SymCog: An open-source toolkit for assessing human symbolic cognition

Maurice Flurie[1,2] · Alexandra Kelly[3] · Ingrid R. Olson[4] · Jamie Reilly[1,2,4]

ABSTRACT

Symbol systems have a profound influence on human behavior, spanning countless modalities such as natural language, clothing styles, monetary systems, and gestural conventions (e.g., handshaking). Selective impairments in understanding and manipulating symbols are collectively known as *asymbolia*. Here we address open questions about the nature of asymbolia in the context of both historical and contemporary approaches to human symbolic cognition. We describe a tripartite perspective on symbolic cognition premised upon (1) mental representation of a concept, (2) a stored pool of symbols segregated from their respective referents, and (3) fast and accurate mapping between concepts and symbols. We present an open-source toolkit for assessing symbolic knowledge premised upon matching animated video depictions of abstract concepts to their corresponding verbal and nonverbal symbols. Animations include simple geometric shapes (e.g., filled circles, squares) moving in semantically meaningful ways. For example, a rectangle bending under the implied weight of a large square denotes "heaviness." We report normative data for matching words and images to these target animations. In a second norming study, participants rated target animations across a range of semantic dimensions (e.g., valence, dominance). In a third study, we normed a set of concepts familiar to American English speakers but lacking verbal labels (e.g., the feeling of a Sunday evening). We describe how these tools may be used to assess human symbolic processing and identify asymbolic deficits across the span of human development.

Keywords Symbolic representation · Symbolic cognition · Asymbolia · Social cognition · Abstract concepts · Symbolism

Introduction

Natural language reflects perhaps the most familiar and intuitive example of a symbolic system in our daily lives.[1] Words are symbols whose constituent phonemes together represent objects, emotions, and ideas assigned through cultural convention (Saussure, 1916). Countless other symbol systems govern our behavior, social hierarchies, and communication. For example, styles of dress often signify wealth, in-group status, and social identity. Other nonverbal behaviors (e.g., direct eye contact, giving someone the middle finger) mark social dominance or aggression. Green slips of paper are routinely exchanged for goods and services as a token of value, and we are even sensitive to the presence of a red and white octagon as a marker for when to stop our cars at intersections. Knowledge and competency with the symbol system of a particular culture is often essential for navigating daily life.

One might be tempted to conceptualize symbols in terms of a system that exists exclusively "out there." For example, words are symbols that are imposed on us during early childhood. Symbolic cognition, in contrast, situates symbol processing squarely within the human brain. That is, much of human cognition relies on a calculus of symbols for representing remote ideas. The extent to which symbolic representation (X-symbol represents Y-concept) mediates cognition remains one of the most

---

[1] Pierce described a symbol as something present that stands for something absent (Bobbitt, 2005).

✉ Maurice Flurie
  maurice.flurie@temple.edu

1   Eleanor M. Saffran Center for Cognitive Neuroscience, Temple University, Philadelphia, PA, USA

2   Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, USA

3   Department of Psychological and Brain Sciences, Drexel University, Philadelphia, PA, USA

4   Department of Psychology, Temple University, Philadelphia, PA, USA

A brief history of symbols

Swiss linguist Ferdinand de Saussure offered one of the earliest theories of symbolic representation in his description of bi-directional sign relationships (Saussure, 1916). That is, any symbolic relationship depends on both a *signifier*, such as a word or icon, and something to be *signified* (i.e., some idea or meaning) (Bobbitt, 2005; Parmentier, 2016). Some symbols achieve meaning through physical similarity to their referent (e.g., icons). Others can be linked to an underlying meaning through more arbitrary relationships. Our ability to understand a symbol like a word or "peace symbol" does not rely on physical similarity between the symbol and exemplar. Instead, these relationships are understood through convention. We are taught (either implicitly or explicitly) to link these signs and signifiers (Deacon, 2011). Across all symbolic relationships an individual is required to recognize that signs are representations for some underlying meaning.

In the twentieth century, some scientists adopted this general framework to explore symbolic representation in humans and nonhuman primates. Ethologists studied the communication patterns and emergence of symbol knowledge in primates. An early example comes from authors Hayes & Hayes (1953), who studied picture and object symbol use in Viki the chimpanzee. Other work has evaluated the use of naturalistic symbolic communication, such as using predator calls in vervet monkeys (Seyfarth et al., 1980) and arbitrary symbol use (random icons) representing food and drink rewards for chimpanzees (Savage-Rumbaugh, 1986). In many cases, primates are able to acquire simple paired associations (e.g., "cup" picture = give cup). Unlike human toddlers, however, nonhuman primates tend to show little generalization to other communicative contexts (Deacon, 1998).

Developmental psychologists have investigated the shift from referential to symbolic learning with attention to symbolic play (Geraghty et al., 2014; Namy et al., 2004). A child demonstrates flexible symbolic reasoning when they are able to pretend or use objects as representations for remote items (e.g., holding a wooden block to your ear to emulate a telephone). As children are able to engage in "substitution" behavior, they learn to segregate object and meaning (Vygotsky, 1967). It is thought that as children move from representing objects with other physically similar objects, to pretending without any object present at all, they are able to achieve symbolic thinking (Bruner, 1964; Werner & Kaplan, 1963).

The trajectory of symbolic cognition and its maturity during early adulthood is not well-understood. As such, developmental scientists have elucidated the acquisition of symbolic processing during early childhood, but we must
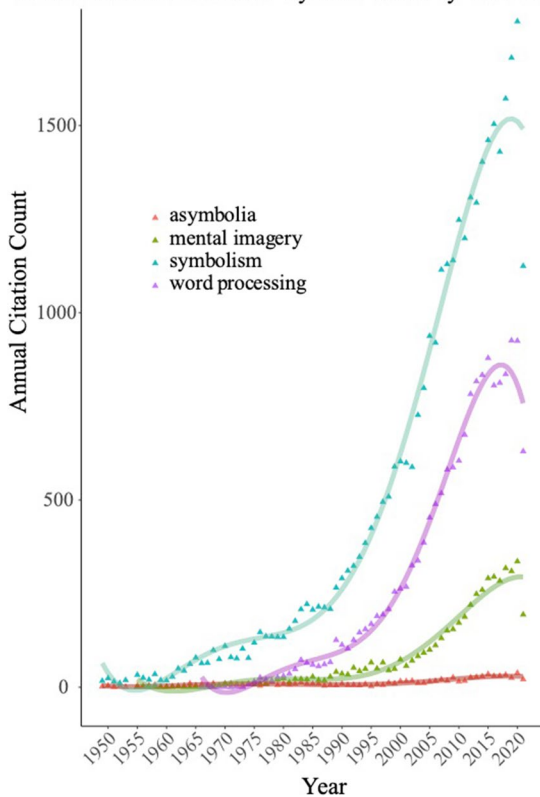

