Developing an Understanding of Emotion Categories: Lessons from Objects

Katie Hoemann1*, Rachel Wu2*, Vanessa LoBue3*, Lisa M. Oakes4*, Fei Xu5*, & Lisa Feldman Barrett1,6,7*

1Department of Psychology, Northeastern University, Boston, MA
2Department of Psychology, UC Riverside
3Department of Psychology, Rutgers University, Newark, NJ
4Department of Psychology, UC Davis, Davis, CA
5Department of Psychology, UC Berkeley, Berkeley, CA
6Department of Psychiatry, Massachusetts General Hospital, Boston, MA
7Martinos Center for Biomedical Imaging, Charlestown, MA

*equal authorship

Correspondence: rachelw@ucr.edu (R. Wu) and l.barrett@northeastern.edu (L. F. Barrett).

Keywords: object categorization, emotional development, emotion concepts, abstract categories
Abstract

How and when infants and young children begin to develop emotion categories is not yet well understood. Research has largely treated the learning problem as one of identifying perceptual similarities among exemplars (typically posed, stereotyped facial configurations). However, recent meta-analyses and reviews converge to suggest that emotion categories are abstract, involving high dimensional, situationally variable instances. In this paper, we consult research on the development of abstract object categorization to guide hypotheses about how infants might learn abstract emotion categories, as the two domains present infants with similar learning challenges. In particular, we consider how a developmental cascades framework offers opportunities to understand how and when young children develop emotion categories.
Categorization: Introducing the Learning Problem

Categorization (see Glossary) is the process of grouping objects or events as similar [1]. The ability to form categories is pervasive in the animal kingdom [2] and facilitates the most basic daily tasks (e.g., identifying objects as edible versus non-edible). Perceptual categories have instances, or exemplars, which are similar in their observable physical (i.e., perceptual) features. For example, apples tend to be round, small enough to grasp in the palm of a human hand, contain light colored flesh, crunch when you bite into them, and raise your blood sugar upon eating. Other categories are referred to as abstract because a perceiver transcends the exemplars’ perceptual dissimilarities to infer their functional similarity in a given situation or context (see Box 1). For example, ‘food’ is an abstract category, because the distinction between edible and non-edible is based on the function of satisfying hunger in a culturally appropriate way (e.g., grasshoppers are eaten as food in some cultures but killed as pests in others). Moreover, the same object or event can be categorized in a flexible, situated manner: a bright yellow dandelion with green leaves might be considered food (i.e., in a salad), a weed (e.g., in the garden to be plucked and thrown away), or a flower (e.g., in a bouquet of wildflowers), depending on the context [3-7]. Even exemplars similar in perceptual features can be categorized in an abstract way (e.g., apples can be grouped together in different ways depending on whether their function is for snacking, such as Fuji or Gala apples, for baking a pie, such as Braeburn or Granny Smith apples, or for target practice, such as anything laying around the yard). There is a large, robust program of research which has established that, from an early age, infants and children learn to infer functional features to form abstract categories [e.g., 8, 9, 10]. The need to create categories that go beyond perceptual features can arise in any domain and is a fundamental human capacity.

In this paper, we build on recent work to hypothesize that a basic aspect of human life – emotion categories and their associated concepts, such as anger, sadness, fear, etc. – are fundamentally abstract in nature [e.g., 4, 11]. Correspondingly, we propose that emotional
development depends on a young learner’s growing capacity to create abstract categories that transcend the perceptual features of physical movements and signals to infer their functional similarity (e.g., a scowl, a wide-eyed gasping face, and a smirk might be functional for signaling threat in anger in different situations; heart rate might increase, decrease or remain unchanged in fear). We survey the research findings on the development of abstract object categories for insights on how emotional development might proceed. Instances of emotions are events, not objects, of course, but the learning problems are similar.

We begin the paper by briefly summarizing the research evidence suggesting that an emotion, such as anger, is an abstract rather than a perceptual category, because exemplars do not share a common facial configuration, pattern of physiological changes, or other physical features. We next consider what is known about the development of abstract object categorization, and introduce the notion of developmental cascades [12] to understand how and when young learners acquire the capacity to infer functional similarities in a flexible, situated way. We then craft a framework for guiding hypothesis formation about how young learners acquire the capacity to create abstract emotion categories.

