labels. In what follows, we discuss each of these points in more detail, while also considering important caveats and future directions.

Our finding of at least 24 varieties of reported emotional experience that primarily mapped onto single category labels is in a similar range as that reported by other investigations of semantic space theory, although here we applied a stricter criterion by simply extracting the number of components needed to account for 95% of the explained variance. Compared to the video elicitation, the hypothetical scenarios seem to be better at capturing more self-reflective emotional states such as pride, disappointment, and guilt, while less representative of more visually evoked emotions such as aesthetic appreciation and entrancement. These differences likely emerge from the self-referential nature of the scenarios, which were written as second-person narratives. Importantly, though, the emotional space we sampled is also dependent upon the stimuli that were provided to participants. Potentially more categories could be reliably captured if more stimuli were rated by participants. As with other evaluations of semantic space theory, the present findings cannot speak to the true number of categorical entities available in emotional self-reports of imagined scenarios but do suggest a higher dimensional space than basic emotion theories typically appreciate. Importantly, across narrative and visual elicitation (using nearly identical experimental designs), a number of consistent categorical entities are observed. These include admiration,

amusement, anger, anxiety, awe, awkwardness, calmness, confusion, craving, disgust, excitement, fear, interest, joy, neutral (labeled 'boredom' in prior studies), nostalgia, relief, romance, sadness, satisfaction, and surprise.

Regarding the distribution of the categorical judgements, our observation of clusters with fuzzy boundaries shows that participants often selected multiple categories to represent their emotional response to an imagined scenario. In most cases there were one or two categories more consistently endorsed across raters than others, although many other categories were still endorsed at lower levels. Nostalgia-inducing scenarios, for instance, appropriately exhibit strong concordance on nostalgia judgements, but also seem to capture a weighted blend of other emotional states that are both pleasant and unpleasant, similar to current empirical characterizations of nostalgia as an emotional state with mixed affect (Batcho, 2013; Sedikides & Wildschut, 2018). Rarely was a single category label agreed upon by all raters, with most scenarios eliciting significant variation in categorical judgements. This finding of blended affective states (similarly found among other visual and auditory modalities) suggests that the distribution of category judgements best reflects how stimuli group together. That is, differential weights placed upon each category judgement explains how semantic meaning smoothly transitions between clusters, rather than only the concordance rates of the highest endorsed category. We note, however, that the scenarios may have clustered differently if they were specifically crafted to reflect theorized functional associations of emotions. Elicitors of disgust, for instance, seem to organize into separate, superordinate domains related to food, pathogens, danger, sex, and morality (Amoroso et al., 2020). Constructing enough scenarios to fully capture each of these domains may have resulted in a more discrete representations of disgust. In a similar vein, positive and negative versions of surprise have been shown to reliably dissociate from one another in vocal bursts, providing more clarity to the structure of surprise within semantic space (Cowen, Elfenbein, et al., 2019). Relatedly, in the present data we show that sadness judgements primarily loaded onto two components, with one of these components representing shared variance with empathic pain and sympathy, while the other was more associated with disappointment. Future work should consider the possibility of additional functional dissociations that may exist within the clusters observed here, as well as the distinctiveness of other category labels that were not examined, such as embarrassment and shame (Miller & Tangney, 1994; Tangney, 1995).

When categorical judgements were combined with the dimensional ratings using CCA, the resulting canonical variates indicated that valence primarily accounted for variation among categories, followed by ratings of arousal and safety. This is to be expected, given that both valence and arousal are considered integral to core affect (Russell, 2003). As noted, although canonical variates other than valence accounted for substantially less variance in the data, they nonetheless captured unique associations among categorical judgements and dimensional appraisals than would be predicted from valence ratings alone. Fearful scenarios, for instance, were rated as less safe but also more arousing when compared to other unpleasant categories, reflecting the intricate link between threat detection and fear expression (Mobbs et al., 2009). By contrast, neutral scenarios were experienced as more safe but less arousing, when compared to their positioning among other categories with regard to valence. Importantly, our CCA replicated the finding from Cowen & Keltner

(2017) that commitment ratings accounted for variation in sadness judgements (and less so in disappointment judgements), reflecting extant empirical perspectives on the link between sadness and attachment processes (Albert & Bowlby, 1982). This finding does not necessarily indicate that sadness was perceived as higher on commitment compared to all other stimuli. In fact, excitement, joy, and romance elicited the highest ratings of commitment (Figure S6). Rather, sad scenarios elicited higher ratings of commitment than would be expected when also considering ratings of valence, safety, and arousal for the same scenarios (as accounted for by the first two canonical variates).

Although our CCA linked the categorical and dimensional ratings among unifying gradients of semantic meaning, most canonical variates primarily accounted for variation in just one or two categories. We also observed strong bias in our redundancy analysis, which showed that dimensional appraisals accounted for substantially less variance in the category loadings than vice versa. This result indicates that category judgements provide a richer representation of reported emotional experience, although the cause of this imbalance remains unclear. In line with basic emotion theory, emotions may be organized among conserved categories of affective states, which is correspondingly reflected in emotional self-reports. However, participants might simply find the category labels more intuitive to judge, whereas the dimensional scales require considering a continuum of possible responses that can be interpreted differently across raters. This limitation is perhaps best reflected in the arousal ratings, which ranged from *more subdued* to *neutral* (midpoint) to *more stimulated* (Cowen & Keltner, 2017). On average, few scenarios elicited ratings below the neutral midpoint, although more unpleasant stimuli were generally rated as less arousing (but see Figure S4 for examples of nonlinear associations in the data). These trends may reflect varying definitions of *subdued*, which can be interpreted as indicating an emotion with lower physiological intensity or alertness as is commonly used in the literature (Posner et al., 2005), or as a form of regulation, restraint, or even depression (Lexico, 2019). Thus, categorical labels might be better conserved in their semantic meaning across individuals, but this is not to say that they are inherently more present than the dimensional properties we intended to measure. Future research on semantic space theory should consider whether the definition of scale endpoints (which have been consistently used across semantic space studies) might arbitrarily amplify differences in variance explained between categorical and dimensional ratings.