Fig. 1 Shows average word usage by year since 1947. Y-axis shows citation count from PubMed. This figure highlights trends in word usage for symbol-based search terms. While other terms show increased counts over time [likely as a function of exponential growth in annual PubMed publications (PubMed total in 1945: 20,513; PubMed total in 2021: 1,769,461)], 'asymbolia' remains relatively low and unchanged.

controversial topics in neuroscience and the philosophy of mind. Neuropsychological deficits in the capacity to use, represent, or understand symbols are collectively known as "asymbolia" (Duffy & Liles, 1979; Varney, 1978).

Figure 1 illustrates historical trends in empirical research on symbolic cognition beginning with asymbolia in the mid-twentieth century (Duffy & Liles, 1979; Parmentier, 2016). Much of our early understanding of symbolic cognition was derived from related disciplines such as semiotics within the purview of philosophy and linguistics. More recent investigations have integrated findings from artificial intelligence, consciousness, and cognitive neuroscience (Deacon, 2011; Deacon et al., 1996; Garcez et al., 2019; Savchenko et al., 2020). Despite the crucial importance of symbolic processing in language and cognition, much remains to be learned about how symbols are represented and linked with concepts within the human brain.

look to research at the other end of the life span (i.e., older adults with aphasia) for further evidence. The aphasiologist Kurt Goldstein hypothesized that asymbolia might underlie some of the language deficits observed in aphasia. Goldstein explored brain-injured patients' ability engage in abstraction of objects (referred to as the "abstract attitude") (Goldstein, 1948). Goldstein noted that in a "stick task," which required patients to hold specific stick configurations in memory (e.g., arranged in the shape of a house), patients could not produce the configuration if they did not initially recognize the configurations as symbols for objects (Goldstein, 1936).

Asymbolia has also been cited in cross-modality investigations of symbol use (e.g., word-to-picture matching). Using a cross-modality approach has been utilized to circumvent reliance on language tasks. Here, when individuals could not process symbolic relations across modalities (e.g., words and gestures), researchers were able to more confidently diagnose a symbolic impairment. Previous studies have followed this approach (see Cicone et al., 1979; Coelho & Duffy, 1987; Duffy et al., 1975; Duffy & Buck, 1979; Duffy & Duffy, 1981; Pickett, 1974), but the use of gesture performance as a corrolary measure to language performance has its own shortcomings. For example, research has shown some patients with aphasia present with co-occurring ideomotor apraxia (i.e., inability to perform pantomimes/gestures) (Goldenberg & Randerath, 2015). In order to characterize an individual with a symbolic deficit, a task should rely on cross-modality confirmation with relatively limited verbal and motoric demand.

In order to evaluate symbolic cognition and identify disorders such as asymbolia, it is crucial to first operationalize a mechanism that lends itself to empirical testing. We hypothesize that symbolic cognition is composed of three dissociable processes: (1) a stored pool of symbols (e.g., verbal labels, icons); (2) a stored pool of corresponding object concepts; (3) a rapid and efficient mapping process between symbols and concepts.

A paradigm to investigate symbolic representation

A longstanding challenge for cognitive neuropsychology has involved developing a reliable assessment of asymbolia. The empirical study of asymbolia and the development of symbolic knowledge have been informed by many different approaches steeped in the traditions of different disciplines (e.g., developmental psychology, comparative cognition, neuropsychology). For example, the study of symbolic cognition in nonhuman primates has leveraged gesture, learned sign systems, and lexigram keyboards consisting of written visual symbols (Lyn, 2007; Savage-Rumbaugh,

1986). Comparative and developmental psychologists have examined symbolic development and representational deficits by characterizing capacities for pretend play (e.g., use this wooden block like a cell phone) and perspective-taking (e.g., pretend that this doll is a teacher), as well as drawing tasks (DeLoache, 1995, 2004; Rakoczy et al., 2005). Even among tasks developed for adults, task variability differs as a function of how various schools of thought operationalize symbols and symbolic relations. For example, researchers interested in symbolic thought as it relates to creativity have used assessments such as Barron's Symbol Equivalence Test (Barron, 1988) in which participants are prompted with a description of a "symbolic image" such as "tall trees in the middle of a field" and are required to generate verbal descriptions of alternative metaphoric images; while this type of metaphorical thinking may be dependent on cognitive abilities underlying symbolic thought such as abstraction, it is not sensitive enough to capture other relevant components of symbolic cognition. Similarly, tasks using symbols have been widely applied in studies of core geometric knowledge that involve simple maps (Dehaene et al., 2006; Shusterman et al., 2008). In contrast, the toolkit described in this paper offers a means of testing the specific proposed relationships between symbol, concept, and symbol-to-concept mapping.

Characterizing the relationships within this triadic interaction could offer a refined asessment of asymbolia. An impairment for naming an apple, for example, might stem from degradation of the concept of *apples* as is common in the semantic variant of primary progressive aphasia (svPPA) (Bozeat et al., 2003; Humphreys & Riddoch, 2003; Reilly et al., 2011, 2020; Ungrady et al., 2019). In contrast, brain damage may also selectively impact symbols (e.g., word forms) in disorders such as pure anomia (Antonucci et al., 2004; McCarthy & Kartsounis, 2000). A global impairment of symbol integration would result in the inability to map multiple symbol modalities (e.g., verbal, nonverbal) to their corresponding referents. A patient with global asymbolia should present with comparable impairment across all representational modalities.

Here we describe a novel experimental paradigm for evaluating components of symbolic cognition across verbal and nonverbal symbol modalities. Participants match novel concept animations to semantically matched symbols in the form of words and images. Our stimuli may bring to mind Heider & Simmel (1944), animations that involved a large triangle interacting with small geometric shapes. Viewers described the movements of shapes as purposeful and social. Our task offers an investigation of semantic dimensionality (e.g., *valence, sociality, dominance*) to better understand the semantic underpinnings of our animation set. In addition, our task is cross-modal, as participants match words and images to a concept animation.