Our proposal – that the capacities to make and use emotion categories and their associated concepts are acquired in infancy via a series of developmental cascades – builds on our prior suggestions that emotional development rests on a young learner’s ability to infer psychological similarities, rather than learn concrete perceptual categories which are then later elaborated [4, 11, 13-15]. For example, a cascades approach suggests that simultaneous, interactive developmental processes create categories full of variation during emotional development, predicting a priori that the instances of a given emotion category, such as anger, will contain considerable variation in their physical, perceptual, affective, and even in their functional features (see Box 2). This broad hypothesis suggests a new set of research questions to guide experimental inquiry, beyond those derived from accounts of
category learning as a perceptual-to-conceptual shift [16-18], which also do not adequately capture how and when infants learn object categories. Testing hypotheses that infants learn emotion categories characterized by substantial within-category variation, in turn, requires innovating the experimental strategies that are currently employed to study emotional development; current approaches mis-specify the learning problem as one of identifying perceptual similarities among highly stereotyped instances of an emotion category (e.g., scowls in anger and smiles in happiness). Instead, the problem may involve learning how and when to transcend perceptual features and infer a functional similarity among exemplars. A cascades approach has the further benefit of unifying research on cognitive and emotional development which have been largely kept separate and therefore unable to inform one another. Our approach also has the potential to unify the study of emotion across infants and adults, enriching each in important ways. For example, most published research remains polarized as to the nature of emotion [19], and research on the nature of emotional development is necessary to help resolve this debate.

**What is an Emotion Category?**

Consider all the things you do when you are angry: you might tremble, freeze, scream, withdraw, attack, cry, and even laugh or joke. The physiological changes in your body will be tied to the metabolic demands that support your actions in a given situation (e.g., cardiac output increases when you are about to run, but not when you freeze and are vigilant for more information to resolve uncertainty or ambiguity; [20]). Sometimes you might feel pleasant but other times unpleasant [21]. Recent meta-analyses and reviews indicate that instances of anger, like the instances of other emotion categories, vary considerably in their associated physiological changes [22], facial movements [23], and neural correlates, whether measured at the level of individual neurons [24, 25], as activity in specific brain regions [26], or as distributed patterns of activity [25, 27]. Instances of an emotion category can vary in their affective features (e.g., some instances of fear can feel pleasant, and some instances of
happiness can feel unpleasant; [28, 29]). Instances of different emotion categories can also be similar in a range of features, which is not surprising: sometimes you might smile when you are sad, cry when you are afraid, or scream when you are happy.

To deal with this feature variation, scientists have moved away from the idea that each emotion category is defined by a set of necessary and sufficient features to propose that emotion categories each have a most typical or frequent instance (a prototype) which possesses a common set of features, while other category instances are graded in their similarity to the prototype [e.g., 30, 31, 32] (see Figure 1). Within-category feature variation around the prototype is usually explained, post hoc, by hypothesizing phenomena that are independent of the emotional event itself, such as display rules, regulation strategies, or stochastic variation [for a discussion, see 23]. Accordingly, a prototype view of emotion categories assumes that the category’s prototype has perceptual features that are valid cues in many, but not all, circumstances (see Box 1). The empirical evidence suggests otherwise, however, demonstrating substantial within-category variation, even in studies that are designed to identify the presumed prototypes.

In our view, the magnitude of observed variation in emotional responding is more consistent with the hypothesis that emotion categories are abstract categories, whose instances are functionally (but not always perceptually) similar to one another [3, 4, 13, 15, 33-35]. This approach predicts that substantial within-category variation is intrinsic to the nature of emotion, such that both perceptual features and functional features of a category are situated [3, 4, 13, 15, 35, 36]. In other words, emotion categories are ad hoc abstract categories whose function varies with the situation [6, 7] (see Box 1). Consider, for example, the category for anger: in situations involving a competition or negotiation, the anger category might be constructed such that instances share the functional goal ‘to win’ [37]; in situations of threat, the anger category might cohere around the functional goal ‘to be effective’ [38] or even ‘to appear powerful’ [39]; and in situations involving coordinated action, the anger category
might include instances that share the functional goal ‘to be part of a group’. We hypothesize that individuals learn to construct situation-specific categories based on what is considered most functional in their immediate cultural context; we do not expect a single, core function associated with anger or any other emotion. From the population of available instances that are designated as emotional by other people, infants must learn which features to foreground and which to background when creating an emotion category in a particular situation or context. Critically, instances of that category include their corresponding context: emotional episodes are always high-dimensional, situated events that cannot be signified by a single facial configuration or bodily change.