Some of the scenarios used in our study contained specific emotional words within the scenario (e.g., upset, glad, amazed, etc.) that were synonymous with some of the category labels provided to participants. It is possible that some participants did not follow instructions to report on their own feeling states but instead simply tried to match the category labels to any emotion words that were contained in the scenario. This possibility may have contributed to the categorical labels providing a better account of the self-report data than the dimensional scales. However, we do not think that this is a major concern with the study results. First, only a small proportion of the total stimulus set (~2%) included words that were *identical* to those from the category label list. Second, in these scenarios, as well as others that contained related emotion words or words that referred to other emotion constructs, the emotion word could alternatively describe self-directed feelings, feelings of other parties in the scene, or some other facets of the scenario (e.g., a soothing

candle). Thus, label matching wouldn't work as a general heuristic. Third, the large list of emotions used for the ratings would make a deliberate strategy of label matching laborious. Fourth, we spot-checked several of these items and the self-reflective endorsements made sense given the context of the scenario (e.g., the scenario "a colleague talks about you at length...everyone notices the admiration in her description of your attributes" was rated low in admiration [0.07] but was higher in pride [0.5] and joy [0.57] concordances). Fifth, participants tended to endorse feeling multiple emotions for each scenario, not just the one that was similar to an emotion word present in the narrative; it is also not the case that only scenarios containing emotion words were endorsed categorically. Sixth, the data analysis strategies (e.g., loadings onto PCA components and location in the t-SNE cluster plot) did not rely on the highest-endorsed category as the sole determinant of the findings but rather considered the entire distribution of concordance rates across all categories. Finally, it is possible that the presence of emotion words could lead to a matching strategy for dimensional constructs as well as categorical labels (e.g., the mere presence of the word 'energize' in the scenario could lead one to endorse high arousal ratings).

For any behavioral study that relies on self-report, there is a concern that we cannot confirm if participants are truly experiencing a given emotion, or merely characterizing the perceived emotion that is portrayed in the stimulus. We note that a similar limitation exists with other stimuli used in evaluations of semantic space theory (e.g., videos, vocalizations, and music) for which it is difficult to dissociate emotional experience from emotion perception. Measuring psychophysiological indices of emotional experience and assessing their correspondence with behavioral reports will help to provide more clarity to this issue (Cowen & Keltner, 2020a). Recent evidence demonstrating categorical representations of neural data in response to emotional videos supports the tenets of semantic space theory beyond just behavioral self-reports (Horikawa et al., 2020). Likewise, many of the scenarios used in the present study have also been used to examine neural responses (measured via electroencephalography recordings and functional magnetic resonance imaging) that track the valence and self-relevance of the imagined events (Fields et al., 2019; Fields & Kuperberg, 2012, 2014, 2016), suggesting that these scenarios do differentially evoke emotional responses that can be captured with non-behavioral measures. Nevertheless, more work is needed to elucidate whether a similar semantic space of emotional experience induced via imaginative scenarios can be captured with other measures of emotional response beyond just self-report data.

We also note that the present study only examined self-reported emotional ratings from English speaking participants, who exclusively read scenarios written in English and provided responses based on English emotion concepts. More research is needed to determine how the geometric structure of emotional reports observed here may shift when examined with other cultures or languages. Recent work with speech prosody, musical stimuli, and facial expressions, for instance, has demonstrated core patterns of emotional expression that is preserved across participants from different cultures, including the United States, China, India, Japan, and Korea (Cordaro et al., 2018; Cowen et al., 2020; Cowen, Laukka, et al., 2019).

In sum, our findings indicate that categorical labels are most useful in organizing semantic meaning among emotional self-reports of imagined scenarios, and that categorical judgements reflect a relatively high-dimensional space of separable affective states. We replicated much of the previous work in semantic space theory that has been applied to visual and auditory stimuli, again concluding that emotional experience is perhaps best modeled as a blend between categorical and dimensional attributes. Although dimensional appraisals explained less variance in the category judgements, the category clusters were also not purely discrete in their representation, instead transitioning smoothly among families of blended emotions. This approach aligns with recent neurocomputational methods that appreciate emotional space as more complex than a single categorical or dimensional model, but rather as the interplay among neural systems that support categorical *and* dimensional representations at different timescales and in different regions of the brain (Giordano et al., 2021).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by a National Institutes of Health grant 1 R01 MH124112 (K.S.L.) and a National Science Foundation Graduate Fellowship (L.F.).

## References

- Albert S, & Bowlby J (1982). Attachment and Loss: Sadness and Depression. *Journal of Marriage and the Family*, 44(1). 10.2307/351282
- Amoroso CR, Hanna EK, LaBar KS, Schaich Borg J, Sinnott-Armstrong W, & Zucker NL (2020). Disgust Theory Through the Lens of Psychiatric Medicine. In *Clinical Psychological Science* (Vol. 8, Issue 1). 10.1177/2167702619863769
- Barrett LF (2006). Solving the emotion paradox: Categorization and the experience of emotion. In *Personality and Social Psychology Review* (Vol. 10, Issue 1). 10.1207/s15327957pspr1001\_2
- Barrett LF (2016). Navigating the Science of Emotion. In *Emotion Measurement*. 10.1016/B978-0-08-100508-8.00002-3
- Barrett LF, Adolphs R, Marsella S, Martinez AM, & Pollak SD (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, 20(1). 10.1177/1529100619832930
- Barrett LF, & Bliss-Moreau E (2009). Chapter 4 Affect as a Psychological Primitive. In *Advances in Experimental Social Psychology* (Vol. 41). 10.1016/S0065-2601(08)00404-8
- Barrett LF, & Satpute AB (2019). Historical pitfalls and new directions in the neuroscience of emotion. In *Neuroscience Letters* (Vol. 693). 10.1016/j.neulet.2017.07.045
- Batcho KI (2013). NOSTALGIA: The bittersweet history of a psychological concept. *History of Psychology*, 16(3), 165–176. 10.1037/a0032427 [PubMed: 23646885]
- Blaney PH (1986). Affect and Memory. A Review. *Psychological Bulletin*, 99(2), 229–246. 10.1037/0033-2909.99.2.229 [PubMed: 3515383]
- Brosch T (2013). Comment: On the role of appraisal processes in the construction of emotion. In *Emotion Review* (Vol. 5, Issue 4). 10.1177/1754073913489752
- Clark-Polner E, Johnson TD, & Barrett LF (2017). Multivoxel Pattern Analysis Does Not Provide Evidence to Support the Existence of Basic Emotions. *Cerebral Cortex*, 27(3), 1944–1948. 10.1093/cercor/bhw028 [PubMed: 26931530]