Study 1: Development and norming of animation, word, and image stimuli

Method:

Overview:

The goal of Study 1 was to create and norm a set of visual stimuli to evaluate symbolic cognition across symbol modalities (i.e., words, images). From crowdsourced survey data, we established normative ratings across multiple visual and semantic dimensions for a series of animations, images, and word stimuli. We first created 35 animations designed to illustrate specific abstract concepts (e.g., *freedom, rising*). We then selected verbal (e.g., words) and nonverbal (e.g., images) symbols as conceptual matches and foils for the animations. Each animation was matched with one target and three foils resulting in a set of 140 words and 140 images. Norms were collected on sensory dimensions (e.g., visual complexity) and name agreement between animations and symbols (e.g., animation-image, animation-word pairs) to facilitate matching between conditions. The resulting stimuli and task are designed to assess individual's ability to identify various symbols for shared concepts (in the form of novel animations).

Stimuli:

Concept animations:

Animations were created that consisted of simple shapes moving about that represented abstract concepts. Each animation was designed to convey a novel illustration of an abstract concept (e.g., *heavy, rise, rejection*) where meaning could be attributed to the shapes and their varying activity. The majority of animations ($N = 30$) were created as "concept pairs" where two animations represented the same concept through different patterns of visual activity. For example, two animations illustrate the concept "multiply" but had unique movements. Paired animations are included to provide multiple contexts to evaluate a singular concept. A smaller number of concepts were represented by only one animation ($N = 5$). This results in 20 different concept exemplars across 35 animations. Each animation was 5 s long.

Symbol images and words:

Symbol stimuli consist of 140 images and 140 words where four images and four words were paired with a single animation. The four images/words were selected such that each animation had (1) a semantic target, (2) a semantically

related (to animation) foil, (3) a visually related (to animation) foil, and (4) an unrelated foil. Word and image stimuli were generated using language databases, author consensus, and survey data with the following procedures:

Images:

All images were selected from the Noun Project (*Noun Project*), a web resource comprised of a collection of more than two million curated icons. Authors chose target images with the goal of selecting images that represent the same concept depicted in animations. Foils were chosen based on shared non-symbolic relationships to animations [excluding unrelated foils, which did not share any relationship (i.e., semantic or visual) with the animation]. Differentiating criteria for each condition were as follows:

- *Targets:* An image was considered a conceptual match of the animation if the image was (1) a sign for the action depicted in the animation (e.g., *division* sign for *Divide* animation), OR (2) depicted the action illustrated in the animation (e.g., *arrow pointing up* image for *Rising* animation, OR (3) depicted an object used to execute the action of the animation (e.g., *vacuum* image for *Consume* animation).
- *Semantic foils*: An image was considered semantically related to the animation if the image did not fit the criteria for a match and was/could be (1) a byproduct or result of the action illustrated in the animation (e.g., *hurt* image for *fall* animation), OR (2) a recipient of the action illustrated in the animation (e.g., *police* image for *violence* animation), OR (3) a descriptive attribute of the animation (*friendly* image for *Competition* animation)
- *Visual foils*: An image was considered visually similar to the animation if the image had at least one physical attribute matching the animation or resembled an object that the animation appeared to portray (e.g., same shape, angle, suggested movement), but did not share any of the semantic relationships of target or semantic foil images.
- *Unrelated foils*: An image was considered unrelated to the animation if the image did not meet the criteria for any of the above.

Words:

Words were selected using a combination of crowdsourced survey and database data. Criteria for each condition are as follows:

- *Targets:* Target words were generated from survey data collected with Amazon Mechanical Turk (MTurk). Participants ($N = 53$) were asked to observe each animation and provide one word representing the activity in

Table 1 Corpus stimuli, dimensions, and prompts

Stimuli type	Dimension	Survey prompt
Image	Visual complexity	How visually complex is this symbol?
	Semantic agreement (with animation)	How well do the animation and symbol represent the same concept?
	Visual similarity (to animation)	How visually similar are the animation and the symbol?
Word	Semantic agreement (with animation)	What word best represents the meaning of this animation?

the animation (see *Data collection procedure* below). The response with the highest frequency across participants was considered a target. To identify additional target options for each animation we used WordNet, an online lexical database used to quantify semantic relationships between words. The online interface of WordNet::Similarity was used to determine relatedness of other words (Pedersen et al., 2004) to the highest frequency participant response. Relatedness here was measured using pairwise path length, a metric reflecting the shortest possible path between two words in a set of synonyms. Words with a path length of 1, the highest possible relatedness value, were added as further targets. If this procedure did not yield three targets (i.e., < 2 words had a path length of 1 to the initial target), the response with the second-highest frequency from the MTurk data was considered and the same similarity procedure in WordNet was implemented. Using WordNet norms provided authors an opportunity to select additional potential targets following a quantitative, data-driven approach. After three potential targets were determined for each concept animation, authors selected one target per animation. For paired animations (e.g., shelter1, shelter2) authors chose the two targets that were most related. For example, shelter1 target options were *rain, shelter*, and *protection* and shelter2 targets were *hide, shoot*, and *cover*. *Shelter* and *cover* were chosen as the respective targets, as they had the highest similarity. For non-paired animations (e.g., *chaos*), the highest frequency word from participants was selected.

- *Semantic foils:* Authors generated semantic foil words using the University of Southern Florida (USF) free association norms (Nelson et al., 2004). Words were chosen that were not a synonym or antonym of targets. If the USF free association norms did not yield an appropriate word for a given concept animation, a word was generated via consensus of the authors following the differentiating criteria described above (*See Images: Semantic foils*)
- *Visual foils*: Visual foils were chosen via author consensus by selecting a word describing a physical feature of the animation unrelated to the semantic context (e.g., triangle, spin) or an object similar to physical shapes

presented in the animation (e.g., a line near a rectangle looked similar to a flag).

- *Unrelated foils:* Unrelated foils were pulled from the SUBTLEX database of word frequency in American English (Brysbaert & New, 2009) by making a pseudorandomized list of words matched in letter length to target words ($M = 5.34$, range $= 3–14$) with a minimum frequency threshold of 100.

