How do children learn novel emotion words? A study of emotion concept acquisition in preschoolers

Holly Shablack,[1] Misha Becker,[2] Kristen A. Lindquist[1]

University of North Carolina at Chapel Hill

1. Department of Psychology and Neuroscience

2. Department of Linguistics

In press at *Journal of Experimental Psychology: General*

Author Note

Correspondence concerning this article should be addressed to Holly Shablack, Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, 235 E. Cameron Ave, Chapel Hill, NC 27599. Email: hshab@live.unc.edu

This research was supported by a grant from the Fostering Interdisciplinary Research Explorations program at UNC Chapel Hill. Findings were presented at the Society for Affective Science Conference in 2018 and Cognitive Development Society in 2019.

Word count: 10,471

Abstract

Understanding emotion words is vital to understanding, regulating, and communicating one's emotions. Yet, little work examines how emotion words are acquired by children. Previous research in linguistics suggests that children use the sentence frame in which a novel word is presented to home in on the meaning of that word, in conjunction with situational cues from the environment. No research has examined how children integrate these cues to learn the meaning of emotion adjectives (e.g., "happy," "sad," "mad"). We conducted two studies examining the role of sentence frame and situational context in children's (ages 3-5) understanding of the meanings of novel words denoting emotions. In Study 1 (N=135) children viewed a conversation wherein a novel "alien" word was presented in one of three sentence frames that varied in how likely the word was to denote an emotion (i.e., *is daxy, feels daxy,* or *feels daxy about*). Children selected the image that represented the meaning of the word in a picture-pointing task. Images depicted aliens experiencing an emotion, a physical state, or performing an action. In Study 2 (N=113) situational context was added via cartoons depicting an emotional scenario. Findings suggest that children are more likely to associate emotion images with a novel word with increasing age, more informative sentence frames, and when the situational context implies that an emotion is present. This provides important insight on how educational and clinical settings can use language and situational context to aide in emotion understanding.

Keywords: emotion development, language, syntax, emotion concepts, concept acquisition

Understanding one's own and others' emotions is crucial to social communication, interpersonal relationships, emotion regulation, and well-being (Eisenberg, Sadovsky, & Spinrad, 2005; Hagelskamp, Brackett, Rivers, & Salovey, 2013; Kashdan, Barrett, & McKnight, 2015; Lindquist & Barrett, 2008; Lindquist, Satpute, & Gendron, 2015; Twenge, Catanese, & Baumeister, 2003). Greater understanding of emotions in childhood predicts later academic and interpersonal success (for reviews see Lindquist, Gendron, & Satpute, 2016; Shablack & Lindquist, in press). Moreover, multiple forms of psychopathology are characterized by difficulties in understanding emotions, such as autism (Baron-Cohen & Wheelwright, 2004; Baron-Cohen, 1991; Dapretto et al., 2006; Wang, Dapretto, Hariri, Sigman, & Bookheimer, 2004), depression (Berenbaum & Oltmanns, 1992; Joormann, 2010; Murphy et al., 1999; Phillips, Drevets, Rauch, & Lane, 2003) and anxiety (Etkin & Wager, 2007; Mennin, McLaughlin, & Flanagan, 2009; Salters-Pedneault, Roemer, Tull, Rucker, & Mennin, 2006). It is thus important to understand the mechanisms by which some people gain a complex understanding of emotions, whereas others fail to.

One hypothesis is that a complex understanding of emotions is learned, primarily during early development via discourse with caregivers (Castro, Halberstadt, Lozada, & Craig, 2015; Dunsmore & Halberstadt, 1997; Garrett-Peters, Castro, & Halberstadt, 2017; Halberstadt & Lozada, 2011; Shablack & Lindquist, in press.; Weinberg, Tronick, Cohn, & Olson, 1999). In this view, interactions with caregivers, but particularly emotion word use by caregivers, helps children to acquire a rich cache of knowledge about the emotion concepts most relevant to their culture (Campos, Frankel, & Camras, 2004; Denham, Zoller, & Couchoud, 1994; Ellis, Alisic, Reiss, Dishion, & Fisher, 2014; Fivush, Brotman, Buckner, & Goodman, 2000; Fivush, Haden, & Reese, 2006; Fogel et al., 1992; Halberstadt, Denham, & Dunsmore, 2001; Pons, Harris, & de

Rosnay, 2004). Indeed, research in cognitive science suggests that words help infants, children, and adults acquire and use concepts of all kinds (Lupyan, 2012a, 2012b; Xu & Kushnir, 2013). In particular, words facilitate the acquisition of socially shared concepts such as emotions (Doyle & Lindquist, 2018), color (Steels & Belpaeme, 2005), and spatial relations (Casasanto, 2008; Casasanto & Bottini, 2014) in adults. Yet the mechanisms by which children learn emotion concepts and their labels remains relatively unknown. Unlike object concepts (e.g., *tiger*, *table*), which are concrete, stable, and typically labeled by nouns, emotion concepts are abstract, transient, and typically labeled by adjectives, which makes them a difficult class of concepts to learn labels for (Gentner, 1982; Pinker, 1984). By combining research in the development of emotion and psycholinguistics, the present studies assess for the first time how children might learn to associate novel words with emotion concepts. We hypothesize that like certain other lexical categories, children use the syntactic structure of language to infer the meanings of novel adjectives labeling emotion concepts. In addition, based on prior research on emotion concept understanding (Kayyal, Widen & Russell, 2015; Widen & Russell, 2010, 2011) we hypothesize that situational cues from the environment will be needed to understand that novel adjectives denote emotion concepts.

The role of language in emotion development

Many psychological models of emotional development argue that the ability to experience, perceive, and ultimately understand specific emotion concepts follows a developmental trajectory and is learned through social relationships and verbal communication between children and caregivers (Castro et al., 2015; Dunsmore & Halberstadt, 1997; Garrett-Peters et al., 2017; Halberstadt & Lozada, 2011; Shablack & Lindquist, in press; Weinberg et al., 1999). In this view, infants start life with the ability to experience in their own bodies and perceive in others very basic feelings such as agitation and excitement (Bridges, 1932) or pleasantness and unpleasantness (Camras, 1992; La Barbera, Izard, Vietze, & Parisi, 1976; Lewis & Brooks, 1978; Russell & Bullock, 1985, 1986; Widen, 2013). It is hypothesized that over time, infants and children learn to make more fine-grained discriminations amongst these basic feelings. According to psychological constructionist approaches, concept knowledge about emotions is what ultimately helps children to learn to differentiate between different types of unpleasantness (e.g., *fear* vs. *anger*) or different types of pleasantness (e.g., *joy* vs. *pride*; Barrett, 2006, 2013; Barrett & Russell, 2015; Lindquist, 2013; Russell, 2003).

One perspective is that language helps children to acquire emotion concept knowledge because words serve as "essence placeholders" that cohere together instances (e.g., feeling aggressive, feeling like one's goals are blocked, feeling one's heart beating more quickly) as members of the same emotion category (e.g., anger; for a review, see Lindquist, MacCormack, et al., 2015). This hypothesis shares much in common with hypotheses about the role of language in the acquisition of other concept types (Lindquist, MacCormack, et al., 2015; Lupyan, 2012a, 2012b; Xu & Kushnir, 2013). It suggests that as children develop a larger emotion vocabulary, they develop more nuanced emotion concept knowledge and are thus able to perceive, express and experience a wider range of emotions.[1]

---

[1] Psychological constructionist models of emotion stand in contrast to "basic emotion" approaches that assume that certain emotions are innate, such that children are able to experience in their own bodies and perceive in others discrete emotion concepts such as *anger, disgust, fear, happiness, sadness* and *surprise* from birth (Ekman, 1992; Ekman & Cordaro, 2011; Izard, 2007; Panksepp, 1998). Although a prominent view, there is relatively little empirical support for the idea that infants and young children reliably perceive these emotions in others or themselves; the evidence is more consistent with the hypothesis that infants and children can differentiate between dimensions such as valence and develop an understanding of different emotion categories over early childhood (Rupa & Repacholi, in press; Shaback & Lindquist, in press; Widen, 2013). Although some basic emotion views acknowledge a role for learning in emotion understanding and experience, they stipulate that language itself is unrelated to emotion save for communication (Ekman & Cordaro, 2011). Nonetheless, these approaches would require that children need to learn how to communicate about emotions with words, a task which requires mapping experienced concepts to word forms during early development.

There is ample evidence to suggest that discourse with parents about emotions predicts children's greater emotional perception and understanding. For instance, correlational evidence suggests that as children age, emotion word knowledge increases (Bretherton & Beeghly, 1982; Ridgeway, Waters, & Kuczaj, 1985; Wellman, Harris, Banerjee, & Sinclair, 1995) as does performance in emotion perception tasks (Astington & Jenkins, 1999; Cutting & Dunn, 1999; Harris, De Rosnay, & Pons, 2005; Wellman et al., 1995). Furthermore, in 3- to 6-year-olds, general verbal ability (when controlling for age, attachment security, and gender) is an important predictor of children's ability to understand the emotions of others (de Rosnay & Harris, 2002; de Rosnay, Pons, Harris, & Morrell, 2004; Pons, Lawson, Harris, & De Rosnay, 2003). When using emotion words more specifically, toddlers initially describe their own feelings and the feelings of others in broad valenced terms by using general words such as *happy* and *sad* or *mad*. Yet by age 5 they additionally use words such as *afraid*, *surprised* and *disgust* to describe a more nuanced range of emotional states (Widen & Russell, 2003). This effect appears to go beyond mere labeling, as children's ability to perceptually identify emotional facial expressions in a nuanced manner increases as emotion labeling ability increases. For instance, early in toddlerhood, children tend to use the words *happy* and *sad*[2] in daily discourse, and correspondingly can only reliably differentiate pleasant and unpleasant facial expressions from one another. Yet, around the time that children begin to use the words *anger* and *fear* in daily discourse to differentiate between different negative states, they also become able to perceive

---

[2] Whereas some children use the word *sad* to label negative states, others use the word *mad* instead (Widen, 2013). Nonetheless, most children at this age only reliably use two words for emotion, meaning that if they use *happy* and *sad,* they do not also use *mad* (and vice versa). It is unknown why this idiolectal difference in use of *sad* v. *mad* occurs, but one possibility is that this difference could be a product of socialized gender stereotypes (Plant, Hyde, Keltner, & Devine, 2000), whereby parents use the term *sad* more with girls and *mad* more with boys when explaining emotions (Fivush, 1991; Fivush, Brotman, Buckner, & Goodman, 2000).

these negative emotions on faces (i.e., distinguish anger from fear) in perceptual sorting tasks (Widen & Russell, 2008). Importantly, asking children to match emotional facial expressions to words, as opposed to other facial expression exemplars, facilitates children's performance, even among children as young as 2 years old (Russell & Widen, 2002b, 2002a).

These findings suggest that caregivers help children to label their own emotional states and the emotional states of others in an effort to scaffold their acquisition of knowledge about the emotion concepts relevant to the present context (Campos et al., 2004; Denham et al., 1994; Ellis et al., 2014; Fivush et al., 2000, 2006; Fogel et al., 1992; Halberstadt et al., 2001; Pons et al., 2004). However, very little research has assessed how children learn that a word denotes an emotion concept in the first place. This is an important next step in understanding the relationship between language and emotion and is the purpose of the present studies.

The role of linguistic cues in word learning

Much of the experimental work on children's vocabulary development focuses on how children acquire words for object concepts, which are primarily labeled by nouns (Bloom, 2000; Gentner, 1982; Huttenlocher & Smiley, 1987; Markman, 1990). This emphasis is logical, as children's earliest vocabulary items are largely nouns that label people and basic objects (Bates et al., 1994). However, words of different lexical categories (verbs, adjectives, etc.) tend to have very different kinds of meanings and are learned in very different ways. For instance, verbs often label actions and events, and adjectives, which modify nouns, typically label properties or attributes. Emotions are internal states that are most frequently labeled by adjectives in everyday speech (Shablack, 2017), and verbs and adjectives are conceptually more complex than nouns (Gentner, 1982). Moreover, while caregivers may label salient objects for children ostensively (e.g., “Look! That's a dog!”), caregivers do this only rarely (if at all) with properties and states of being (Gleitman, 1990).

As a product of their abstract and complex nature and lack of ostensive instruction, verb and adjective meanings are less straightforward than noun meanings for children to learn, and they are learned later (Gentner, 1982). An influential approach called the *syntactic bootstrapping hypothesis* (Fisher, Gleitman, & Gleitman, 1991; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Gleitman, 1990; Naigles, 1996) points to the role of syntactic structure in cueing the meanings of these types of words; work within this approach has focused mainly on verbs. Both within a language and across different languages, the number of noun phrases (i.e., *arguments*) that occur with a verb in a sentence is related in highly regular ways to the lexical meaning of the verb. Children can thus use these arguments to infer the meaning of verbs (see Table 1 for examples). For instance, the novel verb *gorp* in the sentence *Bill gorped Jane a rose* could not mean something like 'sleep' or 'kiss' (note the oddness of **Bill slept Jane a rose*).[3]

The syntactic bootstrapping hypothesis has been tested (Fisher et al., 1991; Gleitman et al., 2005; Gleitman, 1990; Naigles, 1990, 1996) by presenting young children with novel verbs used in particular sentence “frames” (e.g., with one vs. two arguments). Children are then asked to point to, or their eye gaze is tracked towards a visual image depicting the relevant action (Fisher, 2002; Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990; Yuan & Fisher, 2009). The methodology has been adapted to test learners' interpretations of not only verbs denoting observable actions but also relatively more abstract kinds of verbs, such as those that could have a meaning like *think* or *seem*, or to abstract adjectives such as *easy* (Becker, 2006, 2014, 2015; Becker & Estigarribia, 2013; Papafragou, Cassidy, & Gleitman, 2007). In some cases, such as for

[3] Linguists use the symbol * to indicate ungrammaticality or unacceptability.

abstract verbs such as *think*, the sentence frame is even more informative than extragrammatical situational context (what is going on in the world when a particular verb is uttered) for cueing children to the predicate's meaning (Gleitman et al., 2005; Papafragou et al., 2007). However, children can draw such inferences about more abstract verbs only after age 4 (Becker, 2006, 2014).

Although syntactic bootstrapping is a likely mechanism by which children learn predicate meanings, there is little work examining the relative role of sentence frame in the acquisition of adjectives as a lexical category (Booth & Waxman 2003, 2009; Syrett & Lidz, 2010). No work to our knowledge looks at the application of this approach to novel emotion concepts. Insofar as emotions are most often labeled by adjectives (*happy, sad, mad,* etc.) (Shablack, 2017), children should be able to use syntactic bootstrapping to understand the meaning of novel emotion words and concepts. On the other hand, emotions are highly situated concepts in which people are responding to real or imagined events in the world around them (Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011). As such, access to details about the context in which emotions are occurring (Harris, Johnson, Hutton, Andrews, & Cooke, 1989; Kayyal et al., 2015; Widen & Russell, 2010, 2011, 2013) may influence a child’s inferences about the meaning of emotion concept labels. In particular, situational details from the immediate context, such as information about the causes of a certain emotion (e.g., loss) and what a person did as a consequence (e.g., cry) help children to accurately identify the emotion concept being described in a story, often above and beyond other perceptual information such as facial expressions (Kayyal et al., 2015; Widen & Russell, 2010, 2011). When the evidence is taken together, it is likely that children use both linguistic information (sentence frames) and situational cues from the environment (causes and consequences of emotion) in the acquisition of emotion words. Yet no work to our knowledge has assessed how children use syntactic structure in combination with situational context to learn the meaning of emotion concept labels.

Present studies

We conducted two studies examining the extent to which sentence frames and situational context impact children's (aged 3-5) understanding that novel words denote emotion concepts. In Study 1, we manipulated the sentence frame a novel word was presented in to examine whether sentence frame alone helps children perceive that a word denotes an emotion concept as opposed to other predicate meanings such as a physical state or an action, and at what age this occurs. Sentence frames were manipulated by presenting a novel word in a verb structure that limits the possible meaning of the word. For example, when presenting a novel word (*binty*) as the sole complement of the verb *is* (i.e., *Susan is binty*), the meaning of *binty* is fairly unlimited. As an adjective *binty* can denote an emotion, such as 'happy,' a physical state, such as 'cold,' or even a physical characteristic, such as 'tall' (note the grammaticality of *Susan is happy/cold/tall*). However, if *binty* is the complement of the verb *feels* (i.e., *Susan feels binty*), the meaning of *binty* can no longer denote a physical characteristic like 'tall' (excepting figurative uses of this adjective), though it could still denote either an emotion ('happy') or a temporary physical state ('cold') (thus: *Susan feels happy/cold/*tall*). Finally, if *binty* is used within a sentence like *Susan feels binty about something*, the meaning of *binty* is further restricted to denoting a mental or emotional state; it can no longer denote a physical state like 'cold' (*Susan feels happy/*cold/*tall about something*). Therefore, we predicted that if children rely exclusively on linguistic cues, the likelihood for a child to determine that a novel word is an emotion concept label would increase as the sentence frame becomes more restrictive.

In Study 2, we examined the extent to which the sentence frame helped children acquire the meaning of novel emotion concepts when that acquisition was also scaffolded by situational context. We manipulated the sentence frame as in Study 1 and also told children a "background" story that highlighted the causes and consequences of an experience that the main character underwent. For example, prior to a sentence such as "Palooza feels binty," children heard a story about a character receiving a gold star for her drawing, providing situational context for what she was feeling and why. Study 2 thus examined whether the sentence frame a word appeared in is especially useful for understanding the meaning of a novel emotion concept when that word appears in a relevant situation.

Across both studies we hypothesized that older children would be more likely to perceive a novel word as an emotion concept label given appropriate linguistic and situational cues; however, as we hypothesize that sentence frames and situational context provide specific cues in concept acquisition and word disambiguation, we predicted age interactions with sentence frame (Study 1), especially in the presence of situational context (Study 2).

Study 1

In Study 1, we predicted that children would use sentence frame as a cue for their understanding of novel emotion words. Specifically, if children use sentence frame alone to drive emotion concept word acquisition, then children would especially perceive a novel word as labelling an emotion when it is presented in a more restrictive sentence frame (e.g., *feels* or *feels about*). Mirroring past work on syntactic bootstrapping (e.g. Papafragou et al., 2007) we predict that children would become more adept at using the sentence frame as age increases.

Methods

Participants. One hundred sixty-two children participated in the study at the Museum of Life and Science in Durham, NC. Children who did not meet our a priori inclusion criteria were excluded from further analysis. Nine children were not within our desired age range (i.e., were siblings of other children who participated or were found to not be in the proper age range following parent report). Three additional children were removed from the analysis because they lived in households in which languages other than English were spoken greater than 50% of the time; we reasoned that in the case of bilingual children our sentence frame manipulation would be less effective since those children might regularly experience more varied sentence frames. Two children were removed from analysis because their parent/guardian reported that they had a learning disability on the parent questionnaires. An additional 12 children started, but did not complete the video task because they either stopped participation part way through ($n = 9$) or failed the screening task ($n = 3$). One child completed an incorrect survey due to experimenter error and is not included. The final sample consisted of 135 ($M_{age} = 3.97$, $SD_{age} = .79$; 72 female) children: 44 3-year olds (25 female), 51 4-year olds (27 female), and 40 5-year olds (20 female).[4]

Children completed the study individually with a single experimenter. A second experimenter observed the interaction and answered any parent/guardian questions. Parents and/or legal guardians completed a packet of questionnaires about their child's home life and development. Sixteen individuals did not report family income; of the 119 who did, 19.3% reported a household income < 70K, 21% reported between 70K and 100K, 29.4%, between 100K and 150K and 30.3% reported 150K or higher. Thirty-one individuals did not report the

[4] Since the museum's staff wished to create an atmosphere of inclusivity, we permitted all interested children to participate in the task. Children were excluded post hoc only for the reasons listed above (out of age range, not monolingual, learning impairment, etc.), i.e. reasons that would have led to them not participating if we had prescreened them for eligibility. We excluded children who had already started the task if they failed to continue the task at that point without making children feel excluded.

race/ethnicity of one or both parents. Of the 104 who did, 83.7% of the children were Caucasian/white, 13.5% multiracial, 1.9% African American and .9% Asian.

Materials.