Abstract categories are learned despite a wide range of variation in the features of their instances. Although it might seem impossible or extremely difficult to learn categories under conditions of variation, children do this with ease. For example, in the domain of language, every instance of a speech sound (e.g., “ball”) varies in pitch, voice onset time, speed, and a large number of other acoustical features. During language development, infants learn to infer category boundaries so that they can treat variable acoustical instances as the same sound [41] or word [42]. In some cases, infants actually learn better in the face of variation [42, 43], suggesting that learning an abstract, functional regularity (e.g., a function of anger, such as appearing powerful, removing a goal, requesting support, etc.) might actually be facilitated by variation in other features. Correspondingly, we hypothesize that an emotion category is learned despite and perhaps even because of variable patterns of facial movements, vocal cues, actions, sensory data from the individual’s body and the surrounding external context, etc., on the one hand, and predictable functional outcomes in a given situation, on the other.

**Learning Emotion Categories: A Brief Summary of Research Findings**

Much is known about children’s explicit knowledge of abstract emotion categories, which develops during the preschool years and continues to develop until the middle school years when
emotion categories appear more adult-like [for reviews, see 44, 45]. Most of what we know about the development of emotion categories, however, comes from studies of infants’ attention to highly stereotypical, posed facial configurations (e.g., wide-eyed, gasping faces to depict fear; Box 3). Findings demonstrate that infants as young as 4 to 5 months discriminate between stereotypic facial expressions, such as an exaggerated smile and frown posed by the same model [46-48], and see similarity in a particular stereotyped expression posed by multiple models or multiple instances of the same model [49-51]. Notably, this research has ignored the real-world variability in facial movements and other features that infants and young children perceive and experience during episodes of emotion [for reviews, see 11, 23]. As such, the conclusion that emotion categories develop early in infancy as perceptual categories may be an artifact of the limited, lab-based way in which they have been investigated.

The conclusion that infants have perceptual categories for specific emotions is also undermined by alternative explanations. For example, infants can discriminate stereotyped facial configurations (i.e., smiles as the expression of happiness, scowls as the expression of anger, and so on) on the basis of physical features alone, leaving it unclear whether they actually understand the emotional meaning those faces signify. For example, when 4- to 10-month-old infants discriminate smiling and scowling, they may do so by the presence or absence of teeth in a photograph rather than by inferring anger and happiness, per se [52]. Discriminating between narrowed or widened eyes, or between showing teeth or no teeth, is not the same as inferring what those features might predict about how a person feels or what a person might do next. Similarly, individual exemplars (i.e., photos of posed, stereotypic facial configurations) can be discriminated on the basis of other psychological features that are rarely ruled out as alternative explanations. For example, a smile and a frown can be distinguished by a single affective feature, such as valence [e.g., 53], and a scowl and a frown can be distinguished by the degree of arousal they portray [e.g., 54]. People suffering from semantic dementia
can discriminate valence even as they are unable to infer more specific emotional meanings in faces [55]. Correspondingly, the fact that an infant can discriminate a face that is portraying a pleasant state from one that is portraying an unpleasant state is not equivalent to understanding the difference between happiness and sadness. Often, the exemplars can also be discriminated by novelty rather than on their emotional meaning. For example, the wide-eyed expressive stereotype for fear is rarely seen [56], as is the nose-wrinkled expressive stereotype for disgust [57]. Therefore these expressions will be novel for infants [e.g., 58]. For a specific example of these issues related to fear categorization, see Box 4.

There is also evidence that infants learn to group together instances with different features to infer a similar function or goal in a particular situation. Between 12 and 18 months, infants begin to associate affective features with goals or other functions. At this time, infants avoid novel objects or obstacles when an adult or parent responds with negative facial movements and/or vocalizations (i.e., social referencing; [59, 60]), suggesting that they can use negative affective information from others to plan goal-directed actions. But again, this literature is insufficient, on its own, to indicate whether infants are categorizing for specific emotional meaning, rather than more general affective features such as avoid and approach.

To discover whether abstract emotion knowledge begins to develop in infancy, research on emotional development must move beyond asking when infants categorize highly stereotyped, static images of facial configurations to a broader understanding of how infants learn to deal with varying, observable perceptual features and more inferential, abstract features to create and understand highly variable and context-specific abstract emotion categories.