Participants:

A total of 294 individuals participated in six surveys [$M = 49$ ($SD = 7.8$) participants per survey] using Amazon's MTurk platform. Each survey targeted different dimensions of animation, image, and word stimuli. All participants were English speaking from birth with a high school degree (or GED equivalent). All participants were at least 18 years of age ($M = 36.8$, range $= 20–69$). A total of 182 men and 112 women participated in the surveys, with a mean male/female ratio of 1.6:1. There were no statistically significant differences between surveys in participant age as determined by one-way analysis of variance (ANOVA) [$F(1,290) = 0.701$, $p = 0.4$]. Participants all had master-level experience designation as provided by Amazon. Participants were compensated $10/hour for participation. Participants spent an average of 34.26 min to complete a survey which included 415.02 ratings on average.

Data collection procedure:

We utilized Qualtrics survey software to evaluate our animation, image, and word stimuli on a series of dimensions across six surveys. After providing informed consent, participants observed an instructions screen that explained they would be asked to answer questions about a set of picture symbols, short animations, or words. Instructions further stated that some questions would ask them to consider "what they mean" while others questions would ask them to consider "how they look." Table 1 provides each prompt used for each dimension assessed per modality (i.e., image, word). The instructions specified that participants should not spend too long answering questions and should rely on their initial instincts. Responses were provided on a five-point Likert scale where 1 indicated the *least* (e.g., not at all, not well)

Table 2 Visual similarity contrasts between image conditions

Contrast	Means	Test statistic (t)	Effect size (Cohen's d)
Target vs. semantic foil***	Target (2.98), semantic foil (2.14)	− 5.93	− 1.42
Unrelated foil vs. semantic foil	Unrelated foil (1.86), semantic foil (2.14)	2.52	0.60
Visual foil vs. semantic foil***	Visual foil (3.13), semantic foil (2.14)	− 8.53	− 2.05
Unrelated foil vs. target***	Unrelated foil (1.86), target (2.98)	7.72	1.84
Visual foil vs. target	Visual foil (3.13), target (2.98)	− 1.00	− 0.24
Visual foil vs. unrelated foil***	Visual foil (3.13), unrelated foil (1.86)	− 10.57	− 2.52

***$p < 0.001$

and 5 indicated the *most* (e.g., very complex, very much, very well). "Semantic agreement" for words was open-ended (one, single word response).

Data norming procedures:

Word and image stimuli were evaluated to ensure conditions (i.e., targets, foils) did not differ on low-level visual features. Ratings of visual complexity were evaluated using a one-way ANOVA, which revealed no significant differences between image conditions [$F_{(3,136)} = 1.66$, $p = 0.18$]. A one-way ANOVA of word length showed no significant differences in number of letters between conditions [$F_{(3,136)} = 1.93$, $p = 0.13$] (Table 2).

Results:

We calculated summary statistics for all conditions (i.e., target, foils) in each symbol modality (i.e., words, images). Tables 3 and 4 present all resulting stimuli (or descriptions of stimuli for images) with relevant statistics. For words, mean name agreement is the highest frequency word count collapsed across all participants and target words. To convert to a percentage, the count was divided by the number of participants ($N = 53$) who took the survey. For images, mean name agreement and mean visual similarity are the average participant rating (1–5 Likert scale) collapsed across images. Visual similarity refers to similarity between animation and image pairs. A one-way ANOVA of visual similarity by condition revealed significant differences between the four conditions [$F_{(3,136)} = 45.18$, $p < 0.001$] where visual foils achieved the highest visual similarity to animations ($M = 3.13$, $SD = 0.53$) followed by targets ($M = 2.98$, $SD = 0.71$), semantic foils ($M = 2.14$, $SD = 0.44$), and unrelated foils ($M = 1.86$, $SD = 0.48$). Table 4 shows post hoc analyses with Bonferroni correction.

Matching accuracy in Symbol Identification Task:

From our normed stimuli we developed a paradigm to evaluate the symbolic representation of concepts: a forced-choice task requiring participants to select a target symbol in the presence of three foils. MTurk participants ($N = 30$) observed a concept animation and a display of four symbols with the prompt, "Which image/word best represents the meaning of the animation?" The task was presented in two blocks, one presenting word symbols and the other image symbols. Figure 2 demonstrates a single trial in the image block for the animation "competition1." Fig. 3a describes matching accuracy per animation in each block, where the mean accuracy was 78.7% ($SD = 20\%$) for image targets and 94.9% ($SD = 6.9\%$) for word targets. We implemented a 70% cutoff score for matching accuracy in both conditions, resulting in the removal of eight images and one word. In a follow-up survey, we replaced either the target or a foil symbol for these trials to facilitate increased accuracy. Figure 3b shows accuracy for our final target list where all image-targets and word-targets maintained accuracy > 65%.

Interim discussion:

In our first study we aimed to generate a series of symbols (i.e., words and images) as semantic matches and foils for our concept animations. For each concept animation we generated four images and four words: a semantic match, and three foils of varying types. From these stimuli we created a novel paradigm (forced-choice matching task) to evaluate symbolic representation across verbal and nonverbal contexts. The objective of our paradigm was to match concept animations and symbols based on matched semantic meaning. A critical component of norming our stimuli involved low-level visual similarity, as we did not want matching animations and concepts to rely on simple visual similarities between pairs. To norm our image stimuli on low low-visual dimensions, we evaluated (1) overall