Novel Word Videos. Children viewed videos of two animal hand puppets conversing about aliens (see Table 2 for general dialogue) followed by a picture pointing task with three image options. These brief dialogues were modeled after those used by Yuan and Fisher (2009) and Arunachalam and Waxman (2010) in a syntactic bootstrapping study with 2-year-olds. Videos were on average 16s long and contained voices from two out of four different individuals. Voice identity was quasi-randomly distributed across all videos. Conversations about aliens were used both to engage children's interest and to allow the framing of the novel words as new "alien words." This limits any undue influence of already known emotion concept words via mutual exclusivity (i.e., the belief that two words cannot have the same meaning; Clark, 1987; Hutchinson, 1986; Markman, 1990).

Prior to the experiment, each child completed three screening trials. Each screening trial had a single video that discussed an alien engaging in an action (i.e., "I know an alien who likes to eat pizza") and an accompanying descriptive sentence using a verb in the present progressive (i.e., Point to where the alien is eating pizza!"). To continue, children had to correctly answer at least two trials (i.e., point to the picture of the alien eating pizza and not the distractor pictures), demonstrating that they could watch a brief video accompanied by a verbal description and point to a picture that matched the description.

In the subsequent experimental trials, children watched the hand puppet videos with puppets conversing and introducing a novel word (*binty*, *daxy*, *strupy*, *moky*, *joomy*, *gorpy*, *reksy*, *tropy*) four times using one of the following sentence frames:

Is: The alien is [novel word]

Feels: The alien feels [novel word].

Feels about: The alien feels [novel word] about something.[5]

Each child was presented with the same sentence frame for all experimental trials and no novel word was repeated across trials. Each experimental video and novel word was presented in random order.

In between each experimental trial was a filler trial containing a novel verb in the present progressive (ending in *-ing*: *piffing, tayving, serding*) to indicate an action. The filler trials were included to give children a break from the experimental trials and were not intended as control trials. Nonetheless, child performance on these trials was included as a covariate in the event that children's filler performance significantly impacted their performance on the trials of interest. See supplemental materials for analyses of filler trials and main analyses without filler trial performance as a covariate. Over the course of the session, each child saw 10 videos (3 screening, 4 experimental, 3 filler).

Picture Pointing Task. Following each video, a screen containing auditory instructions with accompanying text instructed the child to "Point to where the alien [sentence frame] [novel word]." The voice from a single individual was used in the auditory instruction. Children then saw three images of a cartoon alien presented in a random array, with each image depicting the alien expressing an emotion, a physical state, or engaging in an action (see Table 3 for list of image types and Figure 1 for sample trial). Children were instructed to point to the image that corresponded to the novel word. Their choice served as the dependent variable. Three alien

[5] For the Feels About conditions, story endings were: "about brushing his teeth", "about cleaning her room", "about eating cookies," and "about playing games." The picture pointing task instructed kids to "point to where the alien feels [novel word] about something."

identities were used in screening trials. Four different alien identities were used for the experimental and filler trials, with the limitation that each experimental trial employed a different alien (thus, the same alien may have been seen in both an experimental trial and a filler, but not in two experimental trials). Images were randomized across all trials to appear once (i.e., children saw all seven possible emotion images appearing in either an experimental or filler trial). All alien stimuli were validated in a separate sample of 3-5 year olds, in which children were better than chance at associating the images with the intended physical state, emotion or action (see Supplemental Materials).

Parent questionnaire. While the child completed the computerized task, parents and/or legal guardians completed a voluntary paper questionnaire. Information was gathered on the child's birth, including date, location and whether the mother's pregnancy was normal. Children's general communicative and linguistic development was measured including any information on whether the child had been evaluated for speech problems or learning disabilities, and a rudimentary number of spontaneous word production of a subset of word categories. Information on familial language was gathered, as well as who the child lived with. Parental race/ethnicity, place of birth and native language, education, career, and household income were also gathered. The first was used as a measure of child race/ethnicity.

Procedure. This study was approved by the University of North Carolina Institutional Review Board. Parents of children who looked to be 3-5 years of age were approached at the museum and asked if their child would like to participate in a short video task on word learning. Only children who were 3-5 were subsequently enrolled (with the exception of children who were accidentally enrolled due to a miscommunication about their age or allowed to participate because a sibling participated. These children were not included in analyses). Children were told that they would play a game involving aliens. Following the parent/legal guardian's consent, the experimenter spent the first few minutes getting to know the child and obtaining verbal assent. The experimental task was then administered online on a laptop computer via Qualtrics.

Children who failed two or more of the three screening trials were thanked for their time and told that the game was over. If children passed the screening trials, the experimenter then introduced the experimental task by saying, "You're doing great! Let's keep playing this game! Now we're going to watch some more videos. These videos are going to have a special alien word and I need help to figure out what that word means! Do you want to play/Are you ready?" If the child wanted to continue, the experimenter then played the first experimental video containing a novel "alien" word followed by a screen with visual and auditory instructions for the picture pointing task. Once an image was chosen, the experimenter confirmed the choice and made the selection by clicking the radio button beneath the image the child chose. If a child was hesitant, the experimenter encouraged them up to three times before continuing to the next trial. Following the video task, children had the option to complete an additional task, not discussed here. For their participation, children received a hand stamp and temporary tattoo.

Results

To examine the impact of Sentence frame and Age on Image choice, mean proportions of each Image choice type were created across trial types (e.g., mean proportion of emotion images chosen within each experimental condition).[6] We opted to treat age as a categorical variable rather than a continuous variable due to the ease of interpreting mixed model ANOVAs over regression models with categorical outcomes. We nonetheless also computed multi-level

---

[6] Eight children had technical difficulties during 1 or 2 trials. In these cases, mean proportions were weighted by the number of completed trials.

multinomial logistic regression analyses in which age was a continuous predictor and sentence frame was a categorical predictor of choice outcomes (see supplemental materials). These findings replicated the mixed model ANOVA findings so we report the ANOVA findings in the main text for ease of interpretation.

To examine our main hypothesis that sentence frame and age would interact to influence children's image choice, we conducted a 3 (Image choice: emotion, action, physical state) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANOVA with mean proportions of Image choice as a within subjects factor and Sentence frame and Age as between subjects factors. To control for the effects of participant gender and screening trial performance, we also conducted a 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with gender and performance on the screening trials as covariates. Lastly, to account for filler trial performance, we conducted a 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with gender, performance on the screening trials, and performance on filler trials as covariates. Filler trial performance was computed as the proportion of trials in which an action image was chosen. Findings were largely identical across the three analyses, so we report the most conservative ANCOVA findings controlling for gender, performance on screening and filler trials. Findings from the other ANOVA and ANCOVA are available in the supplemental materials.

See Table 4 for all effects; we explicitly discuss only significant effects and predicted effects. First, we found a main effect of Image choice, $F(2, 246) = 3.10$, $p = .05$, $\eta^2 = .03$.[7]

---

[7] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 6.36, p = .04, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\varepsilon = 0.95$, being greater than .75.

Children chose physical state ($p < .001$) and emotion ($p < .001$) images significantly more than action images. Children chose physical state images significantly more than emotion images over all ($p = .002$) (see Figure 2).

The main effect of Image choice was qualified by a significant interaction between Image choice and Age, $F(4, 246) = 2.53$, $p = .04$, $\eta^2 = .04$. Simple effects show a significant effect of Age on action images, $F(2, 123) = 4.69$, $p = .01$, $\eta^2 = .07$ and a significant effect of Age on physical state images, $F(2, 123) = 3.65$, $p = .03$, $\eta^2 = .06$. Pairwise comparisons reveal that 3-year-olds chose action images ($p = .003$) significantly more than 5-year-olds. Five-year-olds chose physical state images ($p = .008$) significantly more than 3-year-olds did. Emotion images were chosen equally between all ages (see Figure 3).

As predicted, there was a marginal interaction between Image choice and Sentence frame, $F(4, 246) = 2.28$, $p = .06$, $\eta^2 = .04$, suggesting that children relied on more restrictive sentence frames to inform image choice. Simple effects reveal that the proportion of trials in which emotion images were chosen is marginally influenced by the Sentence frame, $F(2, 123) = 2.92$, $p = .06$, $\eta^2 = .05$. Pairwise comparisons demonstrated that emotion images were chosen significantly more in the Feels About ($p = .02$) than the Feels condition, but that the Is condition ($p = .57$) did not differ from Feels About. Emotion images were chosen marginally more during Is trials than during Feels trials ($p = .09$) (see Figure 4).

We did not find a predicted 3-way interaction between Image choice, Age, and Sentence frame, suggesting that children did not rely more on sentence frame with increasing age. Instead, Age and Sentence frame separately moderated image choice.

Discussion

Results provide preliminary support for our hypotheses in Study 1, which is a conservative test of the syntactic bootstrapping hypothesis in which children received no other context for the meaning of a novel word except the sentence frame it was heard in. We found that age influenced children's tendency to perceive a novel word as denoting an internal state (physical state or emotion) over an action. Three-, 4- and 5-year-olds were equally likely to choose emotion images, with the relative proportion of physical state to action choices increasing over age. Three-year-olds chose more action items than 5-year-olds, whereas 5-year-olds were more likely than 3-year-olds to choose physical state images. Since all of our novel words ended in *-y* (*daxy*, *joomy*, *binty*), and words ending in this sound are prototypically adjectives (*happy*, *bouncy*, *hungry*, *thirsty*, *rosy*, *furry*, *smelly*), this feature alone may have been a cue to 5-year-old children that novel words labeled attributes or states. The fact that 3-year-olds chose the action images the most out of the age groups suggests that 3-year-olds may not yet understand that words ending in *-y* are likely to be adjectives.

We also found suggestive evidence that more restrictive sentence frames impact children's understanding that novel words refer to emotions, although this effect was only marginal. Children were more likely to choose emotions in the most restrictive sentence frame (Feels About) as compared to the Feels condition. Children were equally likely to choose emotions in Is and Feels About conditions, but recall that *is [emotion adjective]* is also a grammatically correct choice.

We note that we did not find the predicted three-way interaction between Age, Sentence frame, and Image choice. This null finding could in part be a result of the conservative study design we used, in which children received no other context but the linguistic context. If younger children cannot use sentence frames effectively and thus choose randomly, whereas older children are reasonably concluding that Is, Feels, and Feels About are each grammatically appropriate for a third of the trials, then it may have been difficult to observe a three-way interaction. This interaction was perhaps even more difficult to observe in light of all children's bias towards choosing physical state images.

There are two potential interpretations of children's bias towards physical state images. First, a bias towards choosing the physical state (as opposed to the emotion) may merely reflect the relative salience of these images to children, as physical state images in our study often contained other potentially interesting details such as areas of differently colored skin, bandages, etc., which were necessary to convey meanings such as *cold, hot, hurt,* and *sick.* This bias might have especially occurred in Study 1 since there was no other situational information to drive children's attention to the other images. Without situational information, the older children may have used the linguistic cue of the novel word ending in *-y* (suggestive of an adjective, rather than the verb-indicating *-ing*) to narrow their choices down to the physical state and emotion images, and then fixated on the physical state images due to their additional intrigue. Indeed a few children did note that they chose images based on color (i.e., "that one because it's blue" and "I chose that one cause it's blue and blue is my favorite color").

This interpretation is consistent with the fact that children even showed a bias towards physical state images in the filler condition (see Supplemental Materials). We included filler items since it is common to include filler items to break up the experimental trials on this type of task (Gerken & Shady, 1996). Note that accounting for children's performance on the filler trials did not alter our results, meaning that children's ability to perform on filler trials did not affect our findings. Nonetheless, we cannot be sure why children chose the physical state images on the filler trials. One possibility is that in the absence of any situational context, children inferred that the images had multiple meanings. That is, an image of a physical state in which an alien was sick could have been interpreted as also “disgusted” or “sad” by children, which would have been technically correct. Our validation data are somewhat consistent with this interpretation; although children performed greater than chance on all stimuli, the stimuli that they performed the least accurately on were the physical states of “sick” and “hurt.” In both cases, children were most likely to misinterpret these states as emotional (e.g., disgusted, sad). This fact is not likely to just be a limitation of our stimuli. Mental states have multiple levels of meaning and individuals differ in the complexity of the inferences drawn about those states—a person who has tears coming from their eyes and is emitting sound from their mouth could be “crying,” “sad” or “experiencing grief at a loss” (Vallacher & Wegner, 1987, 1989; Wegner & Vallacher, 1987). Even highly caricatured facial expressions of emotion are perceived as indicative of mental states in some cultures (e.g., sad) but as actions (e.g., crying) in others (Gendron, Roberson, van der Vyver, & Barrett, 2014).

Nonetheless, observations of children’s spontaneous comments, in combination with our validation data, suggests that on average children understood the intended meaning of the alien images. Children were more accurate than chance at identifying the correct emotion, physical state, and action image in our validation study (see Supplementary Materials). Additionally, children’s spontaneous utterances during our experimental trials suggested that they understood the meaning of the stimuli. For instance, for action images children stated, “this is running,” “[s/he is] falling down,” “the one who is walking is moky” or imitated running. For physical state images children proclaimed things such as, “this one’s hot,” “sweaty!,” “he has two Band Aids, he hurts,” “maybe he broke his leg,” and “he’s ouchy.” Finally, for emotion images children made statements such as “this means mad,” “he’s scared!” and “that one is sad.”

Second, it is possible that children had a bias towards the physical state images because the Feels vs. Feels About sentence frame was not sufficiently helpful across the age range to help children move beyond their bias towards physical state images. This interpretation is consistent with our observation of an interaction between Image choice and Sentence frame, but lack of a 3-way interaction between Sentence frame, Age and Image choice. The marginal interaction between Image choice and Sentence frame suggests that children may be able to use the sentence frame to begin to home in on the correct word meaning, supporting prior research (Becker, 2015; Gleitman et al., 2005; Papafragou et al., 2007). However, the fact that older children were not significantly more likely to choose the emotion image over physical state images in the Feels About condition suggests that even 4- and 5-year-olds are not yet able to use Feels About to restrict an adjective meaning to an emotion rather than a physical state.

Taken together, the findings of Study 1 suggest that although linguistic cues may be meaningful to children in this age range, children may not be able to use sentence-level cues alone to understand emotion adjectives. These findings are consistent with recent work showing that the same sentence-level cues children are known to exploit for learning verbs may not be usable for adjectives at this stage of development (Booth & Waxman, 2009; Syrett, Latourrette, Ferguson, & Waxman, 2018). However, as we note, Study 1 is a particularly conservative test of the syntactic bootstrapping hypothesis, as it is atypical for children to experience emotion concepts and their words in the absence of an emotional situation. We thus included situational context in Study 2 to more clearly test the role of sentence frame in children's understanding of novel emotion words when an emotional situation was also present.

Study 2

Since research on emotion concept development demonstrates that children rely on situational context--including the causes and consequences of an emotion when determining the meaning of an emotional facial expression (Kayyal et al., 2015; Widen & Russell, 2010, 2011)--we modified Study 1 to include information about the situational context a novel adjective was occurring in. Study 1 was a conservative test of the syntactic bootstrapping hypothesis, but we reasoned that including situational context would ultimately be a more ecologically valid test of the types of contexts in which children learn about emotions (e.g., when observing an event and hearing an adult talk about it).

To achieve this end, children in Study 2 viewed videos of a cartoon alien experiencing emotional situations. We predicted that with this additional situational context present, we would observe a 3-way interaction such that older children would be more likely to choose emotion images when the sentence frame was maximally restrictive. As the procedure is largely identical to Study 1, we only describe the differences between them below.

Methods

Participants. One hundred forty seven children participated in the study at the Museum of Life and Science in Durham, NC. As in Study 1 children who did not meet our a priori inclusion criteria were excluded from further analysis. Nine were not within our desired age range (had mistakenly been enrolled or allowed to participate because a sibling had done so). Seven were removed from analysis because they lived in households where languages other than English were spoken greater than 50% of the time, and one was later found to have a learning disability as reported by their parent/legal guardian in the parent questionnaires. Eighteen children did not complete the video task either due to screening trials failure (n = 9) or opting not to continue in the midst of the task (n = 9). The final sample consisted of 113 (M_{age} = 4.04, SD_{age}

= .77; 45 females) children: 31 3-year olds (8 female), 46 4-year olds (26 female), and 36 5-year olds (11 female).

Parents and/or legal guardians completed a voluntary packet of questionnaires identical to that in Study 1. Twelve individuals did not report family income; of the 101 who did, 20.8% reported a household income < 70K, 28.7% between 70K and 100K, 26.7% between 100K and 150K and 23.8%, 150K or higher. Sixteen parents/guardians did not fill out the race/ethnicity information for one or both parents. Of the 97 who did, 82.5% of the children were Caucasian/white, 14.4% multiracial, 1.03% African American, 1.03% Asian and 1.03% Hispanic.

Materials.

Novel word videos. Children viewed videos of an alien cartoon character with a narrator describing a short story (see Table 5 for sample stories). Videos were created using GoAnimate (http://www.goanimate.com) and were on average 23.85s. A single individual narrated the story for all videos. The screening video scripts were similar to Study 1, but now had a single narrator, a name associated with the alien, and a cartoon accompanying the auditory stimuli. Children who failed 2 or more of the 3 screening trials were thanked for their time and told that the game was over.

During the experimental trials, children watched seven cartoon videos with accompanying narration: each video employed a different novel word (*daxy*, *moky*, *reksy*, *binty*, *gorpy*, *joomy*, *tropy*) in one of three between-subject experimental sentence frames (Is, Feels, Feels About) as in Study 1. Each alien character in the video had a neutral expression. Aliens, video stories and novel words were fully randomized such that no single alien, story, or word were consistently paired together. In contrast to Study 1, the novel word was presented twice (rather than four times) in the video and a third time (rather than fifth time) in the picture pointing task instructions (see Table 5 for sample stories and Figure 1 for a sample trial). Emotional situations were created by providing a brief story that highlighted a positive or negative discrete emotional experience for the alien (the alien was *happy, sad, afraid, mad, excited, disgusted,* or *surprised*), prior to the introduction of the novel word in the target sentence frame. Stories were kept short to limit cognitive burden and to keep the child's attention. Stories were developed based on prior work (see Widen & Russell, 2010) drawing from the prototypical causes of emotions in a North American setting. A total of 7 experimental trials of a single alien identity were presented. We reasoned that variation across videos (i.e., seeing aliens and stories) would be sufficiently attention-capturing for children so we did not include filler trials to maintain their attention. All sessions began with the *happy* story followed by the remaining 6 emotions in randomized order. We modeled this method after other developmental research on emotion (Russell & Widen, 2002b, 2002a; Widen & Russell, 2010). In these studies, *happy* trials are presented first because *happiness* is a well-understood emotion concept for 3- to 5-year-olds; it is assumed that receiving a more difficult emotion concept first might discourage children.

Picture pointing task. Cartoon alien images were identical to Study 1 (Table 3), however, each matched the alien seen in the video, where s/he was given an identifying name (screening trials: Chrysanthemum, Magenta, Frebedo; Experimental trials: Palooza, Chromia, Wazu, Xylobean). Female participants viewed either Palooza or Chromia (the "female" aliens) and male participants viewed either Wazu or Xylobean (the "male" aliens). We matched the gender of the aliens and children since there is some evidence that interpersonal similarity facilitates mental state inference (Ames, 2004). All children saw Chrysanthemum, Magenta and Frebedo for the screening trials. During the picture pointing task, images of the aliens displaying a particular emotion, action, and physical state image appeared only once throughout the seven trials such that no image was repeated (e.g., children saw a happy alien on only one trial throughout the experiment). The emotional image depicted always matched the story (e.g., children saw a happy alien for a story describing *happiness*) but the particular action and physical state seen were randomly displayed.

Procedure. This study was approved by the University of North Carolina Institutional Review Board. The procedure was identical to Study 1, save that participants saw videos of the alien scenarios, rather than videos of puppets holding short conversations.

Results

Analysis procedures are identical to that in Study 1 with mean proportions of each image choice type calculated across trial types (e.g., mean proportion of emotion images chosen within each experimental condition)[8] entered in mixed model ANOVAs. To examine our main hypothesis, a 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANOVA was conducted with Image choice as a within subjects factor and Sentence frame and Age as between subjects factors. As in Study 1, we also computed a 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with Image choice as a within subjects factor and Sentence frame and Age as between subjects factors, including participant gender and performance on the screening trials as covariates. Findings were largely identical across the two analyses so we report the more conservative ANCOVA findings. See supplemental materials for the ANOVA and multilevel multinomial logistic regression analyses.