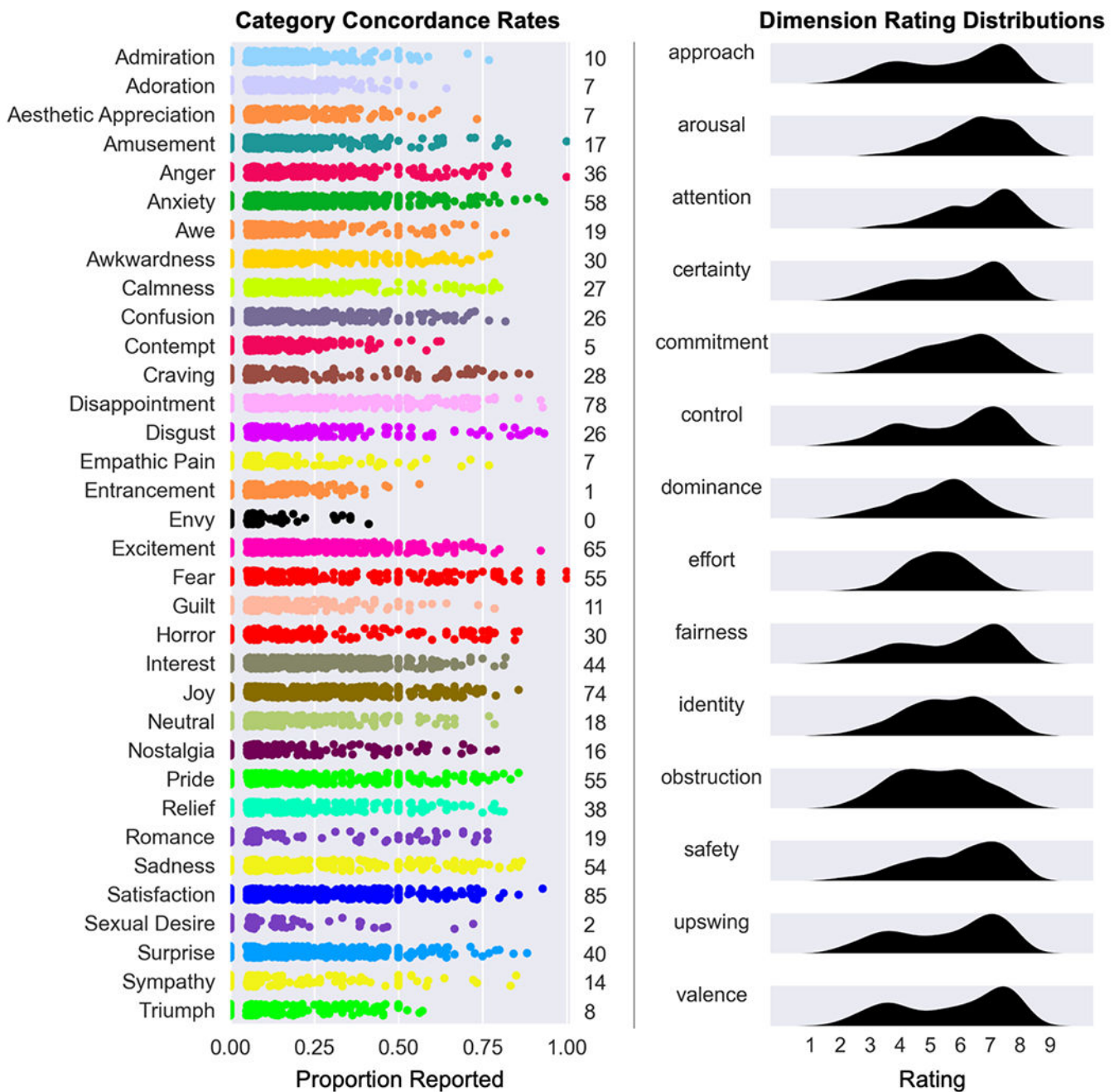
- Cordaro DT, Sun R, Keltner D, Kamble S, Huddar N, & McNeil G (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion*, 18(1). 10.1037/emo0000302
- Cowen AS, Elfenbein HA, Laukka P, & Keltner D (2019). Mapping 24 emotions conveyed by brief human vocalization. *American Psychologist*, 74(6). 10.1037/amp0000399
- Cowen AS, Fang X, Sauter D, & Keltner D (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences of the United States of America*, 117(4). 10.1073/pnas.1910704117
- Cowen AS, & Keltner D (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences of the United States of America*, 114(38). 10.1073/pnas.1702247114
- Cowen AS, & Keltner D (2019). What the Face Displays: Mapping 28 Emotions Conveyed by Naturalistic Expression. *American Psychologist*. 10.1037/amp0000488
- Cowen AS, & Keltner D (2020a). Reply to bowling: How specific emotions are primary in subjective experience. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 117, Issue 18). 10.1073/pnas.2003626117
- Cowen AS, & Keltner D (2020b). Universal facial expressions uncovered in art of the ancient Americas: A computational approach. *Science Advances*, 6(34). 10.1126/sciadv.abb1005
- Cowen AS, & Keltner D (2021). Semantic Space Theory: A Computational Approach to Emotion. In *Trends in Cognitive Sciences* (Vol. 25, Issue 2). 10.1016/j.tics.2020.11.004
- Cowen AS, Laukka P, Elfenbein HA, Liu R, & Keltner D (2019). The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nature Human Behaviour*, 3(4). 10.1038/s41562-019-0533-6
- D'Argembeau A, Xue G, Lu ZL, Van der Linden M, & Bechara A (2008). Neural correlates of envisioning emotional events in the near and far future. *NeuroImage*, 40(1). 10.1016/j.neuroimage.2007.11.025
- Ekman P (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6(3–4). 10.1080/02699939208411068
- Ekman P, Sorenson ER, & Friesen WV (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875). 10.1126/science.164.3875.86
- Ellsworth PC (2013). Appraisal theory: Old and new questions. *Emotion Review*, 5(2). 10.1177/1754073912463617
- Faul L, Baumann M, & LaBar K (2022, April 13). The representation of emotional experience from imagined scenarios. 10.17605/OSF.IO/Y8RW7
- Faul L, St. Jacques PL, DeRosa JT, Parikh N, & De Brigard F (2020). Differential contribution of anterior and posterior midline regions during mental simulation of counterfactual and perspective shifts in autobiographical memories. *NeuroImage*, 215. 10.1016/j.neuroimage.2020.116843
- Fields EC, & Kuperberg GR (2012). It's All About You: An ERP Study of Emotion and Self-Relevance in Discourse. *NeuroImage*, 62(1), 562–574. 10.1016/j.neuroimage.2012.05.003 [PubMed: 22584232]
- Fields EC, & Kuperberg GR (2014). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Social Cognitive and Affective Neuroscience*, 10(9). 10.1093/scan/nsv004
- Fields EC, & Kuperberg GR (2016). Dynamic effects of self-relevance and task on the neural processing of emotional words in context. *Frontiers in Psychology*, 6(JAN). 10.3389/fpsyg.2015.02003
- Fields EC, Weber K, Stillerman B, Delaney-Busch N, & Kuperberg GR (2019). Functional MRI reveals evidence of a self-positivity bias in the medial prefrontal cortex during the comprehension of social vignettes. *Social Cognitive and Affective Neuroscience*, 14(6). 10.1093/scan/nsz035
- Friendly M, & Fox J (2021). candisc: Visualizing Generalized Canonical Discriminant and Canonical Correlation Analysis. R package version 0.8-6.
- Gaddy MA, & Ingram RE (2014). A meta-analytic review of mood-congruent implicit memory in depressed mood. *Clinical Psychology Review*, 34(5), 402–416. 10.1016/j.cpr.2014.06.001 [PubMed: 24980699]