Table 3 Animation and word pairings

Animation ID	Target 1	Semantic foil	Visual foil	Unrelated foil
acceptance1	*join*	*offer*	*flower*	*cost*
acceptance2	*join*	*gift*	*caterpillar*	*escape*
change1	*morph*	*clothes*	*square*	*keeper*
change2	*transform*	*tire*	*post*	*grind*
chaos1	*chaos*	*fire*	*shape*	*candle*
competition1	*race*	*friendly*	*black*	*halt*
competition2	*race*	*enemy*	*circle*	*retrieve*
confinement1	*trap*	*small*	*walk*	*mustache*
consume1	*consume*	*full*	*little*	*throw*
consume2	*consume*	*clean*	*dots*	*reflect*
division1	*divide*	*candy bar*	*triangle*	*tribune*
division2	*divide*	*turkey*	*candy*	*diplomacy*
equality1	*exchange*	*anger*	*bracelet*	*redemption*
equality2	*copy*	*protest*	*cone*	*depend*
exchange1	*exchange*	*call*	*blocks*	*blade*
excrete1	*spitting*	*cup*	*pebble*	*come*
falling1	*fall*	*hurt*	*bug*	*baloney*
falling2	*fall*	*royalty*	*round*	*insert*
freedom1	*escape*	*fighter*	*stamp*	*deal*
heavy1	*bend*	*metal*	*smile*	*stitch*
heavy2	*squash*	*dark*	*lamp*	*paint*
multiply1	*multiply*	*school*	*six*	*swung*
multiply2	*multiply*	*quick*	*flag*	*worship*
order1	*group*	*judge*	*stars*	*consider*
order2	*organize*	*food*	*sharp*	*freshman*
rejection1	*block*	*throw*	*necklace*	*hesitate*
rejection2	*reject*	*nerd*	*line*	*elk*
rising1	*rise*	*sunny*	*angle*	*geese*
rising2	*rise*	*flag*	*black*	*scientist*
shelter1	*shelter*	*bomb*	*arrow*	*finger*
shelter2	*shelter*	*homeless*	*claw*	*kept*
transportation1	*transport*	*camera*	*grey*	*take*
transportation2	*transport*	*luggage*	*motion*	*businessman*
violence1	*chase*	*pain*	*piece*	*pack*
violence2	*destroy*	*police*	*moon*	*gamma*

"Animation ID" is the name provided by authors. "Targets" are those provided by survey participants.

visual complexity for each image and (2) visual similarity between animations and paired images. Visual complexity did not significantly differ between targets and foils. Visual similarity of images to animations were such that visual foils > targets > semantic foils > unrelated foils. This suggests that participants relying on visual similarity on trials are most likely to select visual foils (i.e., visually similar to animation) as matches. For words, we assessed visual complexity by contrasting the number of letters between the four conditions. No significant differences were found between conditions, suggesting word length should not impact target selection.

Before developing our paradigm, we evaluated name agreement for word and image stimuli to ensure selected targets accurately represented their paired animation. For images, mean name agreement was 3.81 on a five-point Likert scale, which showed that our target images were closer to representing animations "very well" (i.e., 5 on scale) than "not well" (i.e., 1 on scale). For words, we did not assess name agreement after selection. Instead, it was assessed during selection by choosing the word that participants provided with the highest frequency when naming concepts depicted in animations.

After norming our stimuli, we created a novel paradigm to assess symbolic representation in abstract concepts. In two blocks (word block, image block), participants were tasked with selecting a symbol that matched the animation in an array of one target and three foils (visual, semantic, unrelated). The mean accuracy for selecting targets was 78.7% ($SD = 20$%) for image targets and 94.9% ($SD = 6.9$%) for word targets. To improve overall accuracy, we implemented a 70% cutoff for each animation-target pair. This resulted in a second survey where eight images and one word were replaced. Overall accuracy of target selection was improved to 84.70% ($SD = 11.0$%) in images and 96.25% ($SD = 4.4$%) in words. All replaced images and words resulted in improved accuracy for target selection. Three animation-image pairs remained below 70% accuracy [e.g., equality2 (61.2%), heavy1 (54.8%), heavy2 (58.1%)], but are still included in the stimuli set as accuracy remained far above chance (i.e., 25% in four-image array).

The development and norming of this stimulus set highlighted the dynamic nature of abstract concept identification. Researchers should be cautious to use these stimuli in an open-ended context. For word symbols, many animations had multiple words which were identified as potential matches through survey results, suggesting numerous words can be used to represent any single animation in our set. In an effort to recognize the fluid relationship between concept animations and target words, we included synonyms as potential candidates in selecting our word targets. The Symbol Identification Task aims to curtail this dynamic relationship between animations and symbols by leveraging a forced-choice task where only one word or image can reasonably represent the animation. For example, while the animation for *chaos* could be represented by the words "chaos," "movement," or "crazy," a trial in the Symbol Identification Task would only include the target "chaos" with three foils (e.g., shape, fire, candle).

Table 4 Animation and image pairings

Animation ID	Target	Semantic foil	Visual foil	Unrelated foil
acceptance1	*two people shaking hands*	*happy face*	*necklace*	*shoes*
acceptance2	*two people holding hands*	*gift*	*bridge*	*tree*
change1	*insect in cocoon*	*fast moving clock*	*pillow*	*calculator*
change2	*arrow changing shape*	*TV*	*vertical lines*	*desk tray*
chaos1	*confetti flying*	*wild tiger*	*multiple shapes*	*triangles*
competition1	*three individuals on medal podium*	*three coins*	*volume mixer*	*stick*
competition2	*three people crossing finish line*	*two enemies facing each other*	*leaf blowing in wind*	*road*
confinement1	*person in jail cell*	*mouse*	*soccer field*	*hand pressing button*
consume1	*mop*	*weight scale*	*ping pong paddle and ball*	*hair dryer blowing*
consume2	*vacuum*	*mess*	*grid*	*camera*
division1	*pie chart*	*candy bar*	*camera lens*	*horizontal lines*
division2	*rectangle breaking in pieces*	*two on die*	*Window*	*cracked screen*
equality1	*two shapes trading locations*	*mixing bowl*	*two kebabs*	*moon*
equality2	*equal sign*	*Measuring tape*	*fence*	*refrigerator*
exchange1	*individual giving money for product*	*number pad on phone*	*stack of shapes*	*picture frame*
excrete1	*pouring kettle*	*empty box*	*speech bubble*	*boiling water*
falling1	*person dropping phone on ground*	*hurt knee*	*battery*	*drawer*
falling2	*down-pointing arrow*	*royalty figure*	*marbles*	*construction hat*
freedom1	*an unlocked lock*	*people fighting*	*clothing tag*	*bubbling beaker*
heavy1	*anvil*	*person stuck in hole*	*banana*	*knotted rope*
heavy2	*obese figure*	*Band aid*	*unicycle*	*kettle*
multiply1	*A circle pointing to three circles*	*complex maze*	*car pedals*	*cog*
multiply2	*increasing volume sign*	*quick movement*	*square*	*lightning bolt*
order1	*shapes in groups*	*calm meditator*	*Boot spur*	*brick pyramid*
order2	*cubes moving to bins*	*cleaning supplies*	*Clown hat*	*ice cubes*
rejection1	*person with crossed arms*	*upset person*	*tire*	*bracelet*
rejection2	*"no access" sign*	*frowning face*	*earth*	*basketball*
rising1	*escalator*	*sun in sky*	*traffic cone*	*train tracks*
rising2	*thermometer*	*person with medal around neck*	*thunderstorm*	*microwave*
shelter1	*umbrella*	*help*	*shoveling*	*office desk*
shelter2	*tent*	*recover*	*prongs*	*brick wall*
transportation1	*bus*	*small window*	*metal trashcan*	*shooting star*
transportation2	*person carrying box up steps*	*travel mug*	*bowl*	*turtle*
violence1	*person standing on a downed person*	*danger*	*French fries*	*three people jumping rope*
violence2	*Persons standing over cowering person*	*storm*	*satellite*	*crane*
Mean name agreement	3.81 (0.54)	NA	NA	NA
Mean visual similarity	2.98 (0.71)	2.14 (0.44)	3.13 (0.52)	1.86 (0.48)