[8] One child experienced technical difficulties during one trial. As in Study 1, their mean proportions were adjusted to be out of the number of completed trials.

See Table 6 for all effects; We discuss only the significant effects and predicted effects here. As predicted, and as in Study 1, we found a significant interaction between Image choice and Age, $F(3.73, 190.38) = 8.29$, $p < .001$, $\eta^2 = .14$.[9] Emotion images were chosen more as Age increases, and both physical state and action images were chosen less as Age increases. Simple effects reveal significant Age differences in the proportion of trials in which emotion images were chosen, $F(2, 102) = 12.63$, $p < .001$, $\eta^2 = .20$, action images were chosen, $F(2, 102) = 5.65$, $p = .005$, $\eta^2 = .10$, and physical state images are chosen, $F(2, 102) = 4.59$, $p = .01$, $\eta^2 = .08$. Pairwise comparisons reveal that 5-year-olds chose emotion images significantly more than 4- (p = .01) and 3- year-olds ($p < .001$). Four-year-olds chose emotion images significantly more than 3-year-olds (p =.006). Three-year-olds chose action images significantly more than 4- ($p = .03$) and 5- ($p = .001$) year-olds. Four- and 5-year-olds did not differ significantly in the proportion of action images chosen ($p = .18$). Three-year-olds chose physical state images significantly more than 5-year-olds ($p = .003$). Three and 4-year-olds did not differ in significantly in the proportion of trials in which they chose physical state images ($p = .12$) (see Figure 5).

Critically, as predicted, we found a 3-way interaction between Image choice, Age and Sentence frame, $F(7.47, 190.38) = 2.49$, $p = .02$, $\eta^2 = .09$ (Figure 5). To further probe this interaction, we examined the 2-way interaction between Sentence frame and Image choice for each age-group. There was no interaction between Sentence frame and Image choice for 3-year-olds, $F(4, 52) = .41$, $p = .80$, $\eta^2 = .03$, suggesting that 3-year-olds were unable to use the sentence frame to guide their choice of images.

[9] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 21.81, p < .001, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\varepsilon = 0.837$, being greater than .75, $df_{error} = 190.38$.

However, as predicted, there was a significant interaction between Sentence frame and Image choice for 4-year-olds, $F(3.80, 77.99) = 2.67$, $p = .04$, $\eta^2 = .12$.[10] Examination of the simple effects for 4-year-olds reveals that emotion images were selected to a different degree based on Sentence frame, $F(2, 41) = 3.37$, $p = .04$, $\eta^2 = .14$. Pairwise comparisons reveal that 4-year-olds chose emotion images significantly less in the Is frame than the Feels frame ($p = .01$), but did not differ between the Is and Feels About frame ($p = .13$). Emotion images did not differ between Feels and Feels About ($p = .17$). Four-year-olds thus chose emotion images similarly across Feels and Feels About sentence frames and least in Is sentence frames. Simple effects also revealed that 4-year-olds chose action images to a different degree based on Sentence frame, $F(2, 41) = 3.90$, $p =.03$, $\eta^2 = .16$. Pairwise comparisons demonstrated that 4-year-olds chose action images significantly more in Is frames than Feels frames ($p = .01$) and Feels About frames ($p = .02$). Sentence frame did not influence the proportion of trials in which physical state images were chosen, $F(2, 41) = 1.00$, p = .38, $\eta^2 = .05$ (see Figure 6). Finally, the interaction between Image choice and Sentence frame for 5-year-olds was not significant, $F(2.79, 43.24) = 2.14$, $p = .11$, $\eta^2 =.12$[11] (see Figure 6).

Discussion

Study 2 revealed that when children could draw on a situational context that highlighted a caused emotion, Age interacted with Sentence frame to alter Image choice. When the situational context suggested that a novel word referred to an emotion concept, the Sentence frame particularly influenced image choices for 4-year-olds. Three-year-olds were unable to use the

---

[10] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 8.81, p = .012, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, ε = 0.84, being greater than .75.

[11] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 17.06, p < .001, therefore degrees of freedom were corrected using Greenhouse-Geisser corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, ε = 0.70, being less than .75.

sentence frame to help guide their image choices and chose randomly. Five-year-olds chose the emotion most frequently regardless of sentence frame. Five-year-olds thus recognized that Feels About is a sentence frame that is compatible only with emotion meanings, but that emotion meanings were otherwise appropriate in all conditions (due to both the story contexts and linguistic framing). However, 4-year-olds were more likely to choose emotion images in the Feels and Feels About conditions than in the Is condition, suggesting that they were using sentence frame in combination with the situational context to understand that a novel word labeled an emotion.

Interestingly, our findings suggest that more restrictive sentence frames are not useful for children of all ages. Four-year-olds chose emotion images more in the Feels condition than the Feels About condition, whereas 5-year-olds chose emotion images less in the Feels condition than in both the Is and Feels About conditions. Together, these findings suggest that Feels About may not be a useful cue for 3- and 4-year-olds, for whom feeling an abstract state that is in response to an unknown cause might be too complicated of a mental state inference. In contrast, 5-year-olds used Feels About easily and show similar facility doing so with Is, which is a much less complex (and less restrictive) sentence frame. It is unclear why 5-year-olds did not choose emotions as frequently in the Feels condition as in the Feels About and Is conditions, but this finding may be related to the fact that 5-year-olds are beginning to use more complex sentence frames, such as Feels About for emotional states. Although these questions should be addressed in future research, Study 2 presents initial evidence that children are using the sentence frame to understand the meaning of novel emotion words when those words are heard in the context of an emotional situation.

General Discussion

Across two studies, we examined the extent to which both linguistic and situational cues are important in children's understanding that novel words denote emotion concepts, as opposed to physical state or action concepts. Study 1 provided preliminary evidence that, consistent with the linguistic literature on verb learning, sentence frames may be informative for children about the meaning of novel emotion concept words. A marginal interaction between Image choice and Sentence frame suggests that children may be able to use more restrictive sentence frames to home in on a word's meaning. However, these data suggest that sentence frame alone may not be sufficient for children to understand that a novel word denotes an emotion concept. Study 2 further examined the role of sentence frames when children had access to situational cues, mirroring more ecologically valid learning contexts. When linguistic input was heard in the presence of situational cues, we observed a predicted 3-way interaction between Image choice, Age and Sentence frame. Sentence frame guided the selection of emotion images over physical state or action images for 4-year-olds, but not 3-year-olds. This finding is interesting insofar as even 2-year-olds can exploit sentence frames for learning verb meanings (Arunachalam & Waxman 2010, Yuan & Fisher 2009). However, adjectives are more difficult to learn than verbs, and other evidence shows that 3-year-olds have difficulty using sentence frames to draw inferences about the meaning of adjectives (Syrett et al. 2018). Mental state categories such as emotions and physical states may be even more difficult to map onto lexical items than physical characteristics. In contrast to 3- and 4-year-olds, 5-year-olds appeared to rely more heavily on the situational context and realized that all sentence frames were grammatically consistent with an emotion interpretation.

One explanation of our findings is that they are separately driven by the development of emotional understanding and syntactic bootstrapping. That is, these processes could be truly interacting as separate phenomena. If this is true, children develop the ability to represent discrete emotion concepts alongside the separate ability to use syntactic bootstrapping to infer meanings of novel predicate words. Another possibility is that both of these abilities—the representation of emotion concepts and the acquisition of emotion words— are constrained by the more general ability to make mental state inferences. Without the ability to draw mental state inferences, 3-year-olds are unlikely to correctly understand what type of mental state the alien in the story is experiencing. By the same token, without the ability to draw more complex mental state inferences about others' communicated intentions, 3-year-olds may be unable to use sentence frame information to map a novel word onto a particular mental state meaning. Although research suggests that emotion understanding and mental state inferences such as beliefs may follow a similar developmental trajectory (Bretherton & Beeghly, 1982; Cutting & Dunn, 1999; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Harris et al., 1989; Weimer, Sallquist, & Bolnick, 2012), more research should examine the simultaneous development of mental state inference, emotional understanding, and syntactic bootstrapping and the extent to which they rely on similar vs. distinct cognitive processes.

Limitations and Future Directions

Given that these are the first studies to our knowledge to examine how children learn novel emotion word meanings, they are not without limitations. One limitation was that we used sentence frames that may not reflect those used most frequently in daily life. In daily life, emotion words are presented to children in a number of different sentence frames by parents, teachers and peers. It is possible that within our targeted age range, children are not exposed to emotion words within *is, feels or feels about* sentence frames with equal frequency. Thus, it is important to examine corpora of child-directed speech to empirically examine the most frequent linguistic input and output used when adults and children are discussing emotions (see Shablack, Stein, Lindquist & Becker, in prep). These data-driven findings might then be used in future research to examine the variety of sentence frames used by children and their parents to talk about emotion concepts in daily life.

Another limitation of our study concerns the sample used in the present studies. In both Study 1 and Study 2, a large portion of the participants are from high socioeconomic households. Evidence suggests that socioeconomic status is influential on learning and developmental trajectories in general (Bradley & Corwyn, 2002; Hoff, 2003). The sample and findings should thus be extended to other populations, where both emotional and linguistic development might be more variable. Additionally, based on conversations with parents/legal guardians following the study, some children were involved in school programs focusing on emotion development and understanding, which may have improved their performance in our task, overall.

Another limitation of our study was the stimuli we used. We focused on alien stimuli in the present experiments for several reasons. First, we were concerned that using human stimuli might cause children to infer that a new word could not name a human emotion concept that they already knew a word for (e.g., *binty* could not refer to the human concept *happiness* since most children know that the word *happy* names this concept by ages 3-5). Additionally, we reasoned that alien stimuli would be maximally entertaining and engaging to children in our sample. However, as we noted in Study 1, children could have had problems understanding the intended meaning of the cartoons. Our validation study, children's spontaneous verbal expressions in Study 1 and choice behavior in Study 2 suggest that this possibility is of limited concern. Yet a broader concern is that cartoons of aliens depicting emotional expressions are not ecologically valid and thus limit the inferences that can be drawn about how children learn about the words that correspond to real human emotional expressions. This being said, the ecological validity of the posed, caricatured human facial expressions used in most psychology experiments is also questionable (Nelson & Russell, 2011; Quigley, Lindquist, & Barrett, 2014). One interpretation of these posed facial expressions is that they are more like symbols than veridical representations of what people do with their faces in daily life (Adams, Albohn, & Kveraga, 2016; Gendron, Mesquita, & Barrett, 2013; Jack, Garrod, Yu, Caldara, & Schyns, 2012). Thus, there may be greater parallels between our studies and studies using human facial stimuli than appears at first glance. Nonetheless, in future research, it is important to replicate and extend our findings with ecologically valid images of human emotional facial expressions, actions and physical states.

In addition to these future directions, future research might consider existing individual differences that influence the learning of novel emotion concept words. For example, as emotion development is correlated with linguistic and verbal ability, future studies might gather validated performance-based measures of children's language ability, such as the MacArthur Communicative Development Inventory score (Fenson et al., 2007) or Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), rather than parent report. It would also be interesting to know whether individuals who know multiple languages, who may be more sensitive to different sentence frames and more adept at disambiguating the meaning of novel words across multiple languages, differ in their abilities to infer that novel words refer to emotions. To the extent that positive rearing environments confer more opportunities for caregiver discourse about emotion, it would also be interesting to explore the extent to which adversity predicts different outcomes in parents' linguistic framing of emotion and the impact on later outcomes in emotional understanding.

Implications

Although this line of work is new, it has important implications for the role of language in children's emotion understanding, communication, and in their ability to perceive emotions in others or experience them firsthand (Lindquist, 2017; Lindquist et al., 2016; Lindquist, MacCormack, et al., 2015; Lindquist, Satpute, et al., 2015). Above all, this work sheds new light on how children are learning about social categories and using them to make meaning of the world around them. Like the research before it, our findings suggest that drawing inferences about emotion concepts may be a gradual process that occurs over the course of early childhood and relies on both the use of language (Widen & Russell, 2008; Widen, 2013), caregiver communication (Cutting & Dunn, 1999; Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski, et al., 1991) and an understanding of the situational context (Widen & Russell, 2010, 2011).

Our work thus has important implications for how caregivers and parents can use both language and the present situation to aid children in understanding adjectives, including abstract adjectives such as emotion words. Increasing emotion understanding through language is an important part of development as it can lead to better social outcomes, such as successful communication about one's own and others' emotional states, in turn leading to better interpersonal relationships, classroom environments, work environments and leadership (Brackett et al., 2013; Hagelskamp et al., 2013; Rivers, Brackett, Reyes, Elbertson, & Salovey, 2013), less risky behavior (Rivers, Brackett, Omori, et al., 2013), improved grades (Brackett, Rivers, Reyes, & Salovey, 2012), less social isolation (Twenge et al., 2003) and more prosocial behavior (Eggum et al., 2011). Understanding a larger range of emotion concepts is associated with greater emotion differentiation, which is also associated with many positive social outcomes such as lower levels of stress, better emotion regulatory strategies, and overall positive well-being (Kashdan, Barrett & McKnight, 2015; Lindquist & Barrett, 2008). This work can also be applied to curricula aimed at enhancing children's learning of emotion concepts (e.g., Nathanson, Rivers, Flynn, & Brackett, 2016; Rivers, Tominey, O'Bryon, & Brackett, 2013; Weimer et al., 2012). We look forward to future work examining how children learn about emotions via language, and interventions that aim to harness this phenomenon to increase emotion understanding.

Acknowledgements

This research was supported by a grant from the Fostering Interdisciplinary Research Explorations program at UNC Chapel Hill, awarded to K. Lindquist and M. Becker. We also thank all of the research assistants who helped finalize study materials and run participants. A subset of these findings were presented at the Society for Affective Science Conference in 2018 and Cognitive Development Society in 2019.

Context paragraph

Lindquist and Shablack study the role of language in emotion perception and experiences across the lifespan and Becker examines the acquisition of predicates and word meanings in children. Upon meeting at a University panel, we realized how little research explores how children acquire emotion words in the first place. Thus, we combined our interdisciplinary expertise in emotion and linguistic development.

Works Cited

Adams, R. B. J., Albohn, D. N., & Kveraga, K. (2016). A social vision account in facial expression perception. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The Science of Facial Expression*. New York: Oxford.

Ames, D. R. (2004). Inside the mind reader's toolkit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology*, *87*(3), 340-353. https://doi.org/10.1037/0022-3514.87.3.340

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446. https://doi.org/10.1016/j.cognition.2009.10.015

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311–1320. https://doi.org/10.1037//0012-1649.35.5.1311

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger Syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175.

Baron-Cohen, S. (1991). Do people with autism understand what causes emotions? *Child Development, 62*(2), 385-395. https://doi.org/10.1111/j.1467-8624.1991.tb01539.x

Barrett, L. F. (2006). Solving the emotion paradox: categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2013). Psychological construction: the Darwinian approach to the science of emotion. *Emotion Review*, *5*(4), 379–389. https://doi.org/10.1177/1754073913489753

Barrett, L. F., & Russell, J. A. (2015). An introduction to psychological construction. In L. F.

Barrett & J. a. Russell (Eds.), *The Psychological Construction of Emotion* (pp. 1–17). New York: Guilford.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(1), 85-123. https://doi.org/10.1017/S0305000900008680

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry, 37*(3), 441-456. https://doi.org/10.1162/ling.2006.37.3.441

Becker, M. (2014). *The Acquisition of Syntactic Structure: Animacy and Thematic Alignment*. Cambridge: Cambridge University Press.

Becker, M. (2015). Animacy and the acquisition of tough adjectives. *Language Acquisition, 22*(1), 68-103. https://doi.org/10.1080/10489223.2014.928298

Becker, M., & Estigarribia, B. (2013). Harder words: How children learn abstract verbs with opaque syntax. *Language Learning and Development*, *9*, 211–244.

Berenbaum, H., & Oltmanns, T. F. (1992). Emotional experience and expression in schizophrenia and depression. *Journal of Abnormal Psychology, 101*(1), 37-44. https://doi.org/10.1037/0021-843X.101.1.37

Bloom, P. (2000). How do children learn the meanings of words? In *How do children learn the meanings of words*.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development, 4*(3), 357-381.

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development, 80*(1), 15-22.

https://doi.org/10.1111/j.1467-8624.2008.01242.x

Brackett, M. A., Bertoli, M., Elbertson, N., Bausseron, E., Castillo, R., & Salovey, P. (2013). Emotional intelligence: Reconceptualizing the cognition-emotion link. In M. D. Robinson, E. R. Watkins, & E. Harmon-Jones (Eds.), *Handbook of cognition and emotion*. New York: Guilford Press.

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*(2), 218–224. https://doi.org/10.1016/j.lindif.2010.10.002

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*(6), 906–921. https://doi.org/10.1037/0012-1649.18.6.906

Bridges, K. M. B. (1932). Emotional development in early infancy. *Child Development*, *3*(4), 324. https://doi.org/10.2307/1125359

Campos, J. J., Frankel, C. B., & Camras, L. (2004). On the nature of emotion regulation. *Child Development, 75*(2), 377-394. https://doi.org/10.1111/j.1467-8624.2004.00681.x

Camras, L. A. (1992). Expressive development and basic emotions. *Cognition and Emotion*, *6*(3/4), 269–283. https://doi.org/10.1080/02699939208411072

Casasanto, D. (2008). Who’s afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. In *Language Learning*. https://doi.org/10.1111/j.1467-

9922.2008.00462.x

Casasanto, D., & Bottini, R. (2014). Spatial language and abstract concepts. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(2), 139–149. https://doi.org/10.1002/wcs.1271

Castro, V. L., Halberstadt, A. G., Lozada, F. T., & Craig, A. B. (2015). Parents' emotion-related beliefs, behaviors, and skills predict children's recognition of emotion. *Infant and Child Development*, *24*(1), 1–22. https://doi.org/10.1002/icd.1868

Chen J. J. (2003). Communicating complex information: the interpretation of statistical interaction in multiple logistic regression analysis. *American Journal of Public Health, 93*(9), 1376-7; author reply 1377.

Clark, E. (1987). The principle of contrast: A constraint on language acquisition. *Mechanisms of Language Acquisition*.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, *70*(4), 853–865. https://doi.org/10.1111/1467-8624.00061

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience, 9*(1), 28. https://doi.org/10.1038/nn1611

de Rosnay, M., & Harris, P. L. (2002). Individual differences in children's understanding of emotion: the roles of attachment and language. *Attachment & Human Development*, *4*(1), 39–54. https://doi.org/10.1080/14616730210123139

de Rosnay, M., Pons, F., Harris, P. L., & Morrell, J. M. B. (2004). A lag between understanding false belief and emotion attribution in young children: Relationships with linguistic ability and mothers' mental-state language. *British Journal of Developmental Psychology*, *22*(2), 197–218. https://doi.org/10.1348/026151004323044573

Denham, S. A., Zoller, D., & Couchoud, E. A. (1994). Socialization of preschoolers' emotion understanding. *Developmental Psychology*, *30*(6), 928–936. https://doi.org/10.1037//0012-1649.30.6.928

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. *Journal of Experimental Psychology: General*, *147*, 62–73.

Dunn, J., Brown, J. R., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology*, *27*(3), 448–455. https://doi.org/10.1037/0012-1649.27.3.448

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Development*, *62*(6), 1352–1366. https://doi.org/10.1111/j.1467-8624.1991.tb01610.x

Dunsmore, J. C., & Halberstadt, A. G. (1997). How does family emotional expressiveness affect children's schemas? In K. C. Barrett (Ed.), *The communication of emotion: Current research from diverse perspectives* (pp. 45–68). San Francisco: Jossey-Bass.

Eggum, N. D., Eisenberg, N., Kao, K., Spinrad, T. L., Bolnick, R., Hofer, C., … Fabricius, W. V. (2011). Emotion understanding, theory of mind, and prosocial orientation: Relations over time in early childhood. *Journal of Positive Psychology*, *6*(1), 4–16. https://doi.org/10.1080/17439760.2010.536776

Eisenberg, N., Sadovsky, A., & Spinrad, T. L. (2005). Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. *New Directions for Child and Adolescent Development*, *2005*(109), 109–118. https://doi.org/10.2964/jsik.kuni0223

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Ellis, B. H., Alisic, E., Reiss, A., Dishion, T., & Fisher, P. A. (2014). Emotion regulation among preschoolers on a continuum of risk: the role of maternal emotion coaching. *Journal of Child and Family Studies*, *23*(6), 965–974. https://doi.org/10.1007/s10826-013-9752-z

Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American Journal of Psychiatry, 164*(1), 1476-1488. https://doi.org/10.1176/appi.ajp.2007.07030504

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55-64. https://doi.org/10.1111/1467-7687.00209

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23*(3), 331-392. https://doi.org/10.1016/0010-0285(91)90013-E

Fivush, R. (1991). Gender and emotion in mother—child conversations about the past. *Journal of Narrative and Life History*, *1*(4), 325–341.