**The Development of Abstract Object Categorization**

Developmental scientists have a deeper – if incomplete – understanding of how infants deal with the complexity of learning to construct abstract object categories. Over the first 2 years, infants
learn to infer features such as associations and relations (e.g., kitchen versus bathroom items [9]), internal, unobservable properties, such as animacy [61], and roles and functions, such as things that fly [10]. In addition to learning how to infer unobservable features, young learners also figure out when to do so in a particular situation or context [11, 62]. Infants – like adults – group objects together in a flexible, situated (i.e., ad hoc) fashion. They use different similarities to categorize objects depending on the context [63, 64] and determine which similarities are most relevant given the way items are labeled [65]. A timeline sampling important discoveries in the development of abstract object learning is illustrated in Figure 2.

**Learning Abstract Categories via Developmental Cascades**

The development of situated abstract knowledge during infancy and early childhood can be understood as a series of developmental cascades across multiple domains [12, and see 66 for a similar perspective], rather than as a perceptual-to-conceptual shift. Developmental progress in one domain (e.g., 3D object perception) necessarily depends on development in other domains (e.g., motor control), which in turn influences yet other domains (e.g., learning about the functional features of objects via play). We will consider three types of cascades as illustrations.

First, infants’ attention to some abstract object features results from the cascading effects of changes in motor abilities. The emergence of self-sitting and increased sophistication in manual exploration during infancy provide opportunities to learning new features about objects, such as three-dimensionality [67] and figure-ground relations [68]. Although such object features may seem to map onto simple perceptual features, they are in fact highly context dependent. The separation of figure from ground, for example, requires prioritizing specific perceptual features to recognize which forms are objects and which are the background [69]. Thus, increasing motor capacities—and the corresponding changes in infants’ interactions with objects—alters the significance and salience of perceptual
features. Infants’ ability to manipulate objects even allows them to focus on more subtle features that signal abstract or functional commonalities [70].

As a second example, an infants’ emerging linguistic abilities scaffolds their developing object categorization. Even before the emergence of the first word, and throughout the second year, infants’ object categorization is shaped by whether and how objects are labeled with words [71]. Infants construe objects that vary in their perceptual features as belonging to the same category when they are labeled with the same word, and they construe identical objects as belonging to different categories when labeled with different words [e.g., 72]. In addition, infants as young as 6 months appear to recognize the link between common labels and their referent [73]. These emerging abilities to recognize that some objects have a common label certainly influences the development of object categories. Indeed, these linguistic abilities may support infants’ attention to abstract categories that are not signaled by an obvious perceptual similarity. Infants and toddlers can use words as a powerful tool for learning that perceptually dissimilar objects are functionally similar in some way [e.g., 8, 72, 74, 75]. Infants are also only able to recognize novel spatial categories when the relation is labeled [76]. Clearly, infants’ attention to context-specific, abstract object features reflects, in some cases, cascades from their language acquisition. Such findings suggest the hypothesis that infants’ ability to impose functional similarities on perceptually variable instances of emotions may cascade from increasing linguistic ability [4, 11].

As a final example, consider the cascading effect of interactions with social partners on infants’ attention to abstract object features. Social interactions can induce both in-the-moment and longer lasting changes in how infants attend to objects and object features [77-79]. Social partners shape infants’ visual inspection of objects by indicating which objects to prioritize among distractors [78, 79] and by the kind of action they perform on the objects [80, 81]. Moreover, infants’ own actions on objects are determined, in part, by the actions and intentions of social partners [e.g., 82, 83]. These
examples illustrate how infants’ attention to – and learning about – different types of object features reflect cascading effects of social interactions when manipulating objects.

Thus, a cascades framework provides deeper insight into infants’ increasing ability to transcend sometimes very salient perceptual similarities and infer less obvious but situationally relevant functional features. Our understanding of abstract object categories in this framework is relatively new, but nonetheless provides a novel framework to generate novel hypotheses and innovate experimental methods to study infants’ growing capacity to create abstract emotion categories, which are intrinsically dependent on context [e.g., 29, 35].