- Ganis G, Thompson WL, & Kosslyn SM (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, 20(2). 10.1016/j.cogbrainres.2004.02.012
- Giordano BL, Whiting C, Kriegeskorte N, Kotz SA, Gross J, & Belin P (2021). The representational dynamics of perceived voice emotions evolve from categories to dimensions. *Nature Human Behaviour*, 5(9). 10.1038/s41562-021-01073-0
- Giuliani NR, & Berkman ET (2015). Craving Is an Affective State and Its Regulation Can Be Understood in Terms of the Extended Process Model of Emotion Regulation. In *Psychological Inquiry* (Vol. 26, Issue 1). 10.1080/1047840X.2015.955072
- Herbert C, Herbert BM, Ethofer T, & Pauli P (2011). His or mine? The time course of self-other discrimination in emotion processing. *Social Neuroscience*, 6(3). 10.1080/17470919.2010.523543
- Holmes EA, Lang TJ, & Shah DM (2009). Developing Interpretation Bias Modification as a “Cognitive Vaccine” for Depressed Mood: Imagining Positive Events Makes You Feel Better Than Thinking About Them Verbally. *Journal of Abnormal Psychology*, 118(1). 10.1037/a0012590
- Holmes EA, & Mathews A (2005). Mental imagery and emotion: A special relationship? *Emotion*, 5(4). 10.1037/1528-3542.5A489
- Holmes EA, & Mathews A (2010). Mental imagery in emotion and emotional disorders. In *Clinical Psychology Review* (Vol. 30, Issue 3). 10.1016/j.cpr.2010.01.001
- Holmes EA, Mathews A, Dalgleish T, & Mackintosh B (2006). Positive Interpretation Training: Effects of Mental Imagery Versus Verbal Training on Positive Mood. *Behavior Therapy*, 37(3). 10.1016/j.beth.2006.02.002
- Horikawa T, Cowen AS, Keltner D, & Kamitani Y (2020). The Neural Representation of Visually Evoked Emotion Is High-Dimensional, Categorical, and Distributed across Transmodal Brain Regions. *IScience*, 23(5). 10.1016/j.isci.2020.101060
- Hotelling H (1935). The most predictable criterion. *Journal of Educational Psychology*, 26(2). 10.1037/h0058165
- Ji JL, Heyes SB, MacLeod C, & Holmes EA (2016). Emotional Mental Imagery as Simulation of Reality: Fear and Beyond—A Tribute to Peter Lang. In *Behavior Therapy* (Vol. 47, Issue 5). 10.1016/j.beth.2015.11.004
- Joseph DL, Chan MY, Heintzelman SJ, Tay L, Diener E, & Scotney VS (2020). The Manipulation of Affect: A Meta-Analysis of Affect Induction Procedures. *Psychological Bulletin*. 10.1037/bul0000224
- Kassam KS, Markey AR, Cherkassky VL, Loewenstein G, & Just MA (2013). Identifying Emotions on the Basis of Neural Activation. *PLoS ONE*, 8(6). 10.1371/journal.pone.0066032
- Keltner D, Sauter D, Tracy J, & Cowen A (2019). Emotional Expression: Advances in Basic Emotion Theory. In *Journal of Nonverbal Behavior* (Vol. 43, Issue 2). 10.1007/s10919-019-00293-3
- Keltner D, Tracy JL, Sauter D, & Cowen A (2019). What Basic Emotion Theory Really Says for the Twenty-First Century Study of Emotion. In *Journal of Nonverbal Behavior* (Vol. 43, Issue 2). 10.1007/s10919-019-00298-y
- Kragel PA, & LaBar KS (2013). Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. *Emotion*, 13(4). 10.1037/a0031820
- Kragel PA, & LaBar KS (2016). Decoding the Nature of Emotion in the Brain. *Trends in Cognitive Sciences*, 20(6), 444–455. 10.1016/j.tics.2016.03.011 [PubMed: 27133227]
- Krijthe JH (2015). Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation.
- Lang PJ (1979). A Bio-Informational Theory of Emotional Imagery. *Psychophysiology*, 16(6). 10.1111/j.1469-8986.1979.tb01511.x
- LeDoux JE, & Hofmann SG (2018). The subjective experience of emotion: a fearful view. In *Current Opinion in Behavioral Sciences* (Vol. 19). 10.1016/j.cobeha.2017.09.011
- Lexico. (2019). Definitions, Meanings, Synonyms, and Grammar by Oxford Dictionary on Lexico.com. In *Lexico Dictionaries* { \textbar } English.
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, & Barrett LF (2012). The brain basis of emotion: A meta-analytic review. In *Behavioral and Brain Sciences* (Vol. 35, Issue 3). 10.1017/S0140525X11000446