Fivush, R., Brotman, M. A., Buckner, J. P., & Goodman, S. H. (2000). Gender differences in parent–child emotion narratives. *Sex Roles*, *42*, 233–253. https://doi.org/10.1023/A:1007091207068

Fivush, R., Haden, C. A., & Reese, E. (2006). Elaborating on elaborations: Role of maternal reminiscing style in cognitive and socioemotional development. *Child Development*, *77*(6), 1568–1588. https://doi.org/10.1111/j.1467-8624.2006.00960.x

Fogel, A., Nwokah, E., Dedo, J. Y., Messinger, D., Dickson, K. L., Matusov, E., & Holt, S. A. (1992). Social process theory of emotion: A dynamic systems approach. *Social Development*, *1*(2), 122–142. https://doi.org/10.1111/j.1467-9507.1992.tb00116.x

Garrett-Peters, P. T., Castro, V. L., & Halberstadt, A. G. (2017). Parents' beliefs about children's emotions, children's emotion understanding, and classroom adjustment in middle childhood. *Social Development*, *26*(3), 575–590. https://doi.org/10.1111/sode.12222

Gendron, M., Mesquita, B., & Barrett, L. F. (2013). Emotion perception: Putting the face in context. In D. Reisberg (Ed.), *Oxford library of psychology. The Oxford handbook of cognitive psychology* (pp. 539–556). New York: Oxford University Press.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*, *14*(2), 251–262. https://doi.org/10.1037/a0036052

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Language Development: Vol. 2. Language, Thought, and Culture*. https://doi.org/10.1109/TAIC-PART.2006.2

Gerken, L., & Shady, M. E. (1996). The picture selection task. In *Methods for assessing children's syntax. Language, speech, and communication*.

Girden, E. R. (1992). *ANOVA: Repeated measures* (No. 84). Sage.

Gleitman, L. Cassidy, K. Nappa, R., Papafragou, A. & Trueswell, J. (2005). Hard words. *Language Learning and Development, 1*, 23-64. https://doi.org/10.1038/sj.bdj.4803768

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*. https://doi.org/10.1207/s15327817la0101_2

Hagelskamp, C., Brackett, M. A., Rivers, S. E., & Salovey, P. (2013). Improving classroom quality with the ruler approach to social and emotional learning: proximal and distal outcomes. *American Journal of Community Psychology*, *51*(3–4), 530–543. https://doi.org/10.1007/s10464-013-9570-x

Halberstadt, A. G., Denham, S. A., & Dunsmore, J. C. (2001). Affective social competence. *Social Development*, *10*(1), 79–119. https://doi.org/10.1111/1467-9507.00150

Halberstadt, A. G., & Lozada, F. T. (2011). Emotion development in infancy through the lens of culture. *Emotion Review*, *3*(2), 158–168. https://doi.org/10.1177/1754073910387946

Harris, P. L., De Rosnay, M., & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science*, *14*(2), 69 -73. https://doi.org/10.1111/j.0963-7214.2005.00337.x

Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition and Emotion*, *3*(4), 379–400. https://doi.org/10.1080/02699938908412713

Hirsh-Pasek, K., & Golinkoff, R. (1996). *The Origins of Grammar*. Cambridge: MIT Press.

Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

Hutchinson, J. E. (1986). Children's sensitivity to the contrastive use of object category terms. *Papers and Reports on Child Language Development*, *25*, 49–56.

Huttenlocher, J., & Smiley, P. (1987). Early word meanings: The case of object names. *Cognitive Psychology, 19*(1), 63-89. https://doi.org/10.1016/0010-0285(87)90004-1

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, *2*(3), 260–280. https://doi.org/10.1111/j.1745-6916.2007.00044.x

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, *109*(19), 7241–7244. https://doi.org/10.1073/pnas.1200155109

Joormann, J. (2010). Cognitive inhibition and emotion regulation in depression. *Current Directions in Psychological Science, 19*(3), 161-166. https://doi.org/10.1177/0963721410370293

Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). Unpacking emotion differentiation: transforming unpleasant experience by perceiving distinctions in negativity. *Current Directions in Psychological Science*, *24*(1), 10–16. https://doi.org/10.1177/0963721414550708

Kayyal, M., Widen, S., & Russell, J. A. (2015). Context is more powerful than we think: Contextual cues override facial cues even for valence. *Emotion*, *15*(3), 287–291. https://doi.org/10.1037/emo0000032

La Barbera, J. D., Izard, C. E., Vietze, P., & Parisi, S. A. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, *47*, 535–538. https://doi.org/10.2307/1128816

Lewis, M., & Brooks, J. (1978). Self-knowledge and emotional development. In *The Development of Affect* (pp. 205–226). Boston: Springer.

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

Lindquist, K. A. (2017). The role of language in emotion: existing evidence and future directions. *Current Opinion in Psychology, 17,* 135-139. https://doi.org/10.1016/j.copsyc.2017.07.006

Lindquist, K. A., & Barrett, L. F. (2008). Emotional complexity. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed.). New York: Guilford.

Lindquist, K. A., Gendron, M., & Satpute, A. (2016). Language and emotion: Putting feelings into words and words into feelings. In M. Lewis, J. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (4th ed.). New York: Guildford.

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology*, *6*(MAR). https://doi.org/10.3389/fpsyg.2015.00444

Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science*, *24*(2), 99–108. https://doi.org/10.1177/0963721414553440

Lupyan, G. (2012a). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology*, *3*(MAR), 54. https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G. (2012b). What do words do? Toward a theory of language-augmented thought. In *Psychology of Learning and Motivation* (Vol. 57, pp. 255–297). Elsevier.

https://doi.org/10.1016/B978-0-12-394293-7.00007-8

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*(1), 57-77. https://doi.org/10.1016/0364-0213(90)90026-S

Mennin, D. S., McLaughlin, K. A., & Flanagan, T. J. (2009). Emotion regulation deficits in generalized anxiety disorder, social anxiety disorder, and their co-occurrence. *Journal of Anxiety Disorders, 23*(7), 866-871. https://doi.org/10.1016/j.janxdis.2009.04.006

Murphy, F. C., Sahakian, B. J., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., & Paykel, E. S. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychological Medicine, 29*(6), 1307-1321. https://doi.org/10.1017/S0033291799001233

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357-374. https://doi.org/10.1017/S0305000900013817

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition, 58*(2), 221-251. https://doi.org/10.1016/0010-0277(95)00681-8

Nathanson, L., Rivers, S. E., Flynn, L. M., & Brackett, M. A. (2016). Creating emotionally intelligent schools with RULER. *Emotion Review*, *8*(4), 305–310. https://doi.org/10.1177/1754073916650495

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology*, *110*(1), 52–61. https://doi.org/10.1016/j.jecp.2011.03.014

Panksepp, J. (1998). The preconscious substrates of consciousness: Affective states and the evolutionary origins of the self. *Journal of Consciousness Studies*, *5*(5–6), 566–582.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition, 105*(1), 125-165. https://doi.org/10.1016/j.cognition.2006.09.008

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biological Psychiatry, 54*(5), 515-528. https://doi.org/10.1016/S0006-3223(03)00171-9

Pinker, S. (1984). *Language Learnability and Language Development*. https://doi.org/10.1007/s13398-014-0173-7.2

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*. https://doi.org/10.1111/j.1471-6402.2000.tb01024.x

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. https://doi.org/10.1080/17405620344000022

Pons, F., Lawson, J., Harris, P. L., & De Rosnay, M. (2003). Individual differences in children's emotion understanding: Effects of age and language. *Scandinavian Journal of Psychology*, *44*(4), 347–353. https://doi.org/10.1111/1467-9450.00354

Quigley, K. S., Lindquist, K. A., & Barrett, L. F. (2013). Inducing and measuring emotion and affect: tips, tricks, and secrets. *Handbook of Research Methods in Social and Personality Psychology*. https://doi.org/http://dx.doi.org/10.1017/CBO9780511996481.014

Ridgeway, D., Waters, E., & Kuczaj, S. a. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology*, *21*(5), 901–908. https://doi.org/10.1037//0012-1649.21.5.901

Rivers, S. E., Brackett, M. A., Reyes, M. R., Elbertson, N. A., & Salovey, P. (2013). Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing the RULER approach. *Prevention Science*, *14*(1), 77–87. https://doi.org/10.1007/s11121-012-0305-2

Rivers, S. E., Brackett, M., Omori, M., Sickler, C., Bertoli, M. C., & Salovey, P. (2013). Emotion skills as a protective factor for risky behaviors among college students. *Journal of College Student Development*, *54*(2), 172–183. https://doi.org/10.1353/csd.2013.0012

Rivers, S. E., Tominey, S. L., O'Bryon, E. C., & Brackett, M. A. (2013). Developing emotional skills in early childhood settings using Preschool RULER. *The Psychology of Education Review*, *37*, 19–25.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *48*(5), 1290–1298. https://doi.org/10.1037/0022-3514.48.5.1290

Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology*, *22*(1), 97–102. https://doi.org/10.1037/0012-1649.22.1.97

Russell, J. A., & Widen, S. C. (2002a). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*(1), 30–52. https://doi.org/10.1111/1467-9507.00185

Russell, J. A., & Widen, S. C. (2002b). Words versus faces in evoking preschool children's knowledge of the causes of emotions. *International Journal of Behavioral Development*, *26*(2), 97–103. https://doi.org/10.1080/01650250042000582

Salters-Pedneault, K., Roemer, L., Tull, M. T., Rucker, L., & Mennin, D. S. (2006). Evidence of broad deficits in emotion regulation associated with chronic worry and generalized anxiety disorder. *Cognitive Therapy and Research, 30*(4), 469-480. https://doi.org/10.1007/s10608-006-9055-4

SAS Institute Inc. 2013. SAS 9.4 Statements: Reference. Cary, NC: SAS Institute Inc.

Shablack, H. (2017). Linguistic and situational context assists emotion concept acquisition in preschoolers (unpublished master's thesis). University of North Carolina at Chapel Hill.

Shablack, H., & Lindquist, K. A. (in press). The role of language in the development of emotion. In V. LoBue, K. Perez-Edgar, & K. Buss (Eds.), *Handbook of Emotion Development*. New York: Springer.

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: a case study for colour. *The Behavioral and Brain Sciences*, *28*(4), 469-89; discussion 489-529. https://doi.org/10.1017/S0140525X05000087

Syrett, K., Latourrette, A., Ferguson, B., & Waxman, S. R. (2018). Crying helps, but being sad doesn't: Verbs, but not adjectives, constrain referent selection for novel nouns through age three. In *Paper presented at the Boston University Conference on Language Development*.

Syrett, K., & Lidz, J. (2010). 30-month-olds use the distribution and meaning of adverbs to interpret novel adjectives. *Language Learning and Development, 6*(4), 258-282.

Twenge, J. M., Catanese, K. R., & Baumeister, R. F. (2003). Social exclusion and the deconstructed state: time perception, meaninglessness, lethargy, lack of emotion, and self-awareness. *Journal of Personality and Social Psychology*, *85*(3), 409–423. https://doi.org/10.1037/0022-3514.85.3.409

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94*(1), 3. https://doi.org/10.1037/0033-295X.94.1.3

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: individual variation in action identification. *Journal of Personality and Social Psychology, 57*(4), 660-671. https://doi.org/10.1037/0022-3514.57.4.660

Wang, A. T., Dapretto, M., Hariri, A. R., Sigman, M., & Bookheimer, S. Y. (2004). Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(4), 481–490.

Wegner, D. M., & Vallacher, R. R. (1987). The trouble with action. *Social Cognition, 5*(3), 179-190. https://doi.org/10.1521/soco.1987.5.3.179

Weimer, A. A., Sallquist, J., & Bolnick, R. R. (2012). Young children's emotion comprehension and theory of mind understanding. *Early Education and Development*, *23*(3), 280–301. https://doi.org/10.1080/10409289.2010.517694

Weinberg, M. K., Tronick, E. Z., Cohn, J. F., & Olson, K. L. (1999). Gender differences in emotional expressivity and self-regulation during early infancy. *Developmental Psychology*, *35*(1), 175–188. https://doi.org/10.1037//0012-1649.35.1.175

Wellman, H. M., Harris, P. L., Banerjee, M., & Sinclair, A. (1995). *Early Understanding of Emotion: Evidence from Natural Language*. *Cognition & Emotion* (Vol. 9). https://doi.org/10.1080/02699939508409005

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

https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*(1), 114–128. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*(2), 291–312. https://doi.org/10.1016/j.cogdev.2008.01.002

Widen, S. C., & Russell, J. A. (2010). Children's scripts for social emotions: Causes and consequences are more central than are facial expressions. *British Journal of Developmental Psychology*, *28*(3), 565–581. https://doi.org/10.1348/026151009X457550d

Widen, S. C., & Russell, J. A. (2011). In building a script for an emotion, do preschoolers add its cause before its behavior consequence? *Social Development*, *20*(3), 471–485. https://doi.org/10.1111/j.1467-9507.2010.00594.x

Widen, S. C., & Russell, J. A. (2013). Children's recognition of disgust in others. *Psychological Bulletin*, *139*(2), 271–299. https://doi.org/10.1037/a0031640

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, *49*(5), 1105–1127. https://doi.org/10.1016/j.neuropsychologia.2010.12.032

Xu, F., & Kushnir, T. (2013). Infants are rational constructivist learners. *Current Directions in Psychological Science*, *22*(1), 28–32. https://doi.org/10.1177/0963721412469396

Yuan, S., & Fisher, C. (2009). "really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*(5), 619-626. https://doi.org/10.1111/j.1467-9280.2009.02341.x

Table 1

Examples of syntactic structure cueing possible meaning

Number of Arguments (NP)	Examples	Verb Denotation
Single NP	John slept; Mary is running	An action/event that an individual engages in by themselves
Two NPs	Sue kissed the baby; Kevin hugged his friend	An action/event that one individual does to another (i.e., the object is affected by the action)
Three NPs	Katrina gave Emily a pen; Roger told Marvin a story	Generally means something about transfer or communication

Table 2

Study 1 sentence frames and general puppet video dialogue.

Sentence frame	**General script**
Is	Puppet 1: I know an alien who is [novel word]. Puppet 2: Really? You know an alien who is [novel word]? Puppet 1: Yes! This alien is [novel word] Puppet 2: Wow! You know an alien who is [novel word]
Feels	Puppet 1: I know an alien who feels [novel word]. Puppet 2: Really? You know an alien who feels [novel word]? Puppet 1: Yes! This alien feels [novel word] Puppet 2: Wow! You know an alien who feels [novel word]
Feels About	Puppet 1: I know an alien who feels [novel word] about brushing his teeth. Puppet 2: Really? You know an alien who feels [novel word] about brushing his teeth.? Puppet 1: Yes! This alien feels [novel word] about brushing his teeth. Puppet 2: Wow! You know an alien who feels [novel word] about brushing his teeth.
Fillers	Puppet 1: I saw an alien who was [novel word]. Puppet 2: Really? You saw an alien who was [novel word]? Puppet 1: Yeah! I saw an alien who was [novel word] Puppet 2: Oh! You saw an alien who was [novel word]

Table 3

Image stimuli for Study 1 and Study 2. Four cartoon alien identities exhibited the listed emotion, action or physical state.

Emotions	happy, excited, sad, mad, scared, disgusted, surprised
Actions	sleeping, jumping, sitting, falling, cartwheeling, walking, running, eating pizza/fruit*, swimming*
Physical states	itchy, hot, cold, sick, burned, hungry, hurt

Note. Screening trials included an emotion image and two action images. Action images that were only included in screening trials are indicated with a *

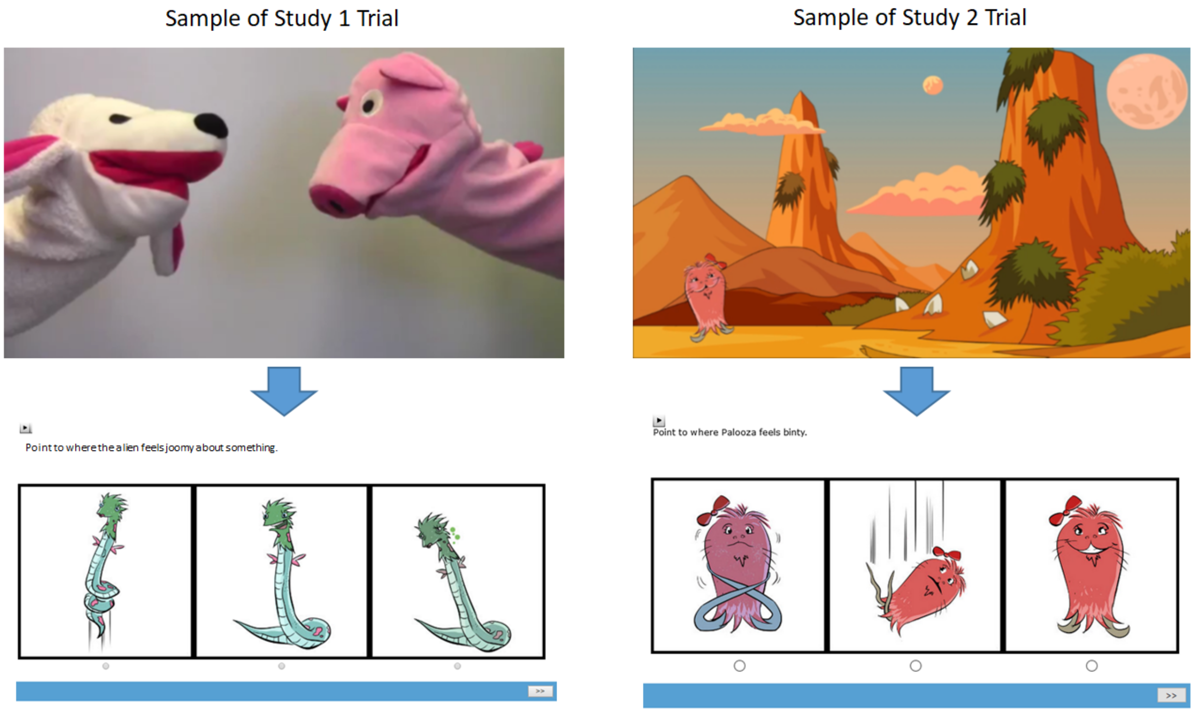


Figure 1. Sample trials for Study 1 and Study 2. For the text of the Study 1 video, see Table 2. In Study 2, children heard a 2-3 sentence story, see Table 4 for examples. Each video was played on a screen by itself, and the next screen presented 3 randomized images of the alien character, accompanied by audio instructions for the child to point to the image they believe depicts the novel word.

Table 4

Study 1 within subjects main effects and interactions for 3 (Image choice: emotion, state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with participant gender, screening trial performance, and filler trial performance as covariates.

	df	*F*	η^2	*p*
Image choice	2	3.10	.03	.05
Image choice x Screening trial performance	2	4.23	.03	.02
Image choice x Gender	2	2.65	.02	.07
Image choice x Filler performance	2	3.55	.03	.03
Image choice x Age	4	2.53	.04	.04
Image choice x Sentence frame	4	2.28	.04	.06
Image choice x Age x Sentence frame	8	1.28	.04	.25

Note. Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 6.36$, $p = .04$, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, ∑ = 0.95, being greater than .75, $df_{error} = 246$.