**Applying a Developmental Cascades Approach to Emotion Category Learning**

We hypothesize that emotional development results from dynamic cascades in other abilities (see Figure 2), which help infants learn when (and when not) to use varying configurations of perceptual features to infer less obvious, situated functions (e.g., when a smile predicts a joyful hug and when it predicts angry aggression). For example, the development of emotion categories, like abstract object categories, may cascade from developmental changes in motor ability. Developing motor abilities provide infants with new information that was previously unavailable, in effect creating the context for learning flexible, ad hoc emotion categories. As infants learn to sit on their own, reach for and manipulate objects, and locomote, facial and body movements, and actions from caregivers might become especially salient as social cues for guiding infants’ actions. Classic work by Campos and colleagues showed that infants’ use of parents’ posed facial expressions to determine whether to traverse a visual cliff was a function of their crawling experience [84]. This pattern suggests a cascade. As infants’ crawling ability emerges and develops, they encounter situations that elicit emotional reactions from parents, and those events provide a context that allow infants to understand how particular features, such as the raise of an eyebrow or the shake of the head, can have different functional meanings (i.e., predict different subsequent actions) in different situations. As another
example, consider how the onset of walking creates a context for developing emotion knowledge. When infants begin to walk independently, there is an increase in their carrying objects to their caregivers in a bid for joint attention [85]. Such bids result in caregivers responding with more action directives [86] and, possibly, providing more information about situation-specific goals and functions that form the basis for emotion categories.

More broadly, all motor movements are accompanied by changes in the bodily systems that support those movements (by a process called allostasis; [87]) as well as the internal sensory data that arises from allostasis (called interoception; [88]). Infants’ developing allostatic abilities may be accompanied by new interoceptive sensations that are, in effect, perceptual features that are available for categorization. We have speculated elsewhere that allostatic development, particularly in the context of loving caregivers, creates another cascade that supports the development of abstract categories, including emotion categories [4, 89].

Similarly, as for object categories, the developmental cascades for emotion categories likely includes language development. During the first year, when infants are developing the ability to discriminate facial configurations [e.g., 49], there are corresponding changes in their ability to use words to link objects into perceptually similar groups [e.g., 75], as well as in their ability to recognize common nouns as labels for familiar items [73]. To the extent that parents and others use emotion words with young infants, these linguistic achievements likely support infants’ emerging ability to learn many-to-many mappings of facial configurations to function. By experiencing that widened eyes with a gaping mouth and narrowed eyes with a wry smile both predict a hug in different situations (which include hearing the same emotion words), infants have the opportunity to learn to infer the same emotional meaning for different facial configurations. Correspondingly, observing the same facial configuration in different situations allows infants to infer different emotional meanings.
At about 12 to 18 months of age – the same time as infants begin to produce their first words [e.g., 90] – they also start to infer other people’s emotions to guide their own actions [59, 60, 84]. This alignment suggests that caregivers’ use of labels referring to emotions and intentions, as well as the infants’ growing understanding of these words, may shape how infants understand and infer similarities across emotional episodes. Emotion word usage may be sparse at first (in particular, directed at the infant; [e.g., 91]), although this is a topic much in need of study. Nonetheless, increases in caregivers’ use of mental state words, and in infants’ vocabulary, may enhance the role that labels play in grouping dissimilar instances. In this way, children’s own language and the action directives used by their caregivers (often embedded in sentences involving emotion labeling) may help shape situated emotion categories over time.

Emotion category development may in fact, be part of the cascade for young children’s emerging ability to infer the mental states of others, known as theory of mind (consistent with [59, 92]). Inferring a function for other people’s facial and body movements, and their vocalizations, as occurs when child is categorizing them for their emotional meaning, is, in effect, a mental inference. Infants begin to show evidence of understanding of others’ intentions between 14 and 18 months of age [e.g., 93, 94], about the same time they start using the emotional reactions of others to guide their own behavior [e.g., 60] – implying that they are able to infer the functional meaning of others’ actions. Some research provides evidence that even young toddlers make emotional inferences about facial configurations [e.g., 45, but see 95]. Thus, the developing ability to make inferences about the mental states of others is likely crucial to the development of adult-like emotion categories, at least in western cultures [96].

**Implications for Future Research**

A cascades approach to understanding emotional development suggests several novel hypotheses about the nature of emotion categories and emotional development. First, emotion
categories and their associated concepts may be learned as abstract and ad hoc from the outset. The variation in facial movements, vocalizations, interoceptive changes, actions and so on, rather than being an obstacle to emotional development, may in fact enhance it by increasing the capacity of those signals to bear emotional information in different situations. Second, the ability to flexibly abstract away from sensory particulars to create ad hoc, functional categories is fundamental to other domains of development, such as spatial categorization [76]. This suggests that emotion categories and concepts are formed via domain-general mechanisms cascading from the development of motor and cognitive processes, such as increasing motor ability, linguistic capacity, and proficiency with mental inference. This second hypothesis, in effect, dissolves the boundary between cognitive and emotional development, allowing debates on the nature of emotion to benefit from research on category learning in children. Both of these hypotheses provide an opportunity to advance emotion research throughout the lifespan.