- Liu G, Cui Z, Yu H, Rotshtein P, Zhao F, Wang H, Peng K, & Sui J (2020). Neural responses to intention and benefit appraisal are critical in distinguishing gratitude and joy. *Scientific Reports*, 10(1). 10.1038/s41598-020-64720-y
- Liu M, Duan Y, Ince RAA, Chen C, Garrod OGB, Schyns PG, & Jack RE (2022). Facial expressions elicit multiplexed perceptions of emotion categories and dimensions. *Current Biology*, 32(1). 10.1016/j.cub.2021.10.035
- Mathews A, Ridgeway V, & Holmes EA (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, 27(2). 10.1080/02699931.2012.698252
- Matsumoto et al. (2008). Handbook of Emotions - Chapter 13 - Facial Expressions of Emotion. In The Guilford Press.
- Matt GE, Vázquez C, & Campbell WK (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, 12(2), 227–255. 10.1016/0272-7358(92)90116-P
- Miller RS, & Tangney JP (1994). Differentiating Embarrassment and Shame. *Journal of Social and Clinical Psychology*, 13(3). 10.1521/jscp.1994.13.3.273
- Mobbs D, Marchant JL, Hassabis D, Seymour B, Tan G, Gray M, Petrovic P, Dolan RJ, & Frith CD (2009). From Threat to Fear : The neural organization of defensive fear systems in humans. 29(39), 12236–12243. 10.1523/JNEUROSCI.2378-09.2009.From
- Moors A, Ellsworth PC, Scherer KR, & Frijda NH (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, 5(2). 10.1177/1754073912468165
- Morey RA, McCarthy G, Selgrade ES, Seth S, Nasser JD, & LaBar KS (2012). Neural systems for guilt from actions affecting self versus others. *Neuroimage*, 60(1). 10.1016/j.neuroimage.2011.12.069
- Moulton ST, & Kosslyn SM (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521). 10.1098/rstb.2008.0314
- Murphy FC, Nimmo-Smith I, & Lawrence AD (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective and Behavioral Neuroscience*, 3(3). 10.3758/CABN.3.3.207
- Northoff G, Heinzel A, de Greek M, Bermpohl F, Dobrowolny H, & Panksepp J (2006) Self-referential processing in our brain-A meta-analysis of imaging studies on the self. *Neuroimage*, 31(1). 10.1016/j.neuroimage.2005.12.002
- Nummenmaa L, & Saarimäki H (2019). Emotions as discrete patterns of systemic activity. *Neuroscience Letters*, 693, 3–8. 10.1016/j.neulet.2017.07.012 [PubMed: 28705730]
- Peters ML, Flink IK, Boersma K, & Linton SJ (2010). Manipulating optimism: Can imagining a best possible self be used to increase positive future expectancies? *Journal of Positive Psychology*, 5(3). 10.1080/17439761003790963
- Phan KL, Wager T, Taylor SF, & Liberzon I (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. In *NeuroImage* (Vol. 16, Issue2). 10.1006/nimg.2002.1087
- Posner J, Russell JA, & Peterson BS (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3). 10.1017/S0954579405050340
- R Core Team. (2021). R: A Language and Environment for Statistical Computing. In R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rasmussen AS, & Berntsen D (2013). The reality of the past versus the ideality of the future: Emotional valence and functional differences between past and future mental time travel. *Memory and Cognition*, 41(2). 10.3758/s13421-012-0260-y
- Robinson MD, & Clore GL (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, 27(11). 10.1177/01461672012711012
- Roese NJ, & Epstude K (2017). The Functional Theory of Counterfactual Thinking: New Evidence, New Challenges, New Insights. In *Advances in Experimental Social Psychology* (Vol. 56, pp. 1–79). 10.1016/bs.aesp.2017.02.001