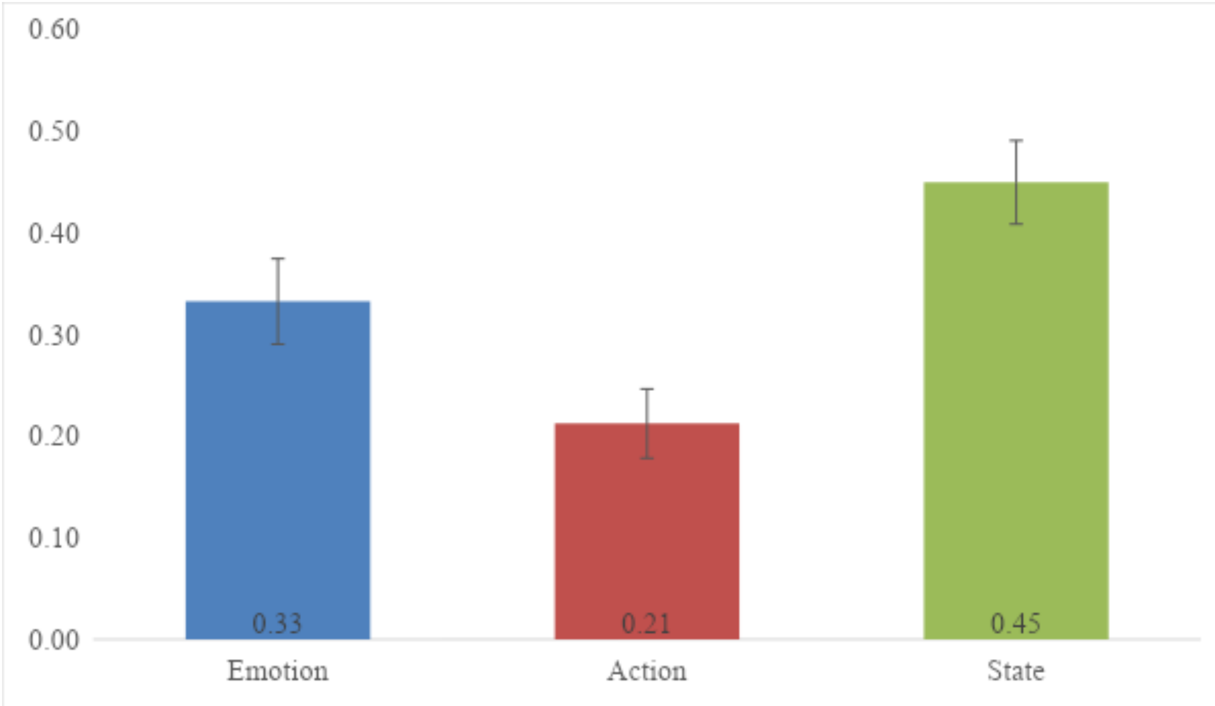


Figure 2. Study 1 estimated marginal means of each Image choice. Bars represent 95% confidence intervals.

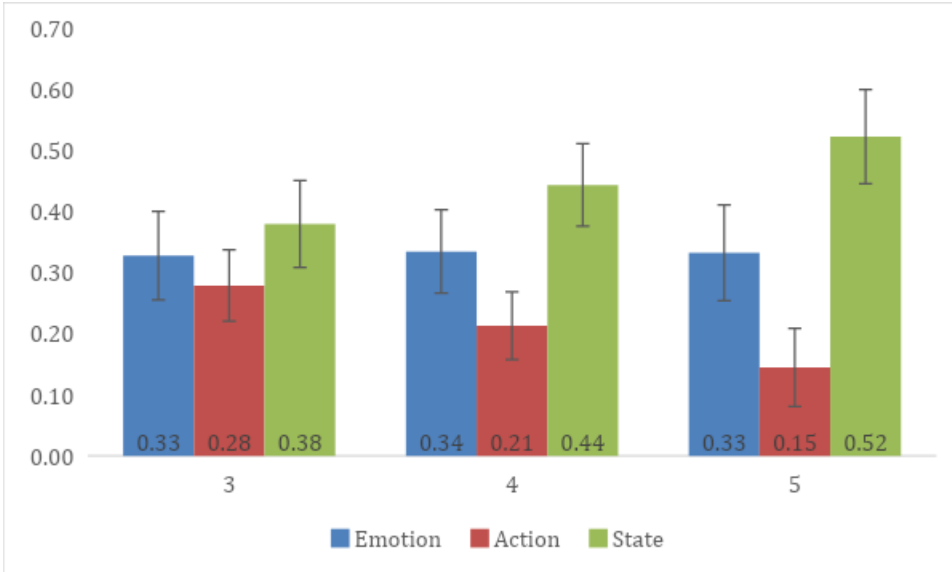


Figure 3. Study 1 estimated marginal means of each Image choice by Age in years. Bars represent 95% confidence intervals.

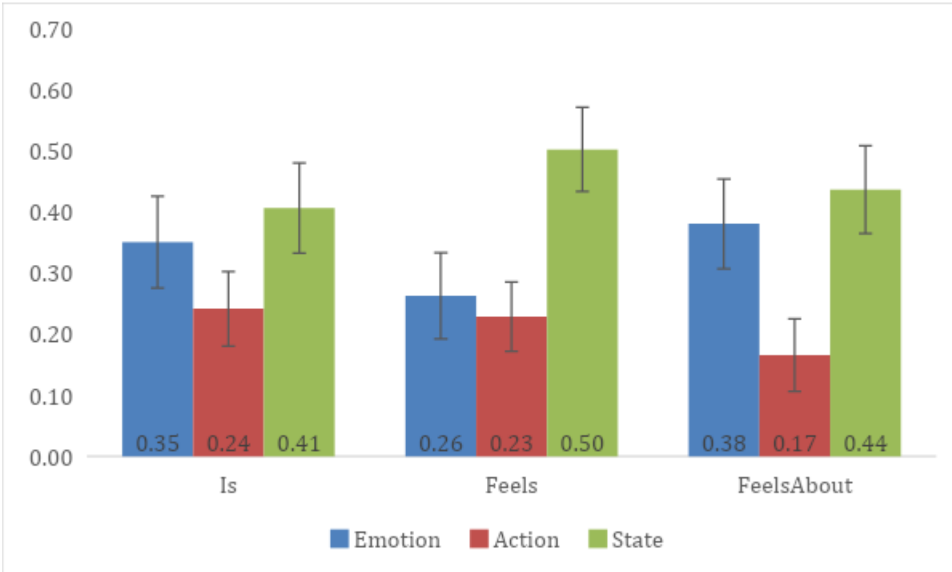


Figure 4. Study 1 estimated marginal means of each Image choice by Sentence frame. Bars represent 95% confidence intervals.

Table 5.

Sample stories for each emotion type of Study 2 and novel words

Emotion	**Story**
Happy	Palooza ran a race at school. She was the fastest alien in the race, so she won first place! Now, Palooza [sentence frame] [novel word]. What do you think [novel word] means?
Sad	Xylobean's best friend moved away to a different planet, and they won't see each other again. Now, Xylobean [sentence frame] [novel word]. What do you think [novel word] means?
Mad	Chromia was reading her favorite book. Then, another alien took it and tore out a page! Now, Chromia [sentence frame] [novel word]. What do you think [novel word] means?
Surprised	One day Wazu came home and all his furniture was turned upside-down. He just stared at his furniture and couldn't figure out how that happened. Now, Wazu [sentence frame] [novel word]. What do you think [novel word] means?
Disgusted	Palooza took a bite of an apple. As soon as she bit into it, she realized it was rotten inside. She didn't want to eat the rest of it. She threw it in the trash. Now, Palooza [sentence frame] [novel word]. What do you think [novel word] means?
Afraid	Wazu heard a loud crashing noise in the distance. Then, the sound started getting closer and closer! Now, Wazu [sentence frame] [novel word]. What do you think [novel word] means?
Excited	Chromia always wanted to fly in a spaceship. Now she was going to get a chance to do it! Now, Chromia [sentence frame] [novel word]. What do you think [novel word] means?

Table 6

Study 2 main effects and interactions for 3 (Image choice: emotion, state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA including participant gender and screening trial performance as covariates.

	df	*F*	η^2	*p*
Image choice	1.87	1.71	.02	.19
Image choice x Screening trial performance	1.87	3.92	.04	.03
Image choice x Gender	1.87	.44	.004	.63
Image choice x Age	3.73	8.29	.14	.001
Image choice x Sentence frame	3.73	1.30	.03	.27
Image choice x Age x Sentence frame	7.47	2.49	.09	.02

Note. Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 21.81$, $p < .001$, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, ∑ = 0.837, being greater than .75, $df_{error} = 190.38$.

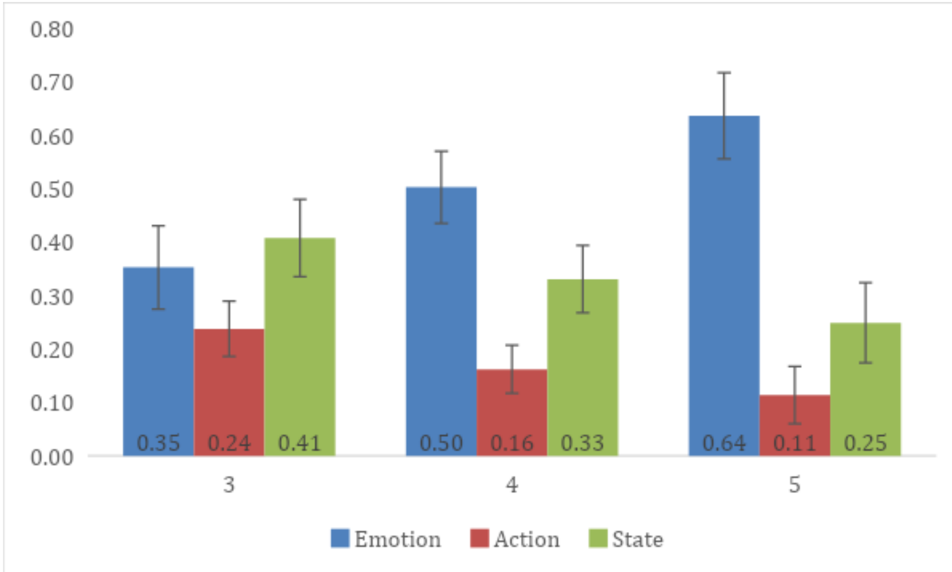


Figure 5. Study 2 estimated marginal means of each Image choice by Age in years. Bars represent 95% confidence intervals.

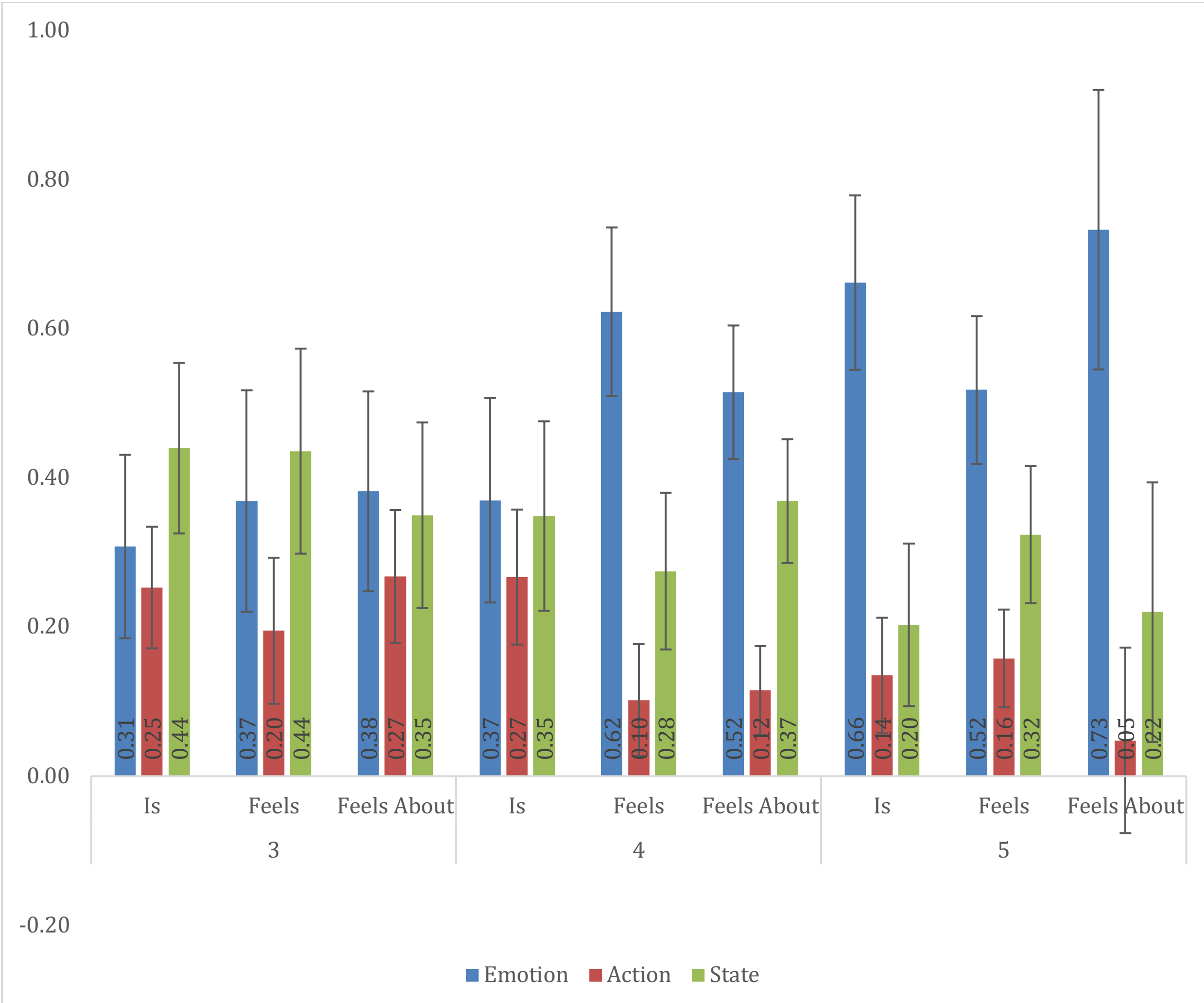


Figure 6. Study 2 3-way interaction between Image Choice, Age in years and Sentence Frame. Bars represent 95% confidence intervals.

Supplemental Material

This supplement reports several additional and follow-up analyses to supplement the main text including: 1) validation results of the image stimuli, 2) results on Filler trial performance from Study 1, and 3) Supplemental analyses of our main hypotheses for both Studies 1 and 2.

Validation of alien image stimuli

To assess whether children were able to infer differential meaning across images types, we conducted a validation study with 17 children (M_{age} = 3.82; SD_{age} = .64; 11 female). The majority were 4-year olds (n=10), followed by 3-year olds (n=5), and 5-year olds (n=2).

Similar recruiting methods were used as in Studies 1 and 2. Following parental consent and child assent, children were randomly presented nine trials in which they were presented with three images (one emotion, one physical state, one action) and asked to "Point to the alien that is [target image]". The target image always matched the instructions and the other images were randomly selected from each of the two remaining categories. For example, if the target image was running, one image was of an alien running and the other two images were a random emotion image and random physical state image of the same alien. Children completed three trials of each image type as a target (e.g., three trials where the target was an action, three where the target was an emotion, and three where the target was a physical state). Study participation lasted at most four minutes.

Frequencies of correct responses for each image individually and for each image type were computed. Proportions of correct choices were calculated based on the number of completed trials for each image type (i.e., if a child answered 2 out of 3 action image trials correctly their action image proportion was .67)[12] (see Supplemental Figure 1).

Proportion correct of image types were then entered into a within-subjects ANOVA to assess whether children were significantly more or less accurate on one image type over another. There were no significant differences in proportion correct between image types, $F(2, 32) = 1.64$, $p = .21$, $\eta^2 = .09$. This suggests that children had a general understanding of the images and no particular image type was easier or harder to understand.

Proportion correct of image types were also entered into a set of t-tests to assess whether children performed above chance (greater than .33). All proportion correct of image types were significantly above chance, meaning that overall children were choosing the correct image across trials, independent of image type (see Supplemental Table 1). However, we note that physical state images had the lowest proportion correct ($M = .69$; $SE = .08$). Thus, we conducted a set of t-tests for each individual image (Supplemental Table 1). We found that children performed the worst when asked to select the “sick” image ($M = .43$; $SE = .20$) and “hurt” image ($M = .63$; $SE = .18$). An error analysis revealed the images that children chose when they were incorrect on “sick” and “hurt” trials. For the sick trials, two children chose the image depicting disgust, one child chose the falling image and one child chose the image depicting sadness. For the hurt trials, one child chose the cartwheeling image, one child chose the image depicting disgust and one child the image depicting sadness. As we mention in the main manuscript, it is possible that for these images, children were conflating alternate visual depictions of the target, such that when an

[12] One child completed less than three action trials and proportions were computed out of completed trials rather than three

individual feels sick or hurt they may feel sad or think about events that could lead to that state (i.e., cartwheeling could cause you to get hurt). See main manuscript for further discussion.

Overall, as our validation sample is largely 4-year olds and they are performing 1) well above chance and 2) do not differ in performance between image types, it is likely that children understood the meaning of the stimuli well enough to ensure that the findings of Studies 1-2 are valid.

Study 1

Filler trial analyses

Study 1 contained three Filler trials. These Filler trials were included to provide a break from the experimental trials, as per prior research (Gerken & Shady, 1996), and we did not intend to analyze these trials. For these Filler trials, children viewed a video that contained the verb Is and a novel word ending in *–ing* (*piffing, tayving, serding*) to indicate a verb. Thus, if a child chose an action image (i.e., *running, walking, cartwheeling*), they answered correctly. Insofar as Filler trials include "-ing" verb forms, and these verb forms should be easier for children of this age to complete than adjective forms, it could be argued that these Filler trials serve as a type of control trial. However, as we discuss in the main text, the Filler trials might have been more difficult for children than we expected since they could infer a verb across multiple image types (i.e., an alien who is *sad* could also be *crying* or *hurting*). To ensure that children's performance on the Filler trials did not impact their performance on the experimental trials of interest, we include the overall performance on these trials as a covariate in our most conservative ANCOVA reported in the main manuscript. As reported in the main text, children's performance on filler trials did not have an impact on their performance on the experimental trials of interest.

Nonetheless it is interesting to examine children's performance on the fillers in their own right. Here, we examine children's performance on the Filler trials by considering the choice of the action image an accurate response. To do so, we conducted an initial 3 (Image choice: emotion, action, physical state) x 3 (Age: 3, 4, 5) mixed model ANOVA with mean proportions of Filler Image choice as a within subject factor and Age as a between subject factor (see Supplemental Table 2). There was a significant main effect of Image choice, $F(2, 264) = 15.47$, $p < .0001$, $\eta^2 = .11$ (see Supplemental Figure 2), replicating our overall findings that children were the most likely to choose physical state images in Study 1. Pairwise comparisons reveal that children chose physical state images (*M* = .47, *SE* = .03, 95% CI [.41, .52]) significantly more than emotion images (*M* = .25, *SE* = .02, 95% CI [.20, .29]; $p < .0001$) and action images (*M* = .27, *SE* = .03, 95% CI [.22, .32]; $p < .0001$) in the Filler trials. This main effect was qualified by a marginal interaction between Image choice and Age, $F(4, 264) = 2.06$, $p = .09$ $\eta^2 = .03$. Simple effects analysis reveal a marginal influence of Age on the proportion of Action images chosen, $F(2, 132) = 2.90$, $p = .06$, $\eta^2 = .04$. Pairwise comparisons show that 4-year-olds (*M* = .29, *SE* = .04, 95% CI [.26, .42]; $p = .018$) chose action images significantly more than 5-year-olds (*M* = .19, *SE* = .05, 95% CI [.10, .28]). The difference between 3-year-olds (*M* = .19, *SE* = .04, 95% CI [.20, .38]) and 4-year-olds was not significant ($p = .39$) nor was the difference between 3-year-olds and 5-year-olds ($p = .14$). We also conducted follow up analyses and controlled for Gender and Screening trial performance (see Supplemental Table 3) and also examined Filler performance across Sentence Frame condition (see Supplemental Table 4) in the event that Sentence Frame interacted with Filler choice performance. The overall preference for physical state images remained across all analyses and our marginal interaction between Image choice and Age became non-significant in our most conservative analysis controlling for Gender and Screening trial performance.

Thus, Filler trials were not easy for children to perform across the age-span included in our study. As we note in the main text, there are multiple explanations for this finding. Without context, it might have been difficult for young children to home in on the action image, as some of the other images could also have been interpreted as engaging in actions. Another possibility is that although *-ing* in the English language typically denotes an action, some physical states are also described by progressive verbs (*is hurting*, *is starving*) or verbal adjectives ending in *-ing* (*is amazing*). Nonetheless, as we note in the main text, Filler performance did not empirically alter children's behavior on the experimental trials, meaning that children's failures or successes on Filler trials were independent of the pattern of their responses on experimental trials.