By advocating for the rigorous study of variation and examination of multiple factors that influence category learning, a developmental cascades approach has the potential to enrich the scientific understanding not only of emotion, but of the mind more generally. The insights of a developmental cascades approach are consistent with constructionist accounts of mind and brain [4, 97], which hypothesize that the experience and perception of emotion are events constructed in the brain by domain-general predictive processes [e.g., 15, 98]. Accordingly, emotion categories and concepts are not distinct from emotional episodes, but may in fact be necessary for constructing experiences and perceptions of emotion. Constructionist accounts have been criticized for not offering plausible hypotheses for emotional development. A cascades approach offers a generative framework for studying the developmental implications of constructionist hypotheses, building on previous accounts of emotional development in important and novel ways [11].
A cascades approach also implicitly suggests, consistent with a constructionist approach, that individual differences in emotional granularity [4, 99] and cultural variation in emotion categories [96] are not moderators of emotional universals, but are instead intrinsic to the nature of emotion and inherently result from the processes that support emotional development. A constructionist approach that depends on developmental cascades does not necessarily imply that there are no human universals, however. A major adaptive advantage of our species is to live in social groups, and as a consequence, all cultures find solutions to common problems of group living, including the capacity to learn categories for motivated action, such as emotion categories. As a consequence, certain categories may be universal even if they are not innate [for discussion, see 3], consistent with evidence about cultural evolution and gene-culture co-evolution [e.g., 100].

Finally, a developmental cascades approach has the potential to suggest novel approaches to understanding emotional disorders. Globally, more than 300 million people of all ages suffer from depression, which is ranked by the WHO as the single largest contributor to disability worldwide [101], and on the rise in adolescents [e.g., 102]. Studying emotion category learning as a dynamic, emergent process has the potential to illuminate critical abilities in developmental trajectories that are necessary for other competencies to develop (e.g., exposure to emotion words invites abstract category formation), offering a more precise way of identifying the various factors that might help treat or prevent mental suffering.

Concluding Remarks

A cascades approach for understanding the development of abstract object categories holds promise for understanding the processes by which infants and young children develop emotion categories, which are fundamentally abstract, as well as flexible and situated. There are many differences between object and emotion categories (emotions are dynamic events, not static objects). But a developmental cascades approach offers important insights and generates new hypotheses for
understanding abstract emotion categorization (see Outstanding Questions). Exploring these and other domains of abstract categorization within a developmental cascades framework can stimulate fruitful lines of inquiry that unite cognitive and emotional development by a common set of developmental principles.
Glossary

**Abstract category**: A collection of objects or events that are grouped based on functional or psychological features rather than observable physical or perceptual features.

**Ad hoc category**: Categories that are constructed, on a situation-by-situation basis, where the similarity among instances is rooted in context-specific goals and functions. Ad hoc construction is related to flexible and situated categorization.

**Allostasis**: The process by which the brain anticipates the needs of the body and attempts to meet those needs before they arise (e.g., increase in blood pressure as a person stands).

**Categorization**: The mental act of grouping a collection of instances, objects, or events according to some set of similarities (and ignoring any differences).

**Category**: A collection of objects or events that are considered similar for a given purpose or function in a given situation.

**Concept**: A mental representation of a category. For classical categories, a concept is a single representation with necessary and sufficient features and remains stable for that category across all situations. The concept for a prototype category is also stable across situations, but it is the instance with the most frequent or most typical features. For ad hoc, abstract categories, the concept is the representation of the exemplar that best meets the function of that category in a given situation. The boundary between concept and category blurs for ad hoc, abstract categories.

**Emotional granularity**: An individual’s ability to experience emotion with specificity and detail, in a way that is precisely tailored to the immediate internal (i.e., bodily) and external context.

**Event category**: A collection of situated, dynamic episodes. Because events are dynamic, they have temporal features, as well as features that demarcate the beginning and end.

**Exemplar**: An object or event serving as an instance (i.e., member) of a category.
**Interoception:** The sensory data that collectively describes the physiological state of the body, arising from the allostatic regulation of various bodily systems, including the autonomic nervous system, the endocrine system, and the immune system.

**Object category:** A collection of real-world, tangible items (or photographs/drawings of items).

**Perceptual category:** A collection of objects or events that are similar in their observable physical or perceptual features.