Targets and foils are descriptions of selected images. For full list of images see all stimuli on Open Science Framework: https://osf.io/9jr6h/. "Mean name agreement" (between animations and target images) and "mean visual similarity" are average participant ratings on a 1–5 Likert scale.
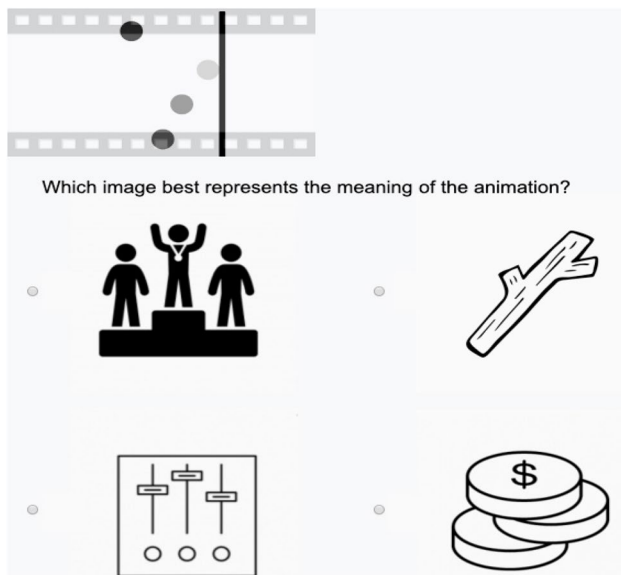
Figure shows "snapshot" of trial for "competition1" animation. In the trial, the animation (top) plays on a loop as participants try to select the correct symbol target.

Fig. 2 Single trial of identification task

Study 2: Animation dimensions and clustering

Overview:

We evaluated our animation set ($N=35$) on a series of semantic dimensions in order to better understand the semantic underpinnings that support creating meaning in each animation. We also sought to understand how animations might cluster into related categories. We selected four dimensions to evaluate the semantic structure of animations: sociality, dominance, theory of mind, and valence. Dimensions were selected based on previous work exploring semantic representations of abstract concepts (Binney, 2016; Crutch, 2013) where concepts tended to cluster. The resulting dimension norms from this study can be utilized to investigate individuals' understanding for specific categories (i.e., clusters) of concepts. For example, an individual could have difficulty processing highly social or highly valenced concept animations, while other categories remain intact.

Method:

Participants Thirty individuals (22 male/8 female) participated in a survey using Amazon's MTurk platform. All participants were at least 18 years of age ($M=37.9$, range $=26–57$).

Data collection We utilized Qualtrics survey software to evaluate animation semantic dimensionality. Participants

first observed an instruction screen explaining the procedure of the survey. Instructions informed participants that they would observe a series of animations and answer questions relating to how they look and what they mean. Participants were informed not to spend too much time on any animation and to rely on initial instinct. Table 5 describes each dimension assessed with the prompt used. All responses were on a five-point Likert scale. For Sociality and Dominance, scale anchors ranged from 1 (*not at all*) to 5 (*very much*). For Theory of Mind and Valence, scales were 1 (*negative*) to 5 (*positive*).

Results

We computed summary statistics for each of the four-dimension ratings (e.g., Sociality, Theory of Mind, Dominance, and Valence) across all animation based on MTurk survey responses ($N=30$ participants). On a five-point Likert scale, animations had an average Sociality rating of 3.36 ($SD=1.41$), Theory of Mind rating of 2.97 ($SD=1.2$), Dominance rating of 2.72 ($SD=1.55$), and Valence rating of 3.05 ($SD=1.22$). Figure 4a displays averages per animation across participants. Figure 4b shows a selection of individual animation ratings. To evaluate similarity between animations based on these semantic dimensions, we performed hierarchical clustering. The resulting dendrogram can be seen in Fig. 5. We calculated gap statistics to determine optimal cluster size (k-means). Animations best aggregated with a three-cluster solution. To determine the dimensions responsible for each cluster we contrasted clusters based on the four dimensions. We coded each animation based on its designated cluster (e.g., cluster 3: violence1&2, rejection1&2, heavy2, shelter2, consume1). A linear mixed-effects model (lme4 R Package) revealed a main effect of cluster [$F(2,32)=21.83$, $p<0.001$], a main effect of dimension ($F(3,96)=12.09$, $p<0.001$), and significant cluster*dimension interaction [$F(6,96)=33.67$, $p<0.001$]. Table 6 shows results of post hoc analysis using paired t-tests with Bonferroni correction to contrast clusters by each dimension.

Interim discussion:

In our second study we sought to better understand the semantic underpinnings of our abstract concept animations. We evaluated four semantic dimensions to characterize animations in a semantic space consisting of *valence, dominance*, *theory of mind*, and *sociality* based on crowdsourced survey data. Animations were then subject to hierarchical clustering analysis to determine if animations tended to group together. Three distinct clusters were defined where
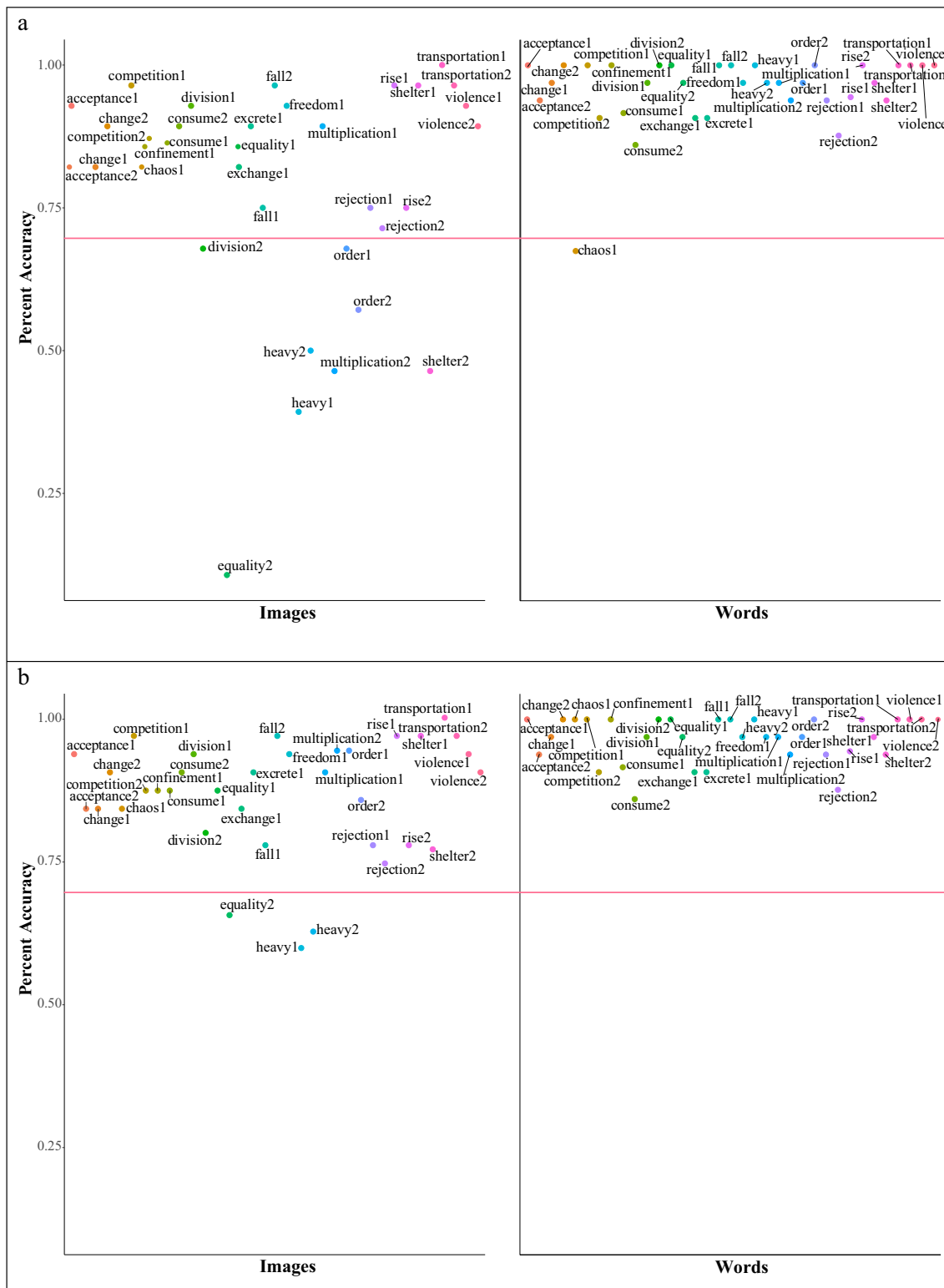
Note: Figure shows differences in matching accuracy of word and images to animations. Cut off score at < 70% used to indicate images/ words needing to be replaced: this includes eight images (left) and one word (right). Matching accuracy improved for all replaced images and words. Three images remained below 70% accuracy (equality1, heavy1, heavy2).

Fig. 3 Animation-image and animation-word matching accuracies

nearly all clusters were significantly different from each other on all four dimensions. Cluster 1 included concept animations that were highly social and positive (i.e., high valence), such as *equality*1, *acceptance*1&2, and *transportation*1&2. Cluster 2 was defined by lower sociality and included concepts *division*1&2, *fall*1&2, and *rise*1&2. Cluster 3 was comparable to Cluster 1 in sociality, but included negative social concept animations like *rejection*1&2 and *violence*1&2. It is possible that some individuals or populations might present with focal difficulty in our matching paradigm based on the semantic nature of animations. For example, populations where individuals often have difficulty with theory of mind and social constructs might have the ability to relate symbols to animations with lower sociality and theory of mind ratings (e.g., *rise, heavy*) but struggle with concepts defined by social dimensions (e.g., *acceptance, rejection*).

Study 3: Developing and norming concepts without words

Overview:

We established familiarity ratings for a series of concept definitions that do not have a representative word in English. Definitions were reduced to only the most familiar/personally experienced based on survey data results. These definitions can be utilized to evaluate semantic knowledge in individuals without relying on specific lexical or visual labels.

Method:

Participants 59 individuals participated in a survey using Amazon's MTurk platform. Six participants were removed due to incomplete survey responses. This resulted in 53 (27 male/26 female) participants who were at least 18 years of age ($M = 36.11$, range $= 20–63$), had at least a high school education, and spoke English from birth.

Stimuli Concepts without words were sourced through members of the Concepts and Cognition Laboratory at Temple University. A list of more than 150 definitions was created for concepts that do not have words in English. The objective was to create a series of definitions for concepts (e.g., ideas, states of thinking, a particular feeling) that might exist in typical or daily life, yet may not have a specific word tied to the definition. These included a range of concepts that related to primary senses (e.g., the voice people use

when talking to pets, the smell of paper money, the sensation of one's stomach rising during freefall), emotions (e.g., the feeling of summer coming to an end, the urge to pinch/squeeze something cute) and ideas (e.g., the realization that two people are thinking or feeling the same thing). Concept definitions were reduced to 95 total after author consensus regarding the existence of the described concept. Definitions were then entered into MTurk surveys for further validation and norming.

Data collection:

To contextualize the survey, we began the instructions screen by stating that "languages have words for culturally shared concepts, yet some concepts we experience may not have names." Participants were then informed that they would be answering questions about concepts. Table 7 illustrates each dimension assessed for each concept with its relative prompt. "Familiarity" (self) responses were dichotomous (yes/no). Familiarity (others) responses were on an eight-point Likert scale, reflecting the likelihood that others experienced the concept. Likert scale responses ranged from 1 (*no one*) to 8 (*everyone in the world*).