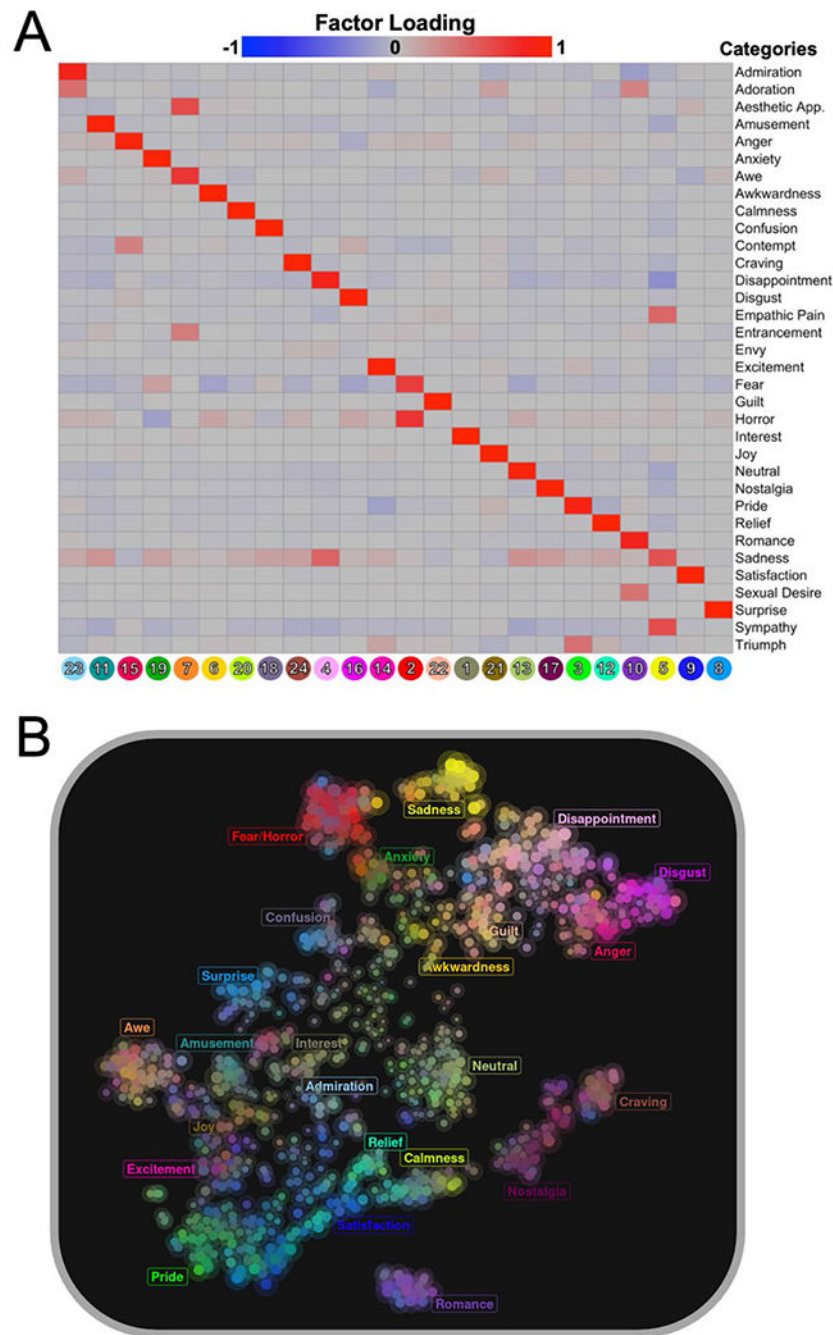


- Russell JA (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 145–172. 10.1037/0033-295X.n0.U45 [PubMed: 12529060]
- Saarimäki H, Gotsopoulos A, Jääskeläinen IP, Lampinen J, Vuilleumier P, Hari R, Sams M, & Nummenmaa L (2016). Discrete Neural Signatures of Basic Emotions. *Cerebral Cortex*, 26(6). 10.1093/cercor/bhv086
- Schacter DL, & Addis DR (2007). On the constructive episodic simulation of past and future events. In *Behavioral and Brain Sciences* (Vol. 30, Issue3). 10.1017/S0140525X07002178
- Sedikides C, & Wildschut T (2018). Finding meaning in Nostalgia. *Review of General Psychology*, 22(1), 48–61. 10.1037/gpr0000109
- Siedlecka E, & Denson TF (2019). Experimental Methods for Inducing Basic Emotions: A Qualitative Review. *Emotion Review*, 11(1), 87–97. 10.1177/1754073917749016
- Siegel EH, Sands MK, Van den Noortgate W, Condon P, Chang Y, Dy J, Quigley KS, & Barrett LF (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, 144(4). 10.1037/bul0000128
- Simon-Thomas ER, Keltner DJ, Sauter D, Sinicropi-Yao L, & Abramson A (2009). The Voice Conveys Specific Emotions: Evidence From Vocal Burst Displays. *Emotion*, 9(6). 10.1037/a0017810
- Takahashi H, Matsuura M, Koeda M, Yahata N, Suhara T, Kato M, & Okubo Y (2008). Brain activations during judgments of positive self-conscious emotion and positive basic emotion: Pride and joy. *Cerebral Cortex*, 18(4). 10.1093/cercor/bhm120
- Takahashi H, Yahata N, Koeda M, Matsuda T, Asai K, & Okubo Y (2004). Brain activation associated with evaluative processes of guilt and embarrassment: An fMRI study. *NeuroImage*, 23(3). 10.1016/j.neuroimage.2004.07.054
- Tangney JP (1995). Recent Advances in the Empirical Study of Shame and Guilt. *American Behavioral Scientist*, 38(8). 10.1177/0002764295038008008
- Thompson B (1984). *Canonical Correlation Analysis*. SAGE Publications, Inc. 10.4135/9781412983570
- Tracy JL, & Robins RW (2007). Emerging insights into the nature and function of pride. *Current Directions in Psychological Science*, 16(3). 10.1111/j.1467-8721.2007.00493.x
- Ugazio G, Lamm C, & Singer T (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion*, 12(3). 10.1037/a0024611
- Wilson-Mendenhall CD, Barrett LF, & Barsalou LW (2013). Neural Evidence That Human Emotions Share Core Affective Properties. *Psychological Science*, 24(6). 10.1177/0956797612464242
- Zhou H, Guo J, Ma X, Zhang M, Liu L, Feng L, Yang J, Wang Z, Wang G, & Zhong N (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience*, 11. 10.3389/fnhum.2017.00451
- Zinck A (2008). Self-referential emotions. *Consciousness and Cognition*, 17(2). 10.1016/j.concog.2008.03.014