**Prototype:** The most typical or frequent instance of a category; other category instances are graded in their similarity to the prototype.
Box 1. A Brief History of Categories

A classical category has exemplars that share observable, perceptual features. Its concept is a single representation consisting of a dictionary definition of necessary and sufficient features. The idea that most categories were classical in structure dominated science and philosophy from antiquity but was replaced in the 1970s by the idea of prototype categories, prompted by observations that a category’s instances vary from one another in their features, some of which are more frequent or more typical (meaning the instance has a majority of the category’s features). The category’s concept (its prototype) is the single most representative instance of the category (i.e., the most frequent or most typical instance).

Abstract categories have exemplars that are similar in their inferred functional features rather than their observable perceptual features. Beginning in the early 1980's, the psychologist Larry Barsalou observed that categories are formed in an ad hoc way, based on the function that the category serves. For example, in playful situations, a person might construct the category “things that fly” with balls, Frisbees, kites, and darts; in situations that require travel, the same category might include an airplane, hot air balloon, and helicopter. In a park, the category will contain birds, bats, bees, and squirrels. The concept for an ad hoc, abstract category is the most representative exemplar (i.e., the prototype) that best describes the category’s function in a given situation. The prototype need not exist in nature; it is the ideal instance that illustrates the category’s function in that situation.

In the science of emotion, different theoretical approaches vary in their hypotheses about the nature of emotion categories and concepts (see Figure 1). For example, certain constructionist [31], functional [103], appraisal [104] and basic emotion [30] approaches propose that emotion categories are structured as prototype categories. But the empirical evidence indicates that the exemplars within an emotion category vary more substantially in their features than can easily be accounted for by a prototype account. For example, a recent meta-analysis found weak to no reliability in the facial
movements that adults (living in western cultures) used to express emotions (average reliability was .22 across all categories, ranging from .11 to .35 for specific categories [105]; see [23] for discussion). Variation of similar magnitudes has been observed in other measurable features, as well (see main text), consistent with the hypothesis that emotion categories are ad hoc, abstract categories [e.g., 4, 6].
Box 2. Features of Emotional Episodes

An instance of emotion can be described according to its features. Some features are physical whereas others are psychological. Physical features are changes in an emoter that can be measured objectively (i.e., independently of a human perceiver), such as facial movements, body movements, vocal acoustics, autonomic nervous system changes, neural activity, chemical changes, etc. Physical features also include changes in the environment, including wavelengths of light, vibrations in the air, chemical olfactants, etc. Perceptual features describe how physical features are perceived by a human brain: e.g., brightness, loudness, color, heat, smell, texture, interoceptions, etc. Affective features capture what a given instance of experience feels like [106]. Valence refers to the feeling of pleasure or displeasure. Arousal refers to a feeling of activation or sleepiness. Appraisal features refer to descriptions of how a situation is experienced during an instance of emotion: e.g., whether or not it is novel, goal-conducive, predictable, and so on [107-109]. Functional features are the goals that a person is attempting to meet in a given situation: e.g., to curry favor, to socially affiliate, to avoid harm [e.g., 103, 110]. Temporal features denote the sequence and structure of events that result as the brain segments continuous activity [111]. The representation of event dynamics drives understanding of intentionality and causality [112], and the demarcation of event boundaries is hypothesized to be one key aspect of emotion categorization [113, 114].
The stimuli that are used in typical emotion perception experiments are static, exaggerated facial configurations that depict stereotyped emotional expressions; they were created by having actors pose specific facial configurations. These configurations of facial movements are presumed to occur in everyday life. However, the facial expressions that occur during real-word emotional episodes are neither static nor reliably express a single emotion category [23]. While the phrase “emotional expression” is often used to describe experimental stimuli, this phrase carries different connotations depending on the audience. To a developmental psychologist, “emotional expressions” may refer only to configurations of (stereotyped) perceptual features presented to infant or child participants. In the adult literature, the use of “emotional expressions” often implies that perceivers have more elaborated concepts for emotion, including the subjective experience of emotion and the ability to infer that experience in others. To the lay reader, the phrase “emotional expression” may further imply that the poser was actually experiencing the emotion in question (i.e., that the pose is a veridical reflection of subjective experience). In considering all audiences, in the present paper we adopt the terminology “(emotional) facial configurations”, which we consider the most neutral description.
Box 4. Is the Categorization of Fear Special?