Supplemental Analyses of Study 1 Data

There were multiple ways of analyzing the experimental trial data from Studies 1 and 2, each of which have advantages and disadvantages. In the main text, we chose to report an ANOVA/ANCOVA framework, as we believe it provides the greatest ease of interpretation given our within subjects design. Nonetheless, the ANOVA/ANCOVA framework requires that we treat Age as a categorical, rather than continuous variable, and dichotomizing continuous variables is not desirable in statistical analyses.

Regardless, to overcome this limitation, we also conducted multilevel multinomial logistic regressions that allow us to treat Age as a continuous, rather than a categorical, variable. Given the design of our studies, a multilevel multinomial analysis allows us to examine both within and between subject effects. It is also advantageous because its multilevel approach nested within individuals makes it less susceptible to differences in cell size across age. There are other drawbacks associated with these types of analyses, however. First of all, they require treating the dependent variable as a categorical outcome, which requires a logistic (in this case, multinomial) approach that models the likelihood of choosing one of multiple categorical options. Logistic models are interpreted in terms of odd ratios, log-odds, or percentages of a predictor predicting one outcome over others, which imposes an additional burden for interpretation. Odds ratios are the easiest metric to interpret, but many software programs do not calculate the odds ratios for interactions (Chen, 2003; SAS Institute Inc., 2013). Given that our hypotheses predict cross-level interactions (e.g., between age and sentence frame), this fact adds an added barrier to testing our hypotheses using age as a continuous variable in a multilevel multinomial approach. One option in this scenario is to use Maximum Likelihood estimates to evaluate the overall Wald's chi-square test for predictors and to interpret parameter estimates in terms of log-odds. Log-odds provide a relative measure of the magnitude of a predictor predicting one outcome over others but are difficult to interpret in absolute terms. With these caveats in mind, we discuss findings in terms of log-odds throughout this supplement. Finally, an additional drawback is that the multilevel multinomial logistic regression approach required that we drop some participants from analysis, preventing us from analyzing our complete sample.

Together, these models generally replicate our ANOVAs/ANCOVAs reported in the main text, documenting that children perform better on the task, as denoted by increased likelihood to choose a "mental state" image over an action, as Age increases and with increasingly restrictive Sentence frames. For ease of comparison between all models, see Supplemental Table 5 for the pattern of findings for the predictors from all tested models (for both Age as a continuous variable in our multilevel multinomial logistic regressions and Age as a categorical variable in our ANOVAs/ANCOVAs).

Multilevel multinomial logistic regression with Age as a continuous variable. For our main hypotheses we examine the effects of Sentence frame and Age on Image choice. We treat Age as a continuous predictor in a series of multilevel multinomial logistic regressions.

We used SAS (SAS Institute Inc., 2013) and the PROC SURVEYLOGISTIC procedure for all analyses. We treated Image choice as a categorical outcome variable for each trial independently, within each child, rather than treating Image choice as a continuous variable computed as a proportion across all trials (as in the ANOVA and ANCOVAs reported in the main text and below). We conducted a total of three models using a model building approach. Model 1 tested the effects of Sentence frame as a categorical predictor and Age as a continuous predictor on Image choice per trial. In Model 2, we added Gender and Screening trial performance as predictors to control for these factors. In Model 3, we added the interaction between Sentence frame and Age to examine whether Image choice differed within each Sentence frame by Age.

Since we treat each trial as a separate instance within each child, it is important to account for those trials in which children chose not to answer or had technical difficulties. To be conservative, we opted to exclude fourteen children who missed one or more of the four experimental trials, resulting in 484 observations from 121 children (M_{age} = 3.97, SD_{age} = .80; 66 female): 40 3-year olds (23 female), 45 4-year olds (26 female) and 36 5-year olds (17 female). Across Sentence frame, 42 participants were in the Is sentence frame, 41 in the Feels sentence frame and 38 in the Feels about sentence frame. See Supplemental Table 6 for the frequency in which each Image choice type was chosen by Age within each Sentence frame.

Across all Models we used action images as the reference category for Image choice, Is as the reference category for Sentence frame, male as the reference category for Gender and the experimental trial with the alien Xylobean as the reference category for each experimental trial. Age is in years, therefore each unit increase in Age refers to a one year increase in age.[13] See Supplemental Tables 7 and 8 for findings from all models. Significant and predicted findings from Models 1-3 are discussed next, followed by a brief discussion of Models 4 and 5.

In Model 1, we found that Sentence frame was a significant predictor of Image choice, Wald's $\chi^2 = 9.78$, $p = .04$. There was no significant impact of Sentence frame on the log-odds of choosing an emotion image over an action image. Reflecting the bias towards choosing physical state images in our main analyses, children in the Feels sentence frame compared to Is sentence frame had a .58 increase in log-odds of choosing a physical state image over an action image ($\beta = .58$, $SE = .27$; Wald's $\chi^2 = 4.48$; $p = .03$). Children in the Feels About sentence frame compared to the Is sentence frame, had a .57 increase in log-odds of choosing a physical state image over an action image ($\beta = .57$, $SE = .29$; Wald's $\chi^2 = 3.74$; $p = .05$).

Age was also a significant predictor of Image choice, Wald's $\chi^2 = 10.32$, $p = .006$. There was a significant impact of Age on both the log-odds of choosing emotion images and physical state images compared to action images. Specifically, with each year increase in Age, there was a .34 increase in log-odds of choosing emotion images over action images ($\beta = .34$, $SE = .17$;

[13] As not all caregivers reported exact birthdates, we were not able to examine Age effects at a more fine-grained level (i.e., month increments).

Wald's χ^2 = 4.12; p = .04) and a .50 increase in log-odds of choosing physical state images over action images (β = .50, *SE* = .16; Wald's χ^2 = 10.32; p = .001).

When we added both Gender and Screening trial performance to the Model, Sentence frame (Wald's χ^2 = 10.79, p = .03) and Age (Wald's χ^2 = 9.24, p = .01) remained significant predictors of Image choice. Patterns remained the same as those seen in Model 1, with no significant impact of Sentence frame on the log-odds of choosing an emotion image choice over an action image and a significant impact of Sentence frame on the log-odds of choosing a physical state image over an action image. Those in the Feels sentence frame had a .54 increase in log-odds of choosing a physical state image over an action image (β = .54, *SE* = .27; Wald's χ^2 = 3.98; p = .05) compared to those in the Is sentence frame. Those in the Feels About sentence frame had a .62 increase in log-odds of choosing a physical state image over action image than those in the Is sentence frame (β = .62, *SE* = .29; Wald's χ^2 = 4.57; p = .03). With each year increase in Age, there was a marginal .32 increase in log-odds of an emotion image being chosen over an action image (β = .32, *SE* = .17; Wald's χ^2 = 3.35; p = .07) and a .47 increase in log-odds of a physical state image being chosen over an action image (β = .47, *SE* = .15; Wald's χ^2 = 9.24; p = .002). Gender was not a significant predictor of Image choice (Wald's χ^2 =2.74, p = .25) and Screening trial performance was a significant predictor (Wald's χ^2 = 13.04, p = .002). With each unit increase in screening trial performance, there was a 1.18 increase in log-odds of choosing an emotion image (β = 1.18, *SE* = .40; Wald's χ^2 = 8.89; p = .003) over an action image and a .61 increase in log-odds of choosing a physical state image (β = .61, *SE* = .25; Wald's χ^2 = 6.20; p = .01) over an action image.

Critically, in Model 3, we added the interaction between Sentence frame and Age, which tests whether there are differential impacts of Age on performance within the Sentence frames

Feels and Feels About compared to the impacts of Age in the Is sentence frame. Sentence frame became a marginal predictor of Image choice, Wald's χ^2 = 8.96, p = .06. There were no differences in the log-odds of choosing emotion compared to action in either Feels or Feels About sentence frames compared to Is and no difference in log-odds of choosing a physical state image over an action in the Feels sentence frame (which was previously seen). Children in the Feels About sentence frame had a 3.66 increase in log-odds of choosing physical state images over action images compared to those in the Is sentence frame (β = 3.66, *SE* = 1.44; Wald's χ^2 = 6.46, p = .01). Age remained a significant predictor of Image choice (Wald's χ^2 = 9.25, p = .01). With each year increase in Age, there was a .58 increase in log-odds of choosing an emotion image over an action image, although this effect was marginal (β = .58, *SE* = .32; Wald's χ^2 = 3.37, p = .07). With each year increase in Age, there was a .79 increase in log-odds of choosing a physical state image over an action image (β = .79, *SE* = .27; Wald's χ^2 = 8.93, p = .003).

The interaction between Sentence frame and Age was marginally significant, Wald's χ^2 = 8.20, p = .08. There was a .78 decrease between the log odds of choosing a physical state image in the Feels About sentence frame compared to those in the Is sentence frame with each increase in Age (β = -.78, *SE* = .37; Wald's χ^2 = 4.51, p = .03), indicating that children understood the more restrictive nature of this sentence frame as they increased in age. There was no significant impact on the log-odds of choosing physical state images in the Feels sentence frame, nor the log-odds of choosing an emotion image in the Feels or Feels About sentence frame compared to Is sentence frame.

ANOVA with Age as a categorical variable. In our main text, we report a repeated measures ANCOVA with Age as a categorical variable. Here we report more lenient versions of the ANOVA (not controlling for demographic or performance variables) and the ANCOVA (controlling for different combinations of demographic and performance variables).

We first conducted an initial 3 (Image choice: emotion, action, physical state) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANOVA with mean proportions of Image choice as a within-subjects factor and Sentence frame and Age as between-subject factors (see Supplemental Table 8 for all effects). We found a main effect of Image choice, $F(2, 252) = 22.75$, $p < .001$, $\eta^2 = .15$, and this effect remained throughout all analyses. Children chose physical state ($M = .45$, $SE = .02$, 95% CI [.41, .49]; $p < .001$) and emotion ($M = .33$, $SE = .02$, 95% CI [.29, .37]; $p = .001$) images significantly more than action images ($M = .21$, $SE = .02$, 95% CI [.18, .25]). Children chose physical state images significantly more than emotion images ($p = .003$).

The main effect of Image choice was qualified by a significant interaction between Image choice and Age, $F(4, 252) = 2.95$, $p = .02$, $\eta^2 = .05$, which also remained across all analyses. Simple effects showed a significant effect of Age on action images, $F(2, 126) = 5.87$, $p = .004$, $\eta^2 = .09$. Pairwise comparisons revealed that 3-year-olds ($M = .28$, $SE = .03$, 95% CI [.22, .35]) chose action images significantly more than 5-year olds ($M = .13$, $SE = .03$, 95% CI [.06, .19]; $p = .001$) did. Four-year olds ($M = .22$, $SE = .03$, 95% CI [.17, .28]) chose action images significantly more than 5-year-olds did ($p = .03$), but not more than 3-year-olds ($p = .16$). Simple effects also revealed a significant effect of Age on physical state images, $F(2, 126) = 3.90$, $p = .02$, $\eta^2 = .06$. Five-year olds ($M = .52$, $SE = .04$, 95% CI [.45, .60]) chose physical state images significantly more than 3-year olds ($M = .38$, $SE = .04$, 95% CI [.31, .45]; $p = .006$). Four-year-olds ($M = .45$, $SE = .03$, 95% CI [.38, .51]) did not differ from 3-year-olds ($p = .17$) or 5-year-olds (p = .13), however. This suggests that as Age increases and given a novel adjective, children are less likely to choose an action image and more likely to choose an image that reflects a physical state, although not necessarily an emotional one. When we did not control for demographic or performance factors, we did not find our predicted 2-way interaction between Image choice and Sentence frame nor our predicted 3-way interaction between Image choice, Age, and Sentence frame.

ANCOVA with Age as a categorical variable and Gender and Screening trial performance as covariates. A more conservative, 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with Gender and Screening trial performance as covariates was conducted to control for the effects of participant Gender and Screening trial performance (see Supplemental Table 9 for all effects). We again found a main effect of Image choice, $F(2, 248) = 3.58$, $p = .03$, $\eta^2 = .03$. The main effect of Image choice was qualified by a significant interaction between Image choice and Age, $F(4, 248) = 2.75$, $p = .03$, $\eta^2 = .04$, see Supplemental Figure 3. Simple effects showed a significant effect of Age on action images, $F(2, 124) = 5.60$, $p = .005$, $\eta^2 = .08$, and a significant effect of Age on physical state images, $F(2, 124) = 3.56$, $p = .03$, $\eta^2 = .05$. Pairwise comparisons revealed that 3-year-olds chose action images (M = .28, SE = .03, 95% CI [.22, .34], p = .001) significantly more than 5-year olds (M = .13, SE = .03, 95% CI [.07, .20]) did. Four-year olds chose action images (M = .22, SE = .03, 95% CI [.16, .28]; p = .04) significantly more than 5-year-olds did. Three and 4-year-olds did not significantly differ (p = .15) in action image choice. Five-year-olds chose physical state images (M = .52, SE = .04, 95% CI [.44, .60]; p = .009) significantly more than 3-year olds (M = .38, SE = .04, 95% CI [.31, .45]) but not significantly more than 4-year-olds (*M* = .45, *SE* = .03, 95% CI [.38, 51]). Three- and 4 year-olds did not differ in physical state image choices (p = .18).

We found a marginal interaction between Image choice and Sentence frame, $F(4, 248) = 2.25$, $p = .06$, $\eta^2 = .04$. Simple effects revealed a marginally significant effect of Sentence frame on emotion images, $F(2, 124) = 2.71$, $p = .07$, $\eta^2 = .04$ (see Supplemental Figure 4). Pairwise comparisons showed that emotion images were chosen significantly more in the Feels About (*M* = .39, *SE* = .04, 95% CI [.31, .46]; *p* = .02) than the Feels sentence frame (*M* = .27, *SE*= .04, 95% CI [.20, .34]). There was no significant difference between Feels About and Is (*M* = .35, *SE* = .04, 95% CI [.27, .42]; *p* = .45) nor between Feels and Is sentence frame (*p* = .14) on emotion image choice.

As in the main analyses (see manuscript), we did not find our predicted 3-way interaction between Image choice, Age, and Sentence frame, suggesting that Age and Sentence frame independently impact Image choice.

Study 2

Supplemental Analyses of Study 2 Data

For our main hypotheses we are again interested in the effects of Sentence frame and Age on Image choice to assess what information is being used by children when acquiring novel concepts, and at what Age this is occurs. Just as in Study 1, to assess general Age-related effects, we treat Age as a continuous predictor in multilevel multinomial logistic regressions as well as a covariate in our repeated measures mixed model ANOVA/ANCOVAs (see Supplemental Table 10 for comparison of all models). Findings from each are discussed below.

Multilevel multinomial logistic regression with age as a continuous variable. Utilizing SAS (SAS Institute, 2013) and the PROC SURVEYLOGISTIC procedure, we treated Image choice as a categorical outcome variable for each trial independently, within each child. We conducted three main models, the first examining the effects of Sentence frame and experimental trial as separate categorical predictors and Age as a continuous predictor on Image choice. In Model 2, we add Gender and Screening trial performance as predictors and in Model 3, we add the interaction between Sentence frame and Age. Across all models we choose action images to be the reference category for Image choice, Is as the reference category for Sentence frame and male as the reference category for gender. Each unit increase in Age refers to a year increase. As such, all log-odds are interpreted in comparison to Action images, the Is sentence frame and males. As the first story was always a "happy" story, we chose the first story as the reference category. See Supplemental Table 11 for findings from all three Models. Here we discuss significant and predicted findings, highlighting any differences between the three.

As we treated each trial as a separate instance within each child it is important to account for those trials in which children chose not to answer or had technical difficulties. We excluded two children who did not complete all seven experimental trials in this analysis resulting in 777 observations from 111 children (M_{age} = 4.05, SD_{age} = .77; 43 female): 30 3-year olds (7 female), 45 4-year olds (25 female) and 36 5-year olds (11 female). Across Sentence frame, there were 34 participants in the Is sentence frame, 39 in the Feels sentence frame and 38 in the Feels About sentence frame. For the frequency in which each Image choice type was chosen by Age within each Sentence frame, see Supplemental Table 6.

In Model 1, we found our predicted effect of Age on Image choice type, Wald's χ^2 = 26.41, $p < .0001$. With each year increase in Age, there was a .70 increase in log-odds of emotion images being chosen over action images (β = .70, *SE* = .15; Wald's χ^2 = 21.84, *p* < .0001). There was no change in the log-odds of physical state images being chosen over action with each year increase in Age. We did not find our predicted effect of Sentence frame on Image choice.

When Gender and Screening trial performance were added in Model 2, patterns remained the same. Age was a significant predictor of Image choice, Wald's χ^2 = 25.38, *p* < .0001. With each unit increase in Age there was a .67 increase in log-odds of choosing emotion images over action images (β = .69, *SE* = .14; Wald's χ^2 = 21.56, *p* < .0001). There was no significant change in log-odds of children choosing physical state images over action images. Screening trial performance was a significant predictor of Image choice, Wald's χ^2 = 19.19, *p* < .0001. With each unit increase in Screening trial performance there was a 3.87 increase in log-odds of choosing an emotion image over an action image (β = 3.87, *SE* = 1.17; Wald's χ^2 = 11.00, *p* = .0009) and a 4.11 increase in log-odds of choosing a physical state image over an action image (β = 4.11, *SE* = 1.02; Wald's χ^2 = 16.16, *p* < .0001).

When we added our interaction between Sentence frame and Age in Model 3, Age remained a significant predictor of Image choice type, Wald's χ^2 = 28.91, *p* < .0001. With each year increase in Age there was a .78 increase in log-odds of choosing an emotion image over an action image (β = .78, *SE* = .23; Wald's χ^2 = 12.11, *p* = .0005). There was again no change in the log-odds of choosing physical state images over an action with each year increase in Age. Contrary to Models 1 and 2, we found that Sentence frame was a significant predictor of Image choice, Wald's χ^2 = 9.73, *p* = .05. Compared to those in the Is sentence frame, those in the Feels sentence frame had a 2.86 increase in log-odds of choosing an emotion images compared to an action image (β = 2.86, *SE* = 1.51; Wald's χ^2 = 3.59, *p* = .06), although this effect was marginal. There was no significant impact on the log-odds of choosing physical state images in the Feels sentence frame, nor the log-odds of choosing an emotion image in the Feels About sentence frame compared to Is. There is however, a marginal 2.34 decrease in log-odds for physical state images being chosen compared to action for those in the Feels About sentence frame compared to the Is sentence frame ($\beta = -2.34$, $SE = 1.40$; Wald's $\chi^2 = 2.79$, $p = .09$). The interaction between Sentence frame and Age is a significant predictor, Wald's $\chi^2 = 10.77$, $p = .03$. As in Study 1, the interaction term tests whether there are differential impacts of Age within Sentence frames Feels or Feels About compared to the impacts of Age in the Is sentence frame on choosing physical state or emotion images compared to action images. The only significant relationship found was that the log-odds of choosing a physical state image over an action image with each increase in year of Age among children in the Feels About sentence frame is .79 more than each increase in year of Age among children in the Is sentence frame ($\beta = .79$, $SE = .38$; Wald's $\chi^2 = 4.38$, $p = .04$).

ANOVA with Age as a categorical variable. As in Study 1, we ultimately opt to treat Age as a categorical variable. Next we present our less conservative repeated measures mixed models; our most conservative test is reported in the main manuscript.

To examine the impact of Sentence frame and Age on Image choice, we conducted an initial 3 (Image choice: emotion, action, physical state) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANOVA with mean proportions of Image choice as a within-subjects factor and Sentence frame and Age as between-subjects factors. See Supplemental Table 12 for all effects.[14] We found a significant main effect of Image choice,

---

[14] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 17.12, p < .001, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\varepsilon = 0.87$, being greater than .75, $df_{error} = 197.28$.