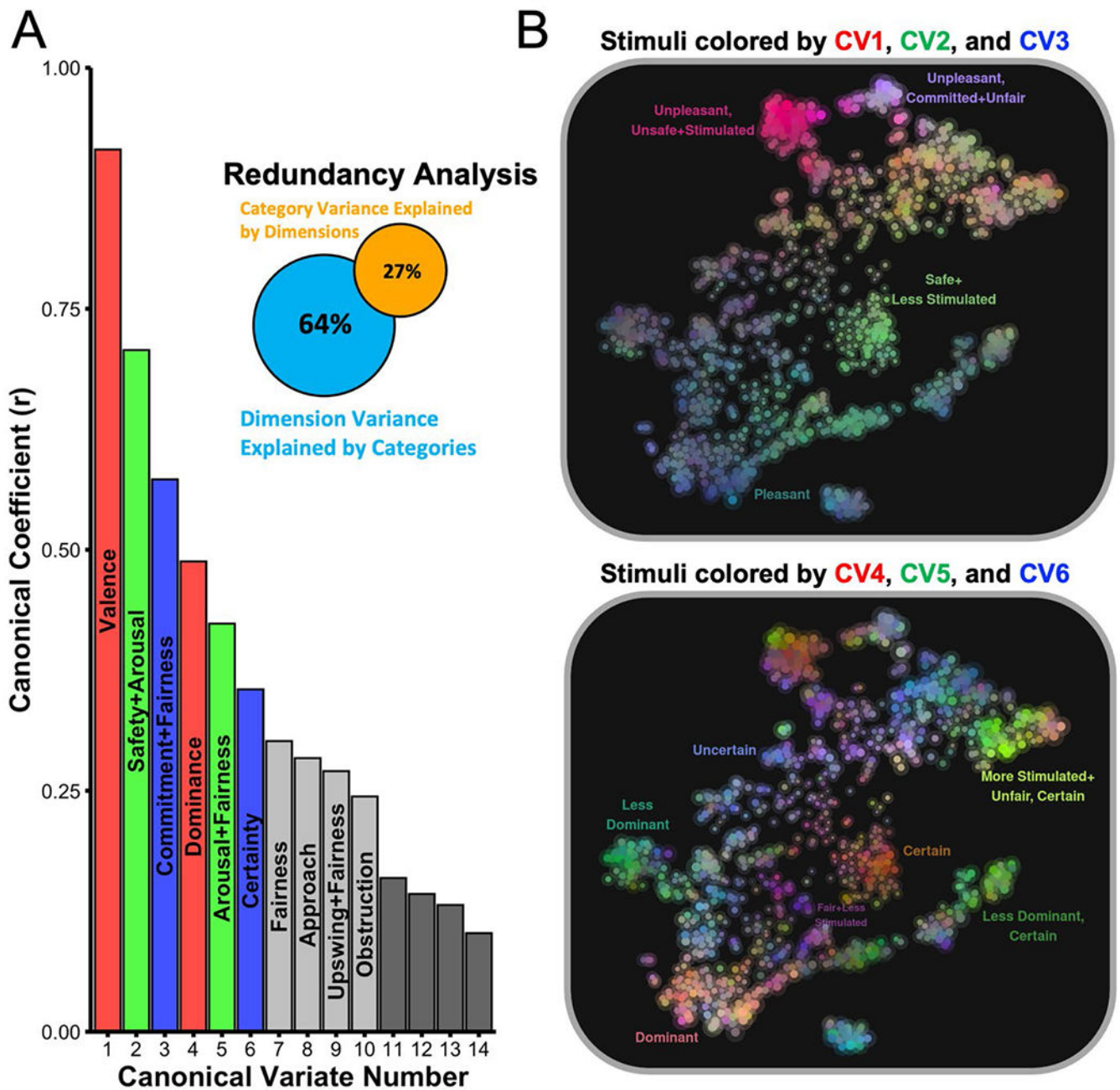
**Figure 1.** Category judgements and dimensional ratings for all 883 scenarios. *Note.* Category concordance rates (left) represent the proportion of participants that chose a given category for an imagined scenario. For each category, all 883 scenarios are represented by different colored circles, the location of which indicates the proportion of participants that chose that category for that scenario. To facilitate transparency, we report to the right of the plot the number of stimuli with concordance rates equal to or above 50% (reliably agreed upon by at least half of raters), instead of using a chance threshold determined via simulation analyses (Cowen & Keltner, 2017). Dimensional distributions

(right) are provided as density plots representing average ratings for each scenario across all dimensions.



**Figure 2.** (A) Principal components analysis (PCA) of the categorical data, followed by varimax rotation of the first 24 components explaining the most variance. (B) The loadings of each stimulus onto the 24 components identified via PCA were mapped into a lower-dimensional space using t-distributed stochastic neighbor embedding (t-SNE). *Note.* (A) Components identified via PCA represent emotional responses that are reliably separate in meaning, and most components consist of loadings from one or two categories. Numbers on the horizontal axis indicate the component numbers after varimax

rotation, which were subsequently assigned categorical labels based on the category that predominantly loaded onto each component. Colors on the heatmap represent the strength of factor loadings for each category onto each principal component, with red indicating positive loadings and blue indicating negative loadings. (*B*) t-SNE is a nonlinear dimensionality reduction technique that seeks to model similar stimuli as nearby points, while separating dissimilar stimuli by larger but more approximate distances. Each point on the chromatic map represents a different stimulus, with size of the points based on the loading of that stimulus onto its primary component. Colors are also based on the varimax-rotated PCA loadings. Specifically, each component was assigned a unique color. Then, for each stimulus, we identified which component(s) that stimulus positively loaded onto and created a weighted blend of colors based on those loadings. In this way, colors represent a gradual change in emotional experience. The resulting chromatic map shows a similar pattern to that observed in other evaluations of semantic space, such that clusters are not fully distinct but rather bridged by continuous gradients of semantic meaning. A larger, interactive map is available at <https://labarlab.shinyapps.io/scenarios/>.