At 7 months, infants begin to behave differentially towards distinct facial configurations, and show an attentional bias towards the wide-eyed gasping face that is the western stereotypical expression of fear (note that this face is the stereotype of threatening someone in Melanesia; [115]). Specifically, infants attend more to the wide-eyed gasping faces when compared to smiling (happy) and neutral faces [116-118]. Some researchers have suggested that this developmental shift indicates that infants appreciate the meaning of a fearful face, that it signals a threat is imminent in the environment. However, it is unclear to what extent other features drive increased looking times toward and difficulty disengaging from wide-eyed gasping faces. This attentional bias holds in comparisons to novel facial configurations that differ in their physical features (e.g., lips closed, cheeks fully blown open; [119]), and even when controlling for the salience of the eyes [e.g., 53]. However, wide-eyed gasping faces can be discriminated from any of these comparisons (and from smiling and neural faces) on the basis of affective features, such as valence. Infants also look longer at wide-eyed gasping faces than at scowling faces [120], which are both negatively valenced but can be discriminated on the basis of novelty (since stereotypic fear faces are rarely seen in real life; [56]). Moreover, any negatively valenced expression may be unfamiliar or novel to most infants in the first 6 months of life, and exposure likely varies based on individual differences and experience. Taken together, findings to date are inconclusive with regard to the meaning of the attentional bias for wide-eyed gasping (fearful) faces that emerges around 7 months.
Acknowledgments

We thank Elizabeth Davis for helpful discussions on ideas in this manuscript. This paper was supported by funding from the U.S. Army Research Institute for the Behavioral and Social Sciences (W911NF-16-1-019) to L. F. Barrett, the National Cancer Institute (U01-CA193632) to L. F. Barrett, the National Institute of Mental Health (R01-MH113234, R01-MH109464) to L. F. Barrett, the National Science Foundation Civil, Mechanical and Manufacturing Innovation (1638234) to L. F. Barrett, the National Eye Institute (R01-EY030127) to L. M. Oakes, and the National Heart, Lung, and Blood Institute (1 F31 HL140943-01) to K. Hoemann.
References


27 Clark-Polner, E., et al. (2017) Multivoxel pattern analysis does not provide evidence to support the existence of basic emotions. *Cerebral Cortex* 27, 1944-1948


76 Casasola, M., et al. (2009) Learning to form a spatial category of tight-fit relations: How experience with a label can give a boost. *Developmental Psychology* 45, 711-723


85 Karasik, L.B., *et al.* (2011) Transition from crawling to walking and infants’ actions with objects and people. *Child Development* 82, 1199-1209


Figure 1. Explanatory frameworks guiding the science of emotion. Theoretical approaches hypothesize the nature of emotion categories and their concepts. Surface similarity: hypotheses about the degree to which instances of an emotion category vary in their observable features. Deep similarity: hypotheses about the similarities in the mechanisms that cause instances of the same emotion category (e.g., the neural circuits or assemblies that cause instances of the same emotion category). The colors represent the type of emotion categories proposed: ad hoc, abstract categories (green zone); prototype or theory-based categories (yellow zone); classical or natural-kind categories (red zone). Adapted with permission from Barrett et al (2019), Psychological Science in the Public Interest. In research on adults, learning to construct abstract emotion categories in an ad hoc, situated manner has been hypothesized as necessary to accurately perceive emotions in other people [55] and experience emotions with any degree of granularity [4, 11, 13]. This hypothesis is based on neuroscientific findings [e.g., 15] and is distinct from functional and prototype hypotheses that emotion categories and concepts are separate from the experience and perception of emotion.

Figure 2. Hypothesized developmental timelines. Core processes (depicted in blue), object categorization (in green), and emotion understanding (in purple). The core processes belong to developmental cascades that might contribute to the development of categorization, including emotion categorization. At each point in time, multiple abilities are emerging; for example, at 9 months, infants may be crawling, are sensitive to contextual cues and functional inferences, and can match affective facial and vocal signals. Each of these reflect development within their own “domain” (e.g., colored row), but they may also reflect potential cascades across rows. For example, the limitations in the newborn visual system, which bias infants to look more at faces than at other stimuli, influence their developing perceptual abilities. Similarly, infants’ increased ability to interact with the world through visual-manual exploration provides them with opportunities to learn new properties of objects (e.g., their weight, how they sound when dropped), as well as to learn facial expressions, vocalizations, and other cues related to emotion during object explorations with caregivers and others. “Bias for stereotyped fear configurations” refers to the wide-eyed gasping face that is the stereotype in western cultures (for discussion, see Box 4).