F(1.90, 197.28) = 48.73, $p < .001$, $\eta^2 = .32$, wherein emotion images (*M* = .50, *SE* = .02, 95% CI [.46, .54]) were chosen significantly more than both physical state (*M* = .33, *SE* = .02, 95% CI [.29, .37]; $p < .001$) and action images (*M* = .17, *SE* = .02, 95% CI [.14, .20]; $p < .001$). Physical state images were chosen significantly more than action images ($p < .001$). This was qualified by a significant interaction between Image choice and Age, *F*(1.95, 197.28) = 9.32, $p < .001$, $\eta^2 = .15$. Simple effects revealed a significant impact of Age on emotion images, *F*(2, 104) = 14.48, $p < .001$, $\eta^2 = .22$, action images, *F*(2, 104) = 7.68, $p = .001$, $\eta^2 = .13$, and physical state images, *F*(2,104) 4.23, $p = .02$, $\eta^2 = .08$. Pairwise comparisons showed that 5-year-olds (*M* = .64, *SE* = .04, 95% CI [.56, .72]) were significantly more likely to choose emotion images than 3-year-olds (*M* = .34, *SE* = .04, 95% CI [.27, .42]; $p < .001$) and 4-year-olds (*M* = .51, *SE* = .03, 95% CI [.45, .58]; $p = .02$). Four-year-olds were significantly more likely to choose emotion images than 3-year-olds ($p = .001$). Five-year-olds were significantly less likely to choose physical state images (*M* = .25, *SE* = .04, 95% CI [.18, .33]; $p = .004$) than 3-year olds (*M* = .41, *SE* = .04, 95% CI [.33, .48]). Five-year-olds were also significantly less likely to choose action images (*M* = .11, *SE* = .03, 95% CI [.05, .16]) than 3-year-olds (*M* = .25, *SE* = .03, 95% CI [.20, .31]; $p < .001$). Similarly, 4-year-olds were significantly less likely to choose action images (*M* = .15, *SE* = .02, 95% CI [.11, .20]; $p = .005$) than 3-year olds.

We did not find a predicted 2-way interaction between Image choice and Sentence frame. However, we found our predicted 3-way interaction between Image choice, Age and Sentence Frame (see Supplemental Figure 6). To probe this interaction, we assessed the 2-way interaction between Sentence frame and Image choice for each age-group. Among 3-year-olds, there was a significant main effect of Image choice, *F*(2, 56) = 3.36, $p = .04$, $\eta^2 = .11$. Pairwise comparisons revealed that 3-year-olds chose physical state images (*M* = .41, *SE* = .04, 95% CI [.33, .48]) significantly more than action images (*M* = .25, *SE* = .03, 95% CI [.19, .32]; p = .02). Emotion images (*M* = .34, *SE* = .03, 95% CI [.27, .41]) were not chosen differently from physical state images (*p* = .31) nor action images (*p* = .12). The interaction between Sentence frame and Age was not significant, $F(2, 56) = .99$, $p = .42$, $\eta^2 = .07$.

For 4-year-olds, we also found a significant main effect of Image choice, $F(1.86, 79.84) = 23.98$, $p < .0001$, $\eta^2 = .36$.[15] Pairwise comparisons revealed that 4-year-olds chose emotion images (*M* = .51, *SE* = .04, 95% CI [.44, .58]) significantly more than action images (*M* = .15, *SE* = .02, 95% CI [.11, .20]; $p < .0001$) and significantly more than physical state images (*M* = .33, *SE* = .03, 95% CI [.27, .39]; $p < .004$). Physical state images were chosen significantly more than action images ($p < .0001$). We found a significant interaction between Sentence frame and Age, $F(3.71, 79.84) = 2.78$, $p = .04$, $\eta^2 = .12$. Simple effects revealed a significant impact of Sentence frame on emotion images, $F(2, 43) = 3.53$, $p = .04$, $\eta^2 = .14$. Pairwise comparisons revealed that 4-year olds-chose emotion images significantly less in the Is Sentence frame (*M* = .39, *SE* = .07, 95% CI [.24, .53]) than in the Feels Sentence frame (*M* = .63, *SE* = .06, 95% CI [.51, .75]; *p* = .01). The frequency of choosing emotion images in the Feels About Sentence frame (*M* = .52, *SE* = .05, 95% CI [.42, 62]) did not significantly differ from the Is Sentence frame (*p* = .13) nor Feels (*p* = .15). Sentence frame significantly impacted action images, $F(2, 43) = 3.92$, $p = .03$, $\eta^2 = .14$). Pairwise comparisons revealed that 4-year-olds chose action images significantly more in the Is Sentence frame (*M* = .26, *SE* = .05, 95% CI [.16, .36]) than in the Feels Sentence frame (*M*

[15] Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 7.80$, $p = .02$, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\varepsilon = 0.86$, being greater than .75, $df_{error} = 79.84$.

= .09, *SE* = .04, 95% CI [.01, .18]; p = .01) and more than the Feels About Sentence frame (*M* = .11, *SE* = .03, 95% CI [.04, .18]; p = .02). Feels and Feels About did not differ (p = .73).

For 5-year-olds, we found a significant main effect of Image choice, $F(1.42, 46.88)$ = 42.57, $p < .0001$, $\eta^2 = .56$.[16] Pairwise comparisons revealed that 5-year-olds chose emotion images (*M* = .64, *SE* = .04, 95% CI [.56, .72]) significantly more than action images (*M* = .11, *SE* = .02, 95% CI [.06, .15]; p = .0001) and physical state images (*M* = .25, *SE* = .04, 95% CI [.18, .33]; $p < .0001$). Action images were chosen significantly less than physical state images (p = .002). We found a marginally significant interaction between Sentence frame and Age, $F(2.84, 46.88)$ = 2.36, p = .09, η^2 = .13. Simple effects revealed that Sentence frame had a marginally significant impact on emotion images, $F(2, 33)$ = 3.00, p = .06, η^2 = .15). Pairwise comparisons revealed that emotion images were chosen marginally more in the Is sentence frame (*M* = .66, *SE* = .06, 95% CI [.54, .78]) than the Feels sentence frame (*M* = .52, *SE* = .05, 95% CI [.41, .62]). Five-year-olds chose emotion images in the Feels About sentence frame (*M* = .74, *SE* = .10, 95% CI [.55, .94]) significantly more than the Feels sentence frame. The Is and Feels About sentence frames did not significantly differ (p = .46). Sentence frame also had a marginally significant impact on action images, $F(2, 33)$ = 2.75, p = .08, η^2 = .14). Pairwise comparisons revealed that 5-year-olds chose action images in the Is sentence frame (*M* = .13, *SE* = .03, 95% CI [.07, .19]) marginally more than in the Feels About sentence frame (*M* = .03, *SE* = .05, 95% CI [-.07, .13]; p = .08). Action images were chosen significantly more in the Feels sentence frame (*M* = .16, *SE*

---

[16] Mauchly's test indicated that the assumption of sphericity was violated, χ^2 (2) = 16.77, $p < .0001$, therefore degrees of freedom were corrected using Greenhouse-Geisser corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\varepsilon = 0.71$, being greater than .75, $df_{error} = 46.88$.

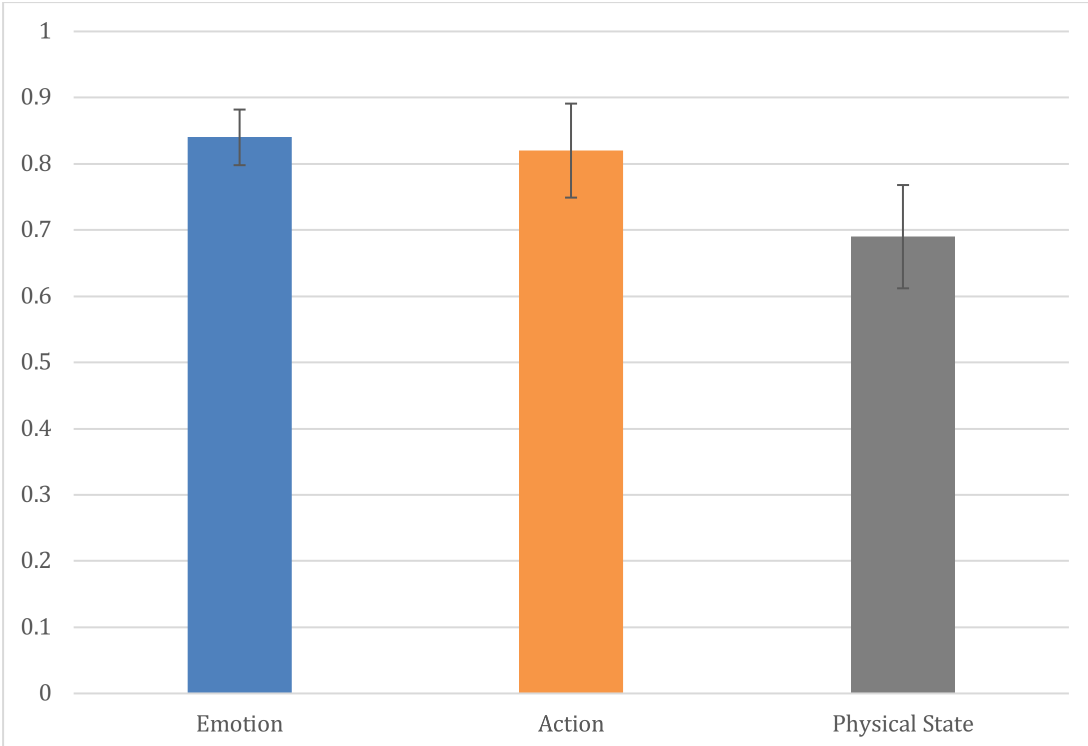
= .03, 95% CI [.11, .21]) than in Feels about (p = .03). The difference between action image choice in the Is and Feels About sentence frames did not differ (p = .51).

Works cited

Chen J. J. (2003). Communicating complex information: the interpretation of statistical interaction in multiple logistic regression analysis. *American Journal of Public Health, 93*(9), 1376-7; author reply 1377.

Gerken, L., & Shady, M. E. (1996). The picture selection task. In *Methods for assessing children's syntax. Language, speech, and communication*.

SAS Institute Inc. 2013. SAS 9.4 Statements: Reference. Cary, NC: SAS Institute Inc.



Supplemental Figure 1. Average correct proportion of each image type in validation study of image stimuli. Bars represent SE.

Supplemental Table 2

t-test for image types and each individual testing whether choice was greater than chance (.33).

Target image type	**M (SD)**	**SE**	***t***	***df***	**95% CI**	***p***
Emotion	***.84 (.17)***	***.04***	***12.34***	***16***	***.43 - .60***	***<.0001***
Mad	.71 (.49)	.18	2.08	6	-.07 - .84	.08
Disgust	.75 (.46)	.16	2.57	7	.03 - .81	.04
Scared	1	-	-	-	-	-
Happy	1	-	-	-	-	-
Excited	1	-	-	-	-	-
Sad	.83 (.41))	.17	3.02	5	.08 - .93	.03
Surprise	.83 (.41)	.17	3.02	5	.08 - .93	.03
Physical state	***.69 (.32)***	***.08***	***4.56***	***16***	***.19 - .52***	***.001***
Hurt	.63 (.52)	.18	1.61	7	-.14 - .73	.15
Hungry	.71 (.49)	.18	2.08	6	-.07 - .84	.08
Burn	.88 (.35)	.13	4.36	7	.25 - .84	.003
Sick	.43 (.53)	.20	.49	6	-.40 - .59	.64
Cold	.75 (.46)	.16	2.57	7	.03 - .81	.04
Hot	.80 (.45)	.20	2.35	4	-.09 - 1.03	.08
Itchy	.71 (.49)	.18	2.08	6	-.07 - .84	.08
Action	***.82 (.29)***	***.07***	***6.98***	***16***	***.34 - .64***	***<.0001***
Sleeping	.86 (.38)	.14	3.69	6	.18 - .88	.01
Jumping	.86 (.38)	.14	3.69	6	.18 - .88	.01
Falling	1	-	-	-	-	-
Cartwheeling	.71 (.49)	.18	2.08	6	-.07 - .84	.08
Walking	.88 (.35)	.13	4.36	7	.25 - .84	.003
Running	1	-	-	-	-	-

Sitting	.71 (.49)	.18	2.08	6	-.07 - .84	.08

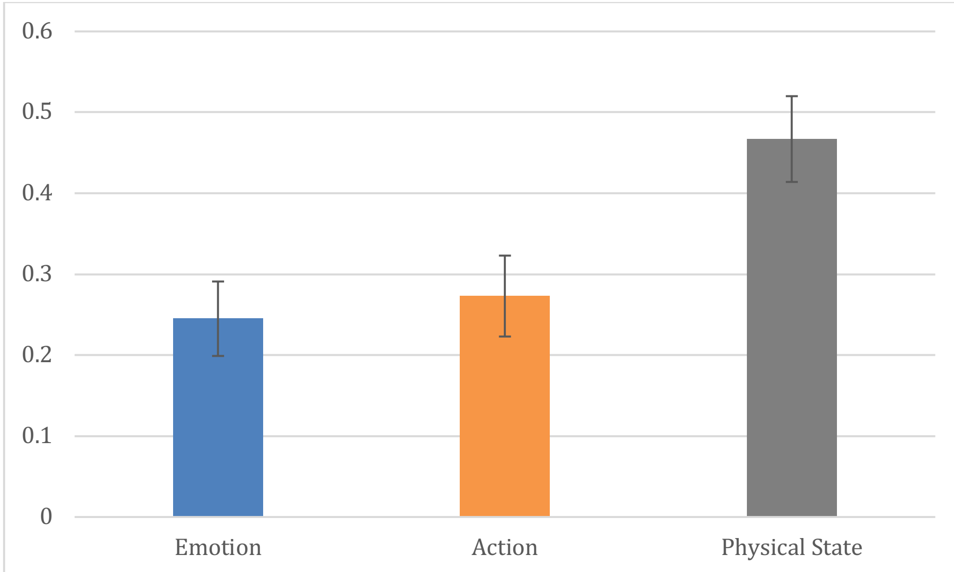
Note: t-test were not conducted for items that were answered correctly across all individuals who received it (scared, happy, excited, falling and running)

Supplemental Table 2

Study 1 within subjects main effects and interactions for 3 (Image choice: emotion, physical state, action) x 3 (Age: 3, 4, 5) mixed model ANOVA on Filler trials.

	df	*F*	η^2	*p*
Image choice	2	15.47	.11	< .0001
Image choice x Age	4	2.06	.03	.09

Note: Mauchly's test indicated that the assumption of sphericity was not violated, $\chi^2(2) = 3.29$, $p = 0.19$, thus sphercity is assumed with $df_{error} = 264$.



Supplemental Figure 2. Study 1 Filler trial estimated marginal means of Image choice from 3 (Image choice: emotion, action, physical state) x 3 (Age: 3, 4, 5) mixed model ANOVA. Bars represent 95% confidence intervals.

Supplemental Table 3

Study 1 within subjects main effects and interactions for 3 (Image choice: emotion, physical state, action) x 3 (Age: 3, 4, 5) mixed model ANCOVA on Filler trials with screening trial performance and gender as covariates.

	df	*F*	η^2	*p*
Image choice	2	.09	.001	.91
Image choice x Screening trial performance	2	.02	<.0001	.98
Image choice x Gender	2	.04	<.0001	.96
Image choice x Age	4	2.02	.03	.09

Note: Mauchly's test indicated that the assumption of sphericity was not violated, $\chi^2(2) = 3.26$, $p = 0.20$, thus sphercity is assumed with $df_{error} = 260$.

Supplemental Table 4

Study 1 within subjects main effects and interactions for 3 (Image choice: emotion, physical state, action) x 3 (Age: 3, 4, 5) x 3 (Sentence frame: is, feels, feels about) mixed model ANCOVA with participant gender and performance on the screening trials as covariates on Filler trials.

	df	*F*	η^2	*p*
Image choice	2	.098	.001	.91
Image choice x Screening trial performance	2	.026	<.001	.98
Image choice x Gender	2	.036	<.001	.97
Image choice x Age	4	1.88	.03	.12
Image choice x Sentence frame	4	.70	.01	.59
Image choice x Age x Sentence frame	8	.15	.005	.97

Note: Mauchly's test indicated that the assumption of sphericity was not violated, $|^2 (2) = 2.86$, $p = 0.24$, thus sphercity is assumed with $df_{error} = 248$.

Supplemental Table 5.

Summarized findings from models tested in Study 1 with Age as a continuous variable and Age as a categorical variable.

	Age continuous			**Age categorical**		
Predictor/covariates	**MMLR Model 1**	**MMLR Model 2**	**MMLR Model 3**	**ANOVA**	**ANCOVA**	**ANCOVA***
IC	-	-	-	Sig	Sig	Sig
Screening trials (x IC)	-	Sig	Sig	-	Sig	Sig
Gender (x IC)	-	Not sig	Not sig	-	Marginal	Marginal
Filler performance (x IC)	-	-	-	-	-	Sig
Experimental trials (x IC)	Marginal	Not sig	Not sig	-	-	-
Age (x IC)	Sig	Sig	Sig	Sig	Sig	Sig
SF (x IC)	Sig	Sig	Marginal	Not sig	Marginal	Sig
Age x SF (x IC)	-	-	Marginal	Not sig	Not sig	Not sig

Note: Sig refers to the corresponding predictor or covariate as a significant predictor or covariate. Marginal indicates that the predictor or covariate was marginally significant with a *p*-value greater than .05 and less than .10. Image choice is included as a predictor for all repeated measures ANOVA/ANCOVAs. Image choice is not a predictor in the MMLRs, as it is the dependent variable for each trial.

IC = Image choice

SF = Sentence frame

MMLR = multilevel multinomial logistic regression

*reported in the main manuscript

Supplemental Table 6.

Frequency of Image choice type for Studies 1 and 2 across all individual experimental trials by Gender, Age and Sentence Frame.

		Study 1				Study 2			
Sentence Frame		**Action**	**Emotion**	**Physical State**	**Total**	**Action**	**Emotion**	**Physical State**	**Total**
Is	Female	26	35	27	88	22	36	26	84
	Male	18	25	37	80	27	75	52	154
	3-yr olds	21	17	18	56	20	26	38	84
	4-yr olds	19	29	24	72	17	25	21	63
	5-yr olds	4	14	22	40	42	60	19	91
	Total	**44**	**60**	**64**	**168**	**49**	**111**	**78**	**238**
Feels	Female	20	23	45	88	11	66	28	105
	Male	13	22	41	76	28	79	61	168
	3-yr olds	12	18	22	52	10	18	21	49
	4-yr olds	16	17	31	64	9	62	27	98
	5-yr olds	5	10	33	48	20	65	41	126
	Total	**33**	**45**	**86**	**164**	**39**	**145**	**89**	**273**
Feels About	Female	14	36	38	88	16	57	39	112
	Male	12	21	31	64	27	76	51	154
	3-yr olds	9	19	24	52	25	27	25	77
	4-yr olds	9	13	22	44	17	80	57	154
	5-yr olds	8	25	23	56	1	26	8	35
	Total	**26**	**57**	**69**	**152**	**43**	**133**	**90**	**266**
Total		**103**	**162**	**219**	**484**	**131**	**389**	**257**	**777**

Supplemental Table 7.

Study 1 estimated regression coefficients and variance components for multilevel multinomial logistic regression with Age as a continuous predictor utilizing SAS SURVEYLOGISTIC.

	Model 1					**Model 2**					**Model 3**				
	$AIC_{intercepts}$ = 1024.71					$AIC_{intercepts}$ = 1024.71					$AIC_{intercepts}$ = 1024.71				
	$AIC_{intercept\ and\ covariates}$ = 1018.22					$AIC_{intercept\ and\ covariates}$ = 1016.21					$AIC_{intercept\ and\ covariates}$ = 1017.68				
Predictor	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**
Type III Analysis															
Sentence frame		9.78	4	.04			10.79	4	.03			8.96	4	.06	
Age		10.32	2	.006			9.24	2	.01			9.25	2	.01	
Separate trials		10.86	6	.09			10.78	6	.10			10.60	6	.10	
Gender							2.74	2	.25			1.90	2	.39	
Screening trials							13.04	2	.002			11.61	2	.003	
Sentence frame x Age												8.20	4	.08	
Emotion															
Intercept	-.98 (.77)	1.63	1	.20		-4.32 (1.36)	10.04	1	.002		-5.20 (1.69)	9.48	1	.002	

Feels	-.01 (.31)	.002	1	.97	.99	-.07 (.31)	.06	1	.81	.93	1.68 (1.59)	1.11	1	.29	
Feels about	.45 (.34)	1.80	1	.18	1.57	.50 (.34)	2.19	1	.14	1.65	2.12 (1.79)	1.40	1	.24	
Age	.34 (.17)	4.12	1	.04	1.41	.32 (.17)	3.35	1	.07	1.37	.58 (.32)	3.37	1	.07	
Gender (female)						.01 (.27)	.001	1	.97	1.01	.03 (.27)	.009	1	.92	1.03
Screening trials						1.18 (.40)	8.89	1	.003	3.27	1.14 (.41)	7.91	1	.005	3.13
Separate trials															
Chromia trial	-.01 (.38)	.001	1	.97	.99	-.01 (.38)	.001	1	.97	.99	-.02 (.39)	.002	1	.97	.99
Palooza trial	-.07 (.34)	.04	1	.84	.93	-.07 (.34)	.04	1	.84	.93	-.07 (.35)	.04	1	.83	.93
Wazu trial	.03 (.40)	.005	1	.94	1.03	.03 (.41)	.005	1	.94	1.03	.03 (.41)	.005	1	.95	1.03
Sentence frame x Age															
Feels											-.46 (.40)	1.36	1	.24	
Feels about											-.43 (.45)	.94	1	.33	

Supplemental Table 7 (continued).