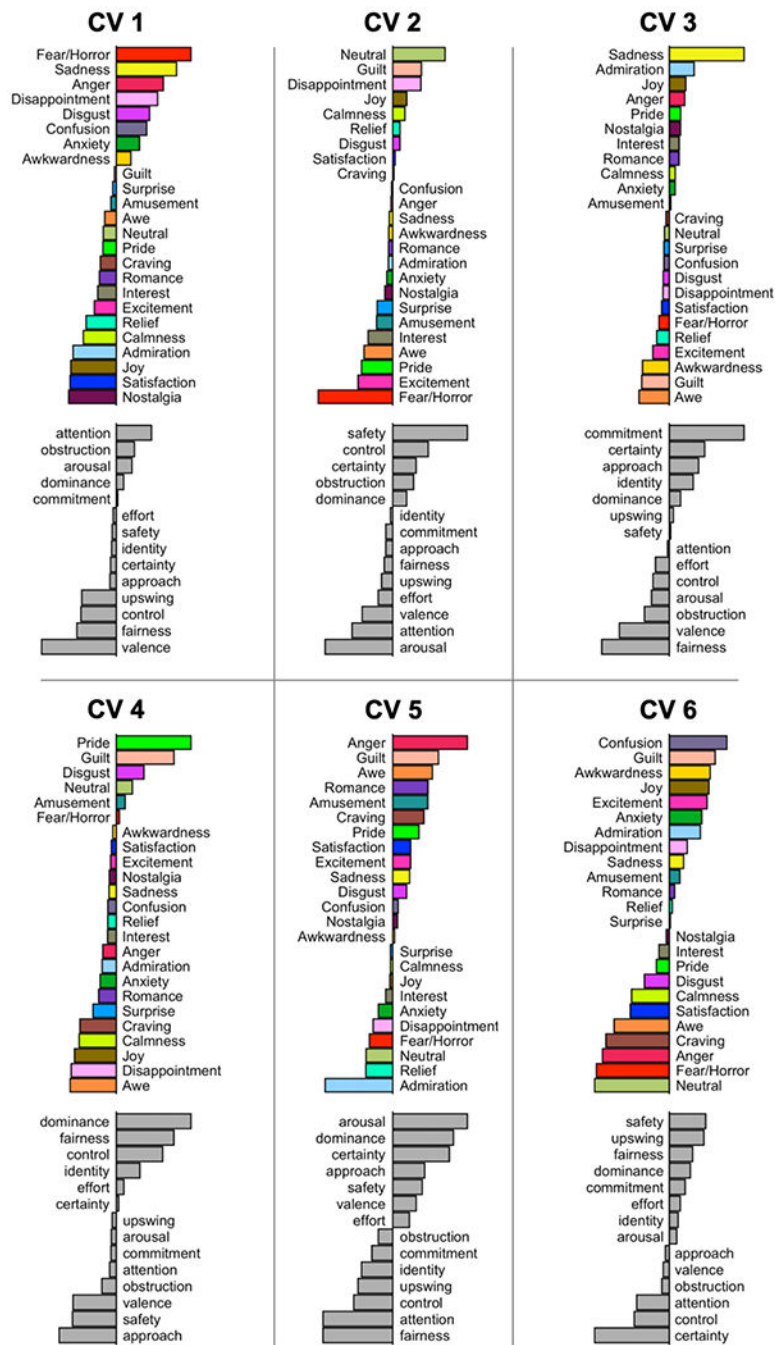


**Figure 3.**

(A) Canonical correlations between the categorical and dimensional ratings and (B) distribution of the first six canonical variate loadings.

*Note.* (A) Ten canonical variates exhibited significant correlations, suggesting that these latent factors unify variance in affective meaning among the two datasets. Bars are labeled according to the primary dimension(s) associated with each canonical variate (for those that were significant). Colors correspond to the red, blue, green (RGB) color channel assignment in the chromatic maps to the right. (B) The same t-SNE plot from Figure 2 was plotted such that the color of each stimulus (RGB) corresponds to the loading of that stimulus onto

each of three canonical variates. The top chromatic map bases the coloring on the first three canonical variates, while the bottom chromatic map bases the coloring on the next three canonical variates. Interactive maps are available at <https://labarlab.shinyapps.io/scenarios/>.



**Figure 4.** Coefficients of all category components and dimensional attributes from the first six significant canonical variates.  
*Note.* For each set of coefficients, the top bar plot shows weights from the categorical components on their canonical variate, while the bottom bar plot shows weights from the dimensions on their canonical variate. Weights represent the unique contribution of variables onto each canonical variate. Values to the right of each plot indicate positive weights, while values to the left indicate negative weights. For example, valence negatively associates with



the first canonical variate, suggesting that this variate generally represents unpleasantness. The range of the axis for each canonical variate bar plot is constrained by the maximum weight of a categorical component or dimensional attribute on that variate. Color assignment in the categorical bar plots reflect the same color assignments to each principal component shown in Figure 2. Each canonical variate explains additional variance in shared affective meaning between category and dimensional ratings that was unaccounted for by previous canonical variates. Coefficients for canonical variates 7-10 are provided in supplemental materials (Figure S7).