Study 1 estimated regression coefficients and variance components for multilevel multinomial logistic regression with Age as a continuous predictor utilizing SAS SURVEYLOGISTIC.

	Model 1	Model 2	Model 3

	AIC$_{intercepts}$ = 1024.71; AIC$_{intercept\ and\ covariates}$ = 1018.22					AIC$_{intercepts}$ = 1024.71; AIC$_{intercept\ and\ covariates}$ = 1016.21					AIC$_{intercepts}$ = 1024.71; AIC$_{intercept\ and\ covariates}$ = 1017.68				
Predictor	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**	**β (*SE*)**	**Wald's χ^2**	***df***	***p***	**OR**
Physical State															
Intercept	-1.23 (.69)	3.16	1	.08		-2.74 (.98)	7.87	1	.005		-4.04 (1.16)	12.24	1	.0005	
Feels	.58 (.27)	4.48	1	.03	1.78	.54 (.27)	3.98	1	.05	1.72	1.21 (1.51)	.64	1	.42	
Feels about	.57 (.29)	3.74	1	.05	1.76	.62 (.29)	4.57	1	.03	1.86	3.66 (1.44)	6.46	1	.01	
Age	.50 (.16)	10.32	1	.001	1.65	.47 (.15)	9.24	1	.002	1.60	.79 (.27)	8.93	1	.003	
Gender (female)						-.29 (.23)	1.57	1	.21	.75	-.23 (.23)	.96	1	.32	.79
Screening trials						.61 (.25)	6.20	1	.01	1.85	.63 (.25)	6.27	1	.01	1.88
Separate trials															
Chromia trial	-.63 (.36)	3.01	1	.08	.53	-.63 (.37)	2.96	1	.09	.53	-.64 (.37)	2.94	1	.09	.53
Palooza trial	-.66 (.33)	4.07	1	.04	.52	-.67 (.33)	4.02	1	.05	.51	-.68 (.34)	3.96	1	.05	.51
Wazu trial	-.02 (.38)	.002	1	.97	.98	-.02 (.39)	.002	1	.97	.98	-.02 (.39)	.002	1	.97	.98
Sentence frame x Age															
Feels											-.19 (.38)	.24	1	.62	
Feels about											-.78 (.37)	4.51	1	.03	
Overall model evaluation		**χ^2**	***df***	***p***			**χ^2**	***df***	***p***			**χ^2**	***df***	***p***	

Likelihood ratio test	30.49	12	.002	40.50	16	.0007	47.03	20	.0006
Score test	29.79	12	.003	40.15	16	.0007	47.50	20	.0005
Wald test	29.89	12	.003	48.19	16	<.0001	54.41	20	<.0001

Note: OR = odds ratio

Supplemental Table 8.

Study 1 main effects and interactions for 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) ANOVA.

	df	*F*	η^2	*p*
Image choice	2	22.75	.15	< .001
Image choice x Age	4	2.95	.05	.02
Image choice x Sentence frame	4	1.89	.03	.11
Image choice x Age x Sentence frame	8	1.63	.05	.12

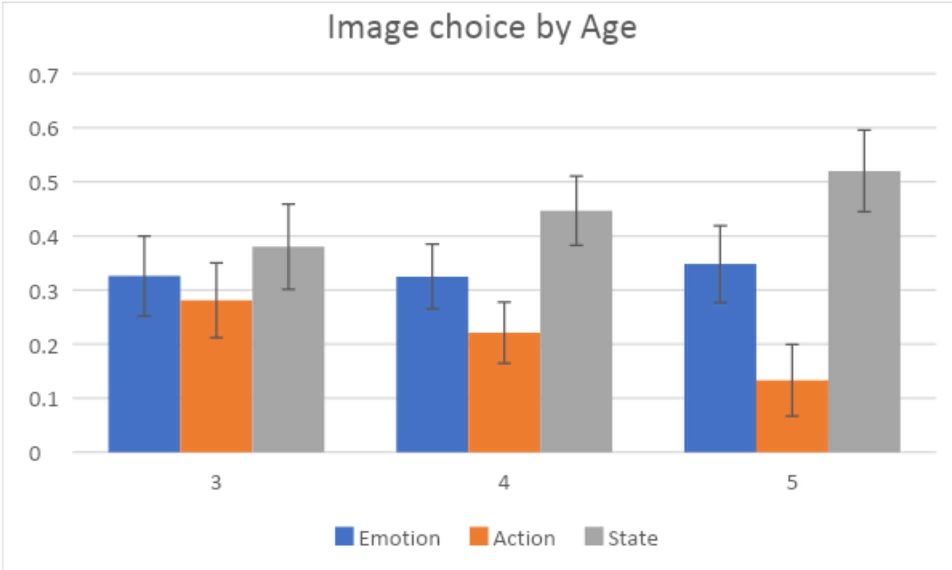
Note: Mauchly's test indicated that the assumption of sphericity was not violated, χ^2 (2) = 5.52, p = 0.06, thus sphercity is assumed with df_{error} = 252.

Supplemental Table 9.

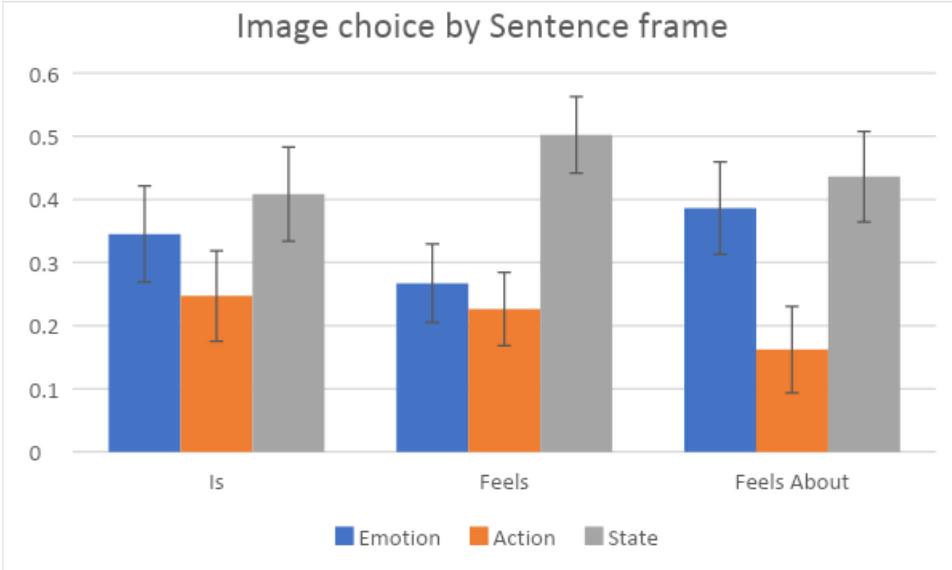
Study 1 within subjects main effects and interactions for 3 (Image choice: emotion, physical *state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with participant gender and performance on the screening trials as covariates.*

	df	*F*	η^2	*p*
Image choice	2	3.58	.03	.03
Image choice x Screening trial performance	2	4.02	.03	.02
Image choice x Gender	2	2.67	.02	.07
Image choice x Age	4	2.75	.04	.03
Image choice x Sentence frame	4	2.25	.04	.06
Image choice x Age x Sentence frame	8	1.32	.04	.24

Note. Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 6.07$, $p = .05$, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, $\Sigma = 0.95$, being greater than .75, $df_{error} = 248$.



Supplemental Figure 3. Study 1 experimental trial estimated marginal means of Image choice by Age in years from our 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with Gender and Screening trial performance as covariates. Bars represent 95% confidence intervals.



Supplemental Figure 4. Study 1 experimental trial estimated marginal means of Image choice by Sentence frame from our 3 (Image choice: emotion, physical state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANCOVA with Gender and Screening trial performance as covariates. Bars represent 95% confidence intervals.

Supplemental Table 10.

Summarized findings from models tested in Study 2 with Age as a continuous variable and Age as a categorical variable.

Test	**Age continuous**			**Age categorical**	
	MMLR Model 1	**MMLR Model 2**	**MMLR Model 3**	**ANOVA**	**ANCOVA***
IC	-	-	-	Sig	Sig
Screening trials (x IC)	-	Sig	Sig	-	Sig
Gender (x IC)	-	Not sig	Not sig	-	Not sig
Experimental trials (x IC)	Sig	Sig	Sig	-	-
Age (x IC)	Sig	Sig	Sig	Sig	Sig
SF (x IC)	Not sig	Not sig	Sig	Not sig	Not sig
Age x SF (x IC)	-	-	Sig	Sig	Sig

Note: Sig refers to the corresponding predictor or covariate as a significant predictor or covariate. Marginal indicates that the predictor or covariate was marginally significant with a *p*-value greater than .05 and less than .10. Image choice is included as a predictor for all repeated measures ANOVA/ANCOVAs. Image choice is not a predictor in the MMLRs, as it is the dependent variable for each trial.

IC = Image choice

SF = Sentence frame

MMLR = multilevel multinomial logistic regression

*reported in main manuscript

Supplemental Table 11.

Study 2 estimated regression coefficients and variance components for multilevel multinomial logistic regression with Age as a continuous predictor.

	Model 1					Model 2					Model 3				
	$AIC_{intercepts}$ = 1577.36					$AIC_{intercepts}$ = 1577.36					$AIC_{intercepts}$ = 1577.36				
	$AIC_{intercept\ and\ covariates}$ = 1511.05					$AIC_{intercept\ and\ covariates}$ = 1496.26					$AIC_{intercept\ and\ covariates}$ = 1492.84				
Predictor	**β (SE)**	**Wald's χ^2**	***df***	***p***	**Odds ratio**	**β (SE)**	**Wald's χ^2**	***df***	***P***	**Odds ratio**	**β (SE)**	**Wald's χ^2**	***df***	***p***	**Odds ratio**
Type III Analysis															
Sentence frame		2.56	4	.63			6.17	4	.19			9.73	4	.05	
Age		26.41	2	<.0001			25.38	2	<.0001			28.91	2	<.0001	
Trials		61.26	12	<.0001			60.22	12	<.0001			59.50	12	<.0001	
Gender							.89	2	.64			.93	2	.63	
Screening trials							19.19	2	<.0001			12.25	2	.002	
Sentence frame x Age												10.77	4	.03	
Emotion															
Intercept	-2.95 (.68)	18.67	1	<.0001		-6.89 (1.32)	27.33	1	<.0001		-6.73(1.60)	17.79	1	<.0001	
Feels	.36 (.32)	1.26	1	.26	1.43	.40 (.31)	1.63	1	.20	1.49	2.86 (1.51)	3.59	1	.058	

Feels about	.46 (.30)	2.26	1	.13	1.58	.71 (.31)	5.14	1	.02	2.04	-1.15(1.53)	.57	1	.45	
Age	.70 (.15)	21.84	1	<.0001	2.01	.67 (.14)	21.56	1	<.0001	1.95	.78 (.23)	12.11	1	.0005	
Gender (female)						.19 (.25)	.59	1	.44	1.21	.19 (.24)	.61	1	.44	1.20
Screening trials						3.87 (1.17)	11.00	1	.0009	48.00	3.19 (1.21)	6.96	1	.008	24.31
Trial															
Afraid	.77 (.35)	4.81	1	.03	2.16	.79 (.36)	4.65	1	.03	2.19	.79 (.37)	4.63	1	.03	2.21
Disgusted	2.19 (.50)	19.51	1	<.0001	8.95	2.25 (.50)	19.91	1	<.0001	9.45	2.27 (.51)	19.76	1	<.0001	9.65
Excited	.34 (.35)	.94	1	.33	1.41	.35 (.37)	.94	1	.33	1.43	.36 (.37)	.94	1	.33	1.43
Mad	1.55 (.44)	12.64	1	.0004	4.71	1.59 (.45)	12.23	1	.0005	4.89	1.60 (.46)	12.30	1	.0005	4.98
Sad	1.44 (.40)	13.03	1	.0003	4.20	1.47 (.41)	12.70	1	.0004	4.33	1.48 (.42)	12.77	1	.0004	4.41
Surprised	.76 (.37)	4.31	1	.04	2.14	.79 (.38)	4.36	1	.04	2.20	.79 (.38)	4.34	1	.04	2.21
Sentence frame x Age															
Feels											-.58 (.35)	2.75	1	.10	
Feels about											.52 (.39)	1.80	1	.18	
physical *State*															
Intercept	-.51 (.67)	.57	1	.45		-4.46 (1.21)	13.67	1	.0002		-3.08(1.44)	4.59	1	.03	
Feels	.32 (.29)	1.20	1	.27	1.38	.37 (.28)	1.76	1	.19	1.45	.13 (1.55)	.007	1	.93	
Feels about	.30 (.28)	1.12	1	.29	1.58	.59 (.27)	4.98	1	.03	1.81	-2.34(1.40)	2.80	1	.09	
Age	.18 (.15)	1.41	1	.24	1.20	.15 (.14)	1.15	1	.28	1.17	-.05 (.20)	.05	1	.82	
Gender (female)						.007 (.23)	.001	1	.97	1.01	.0007(.22)	0	1	.998	1.00

Screening trials						4.11 (1.02)	16.16	1	<.0001	60.82	3.48 (1.05)	10.86	1	.001	32.33
Trial															
Afraid	-.09 (.35)	.07	1	.79	.91	-.08 (.36)	.05	1	.83	.92	-.08 (.36)	.04	1	.83	.93
Disgusted	1.23 (.51)	5.85	1	.02	3.43	1.29 (.51)	6.24	1	.01	3.62	1.30 (.52)	6.22	1	.01	3.66
Excited	.27 (.33)	.67	1	.42	1.31	.28 (.35)	.66	1	.42	1.33	.29 (.35)	.66	1	.42	1.33
Mad	.46 (.42)	1.19	1	.28	1.58	.49 (.43)	1.30	1	.26	1.64	.50 (.44)	1.31	1	.25	1.65
Sad	-.11 (.43)	.06	1	.81	.90	-.08 (.44)	.03	1	.86	.93	-.07 (.45)	.03	1	.87	.93
Surprised	.59 (.36)	2.69	1	.10	1.80	.61 (.37)	2.78	1	.10	1.84	.62 (.37)	2.77	1	.10	1.85
Sentence frame x Age															
Feels											.07 (.35)	.04	1	.84	
Feels about											.79 (.38)	4.38	1	.04	

Overall model evaluation		χ^2	***df***	***p***			χ^2	***df***	***p***			χ^2	***df***	***P***	
Likelihood ratio test		102.31	18	<.0001			125.11	22	<.0001			136.53	26	<.0001	
Score test		96.90	18	<.0001			121.67	22	<.0001			130.41	26	<.0001	
Wald test		90.79	18	<.0001			107.54	22	<.0001			132.68	26	<.0001	

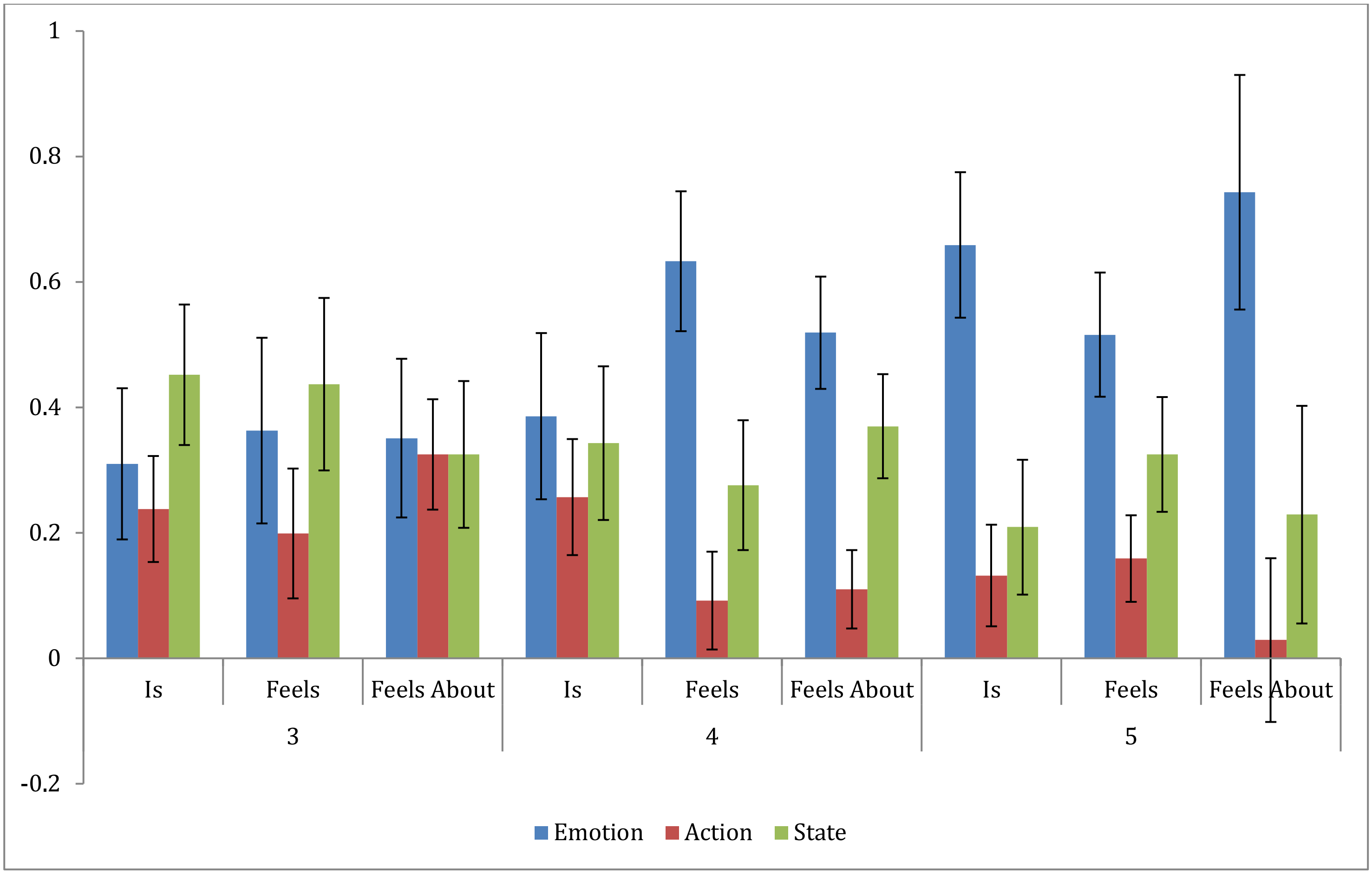
Note: OR = odds ratio

Supplemental Table 12.

Study 2 within subjects main effects and interactions for 3 (Image choice: emotion, physical *state, action) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) ANOVA.*

	df	*F*	η^2	*p*
Image choice	1.90	48.73	.32	< .001
Image choice x Age	3.79	9.32	.15	< .001
Image choice x Sentence frame	3.79	1.02	.02	.40
Image choice x Age x Sentence frame	7.59	2.85	.10	.006

Note: Mauchly’s test indicated that the assumption of sphericity was violated, $|^2$ (2) = 17.12, *p* < .001, therefore degrees of freedom were corrected using Huynh-Feldt corrections follow Girden (1992) suggestion based on Greenhouse-Geisser estimate, ∑ = 0.867, being greater than .75, df_{error} = 197.28.



Supplemental Figure 5. Study 2 estimated marginal means of Image choice, by Sentence frame by Age in years from our 3 (Image choice: emotion, action, physical state) x 3 (Sentence frame: is, feels, feels about) x 3 (Age: 3, 4, 5) mixed model ANOVA. Bars represent 95% confidence intervals.