Results:

We first generated descriptive statistics for each dimension. For Familiarity (self), the mean frequency of "Yes" ratings (i.e., they have experienced the concept) across concepts was 36.9 ($SD = 7.34$) out of the 53 participants. Figure 6a shows the average percentage of "Yes" responses per concept. For each concept we calculated frequency percentiles by dividing the frequency of "Yes" responses by the number of total participants ($N = 53$). We imposed a 70% frequency cutoff, where concept definitions had to be familiar to at least 70% of participants to remain. This yielded 51 concept definitions (see https://osf.io/9jr6h for full list of concept definitions). An unpaired *t*-test revealed significant differences between high and low Familiarity (self) groups, $t(55) = x$, $p < 0.001$, $d = -3.14$ where mean familiarity in the high group was 80.5% ($SD = 4.2\%$) and 57.1% ($SD = 9.9\%$) in the low group across participants and concept definitions. Figure 6b summarizes "Familiarity (others)" across participants while continuing to differentiate high vs. low "Familiarity (self)" contrasts. For the high Familiarity (self) group, mean familiarity was 5.89 ($SD = 0.52$) and 4.78 ($SD = 0.49$) for the low Familiarity (self) group. An unpaired *t*-test revealed significant differences between groups, $t(91.9) = x$, $p < 0.001$, $d = -2.16$.

Table 5 Animation dimensions and prompts

Dimension	Survey prompt
Sociality	"How much could this represent social interactions between humans?"
Dominance	"How much do you see actions suggesting dominance?"
Theory of Mind	"What are the shapes feeling in this animation?"
Valence	"How does this animation make you feel?"

Interim discussion:

Our third experiment consisted of norming a set of concept definitions in order to generate a series of concepts that exist in society, but presently do not have names. To accomplish this aim, 95 concept definitions were generated and normed on two dimensions: "Familiarity (self)" (i.e., personal experience with the concept) and "Familiarity (others)" (i.e., experienced amongst other people). When including only the most frequently experienced concepts definitions in our final set, concepts had been experienced by 80.7% of participants on average, and had an average rating of 5.89 (on an eight-point Likert scale) for familiarity. Together, these norming procedures provide a set of 51 concepts without words to be used by researchers

interested in tapping concept knowledge unrelated to specific words or other symbols.

General discussion:

We developed an open-source toolkit to evaluate symbolic cognition and reported three normative studies characterizing the stimuli. This suite of experiments was designed to assess symbol-semantic access from abstract concept animations ($N = 35$), words, and static images. In our first study, we acquired matching agreement norms between concept animations and paired symbols (i.e., words and images). In our second study we evaluated clustering properties and semantic dimensionality of the animation stimuli. In the final study, we acquired familiarity norms for a set of concepts without words. This toolkit may prove useful for elucidating three broad cognitive processes that allow us to make use of symbols, including (1) integrity of semantic representations, (2) integrity of corresponding verbal and nonverbal symbols, and (3) a durable and efficient mapping between concepts and symbols.

One major challenge for the study of symbolic cognition involved disentangling co-occurrence statistics and paired associations from symbolic representation. For example,
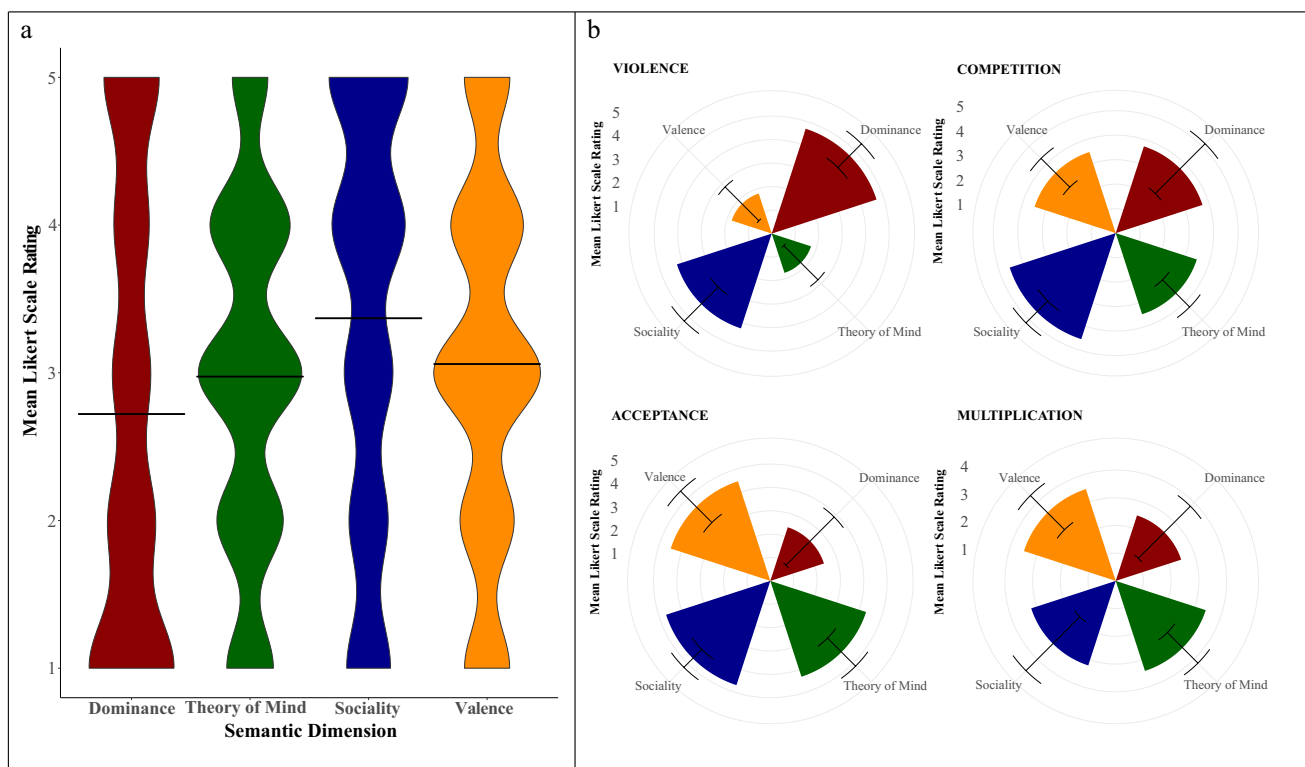

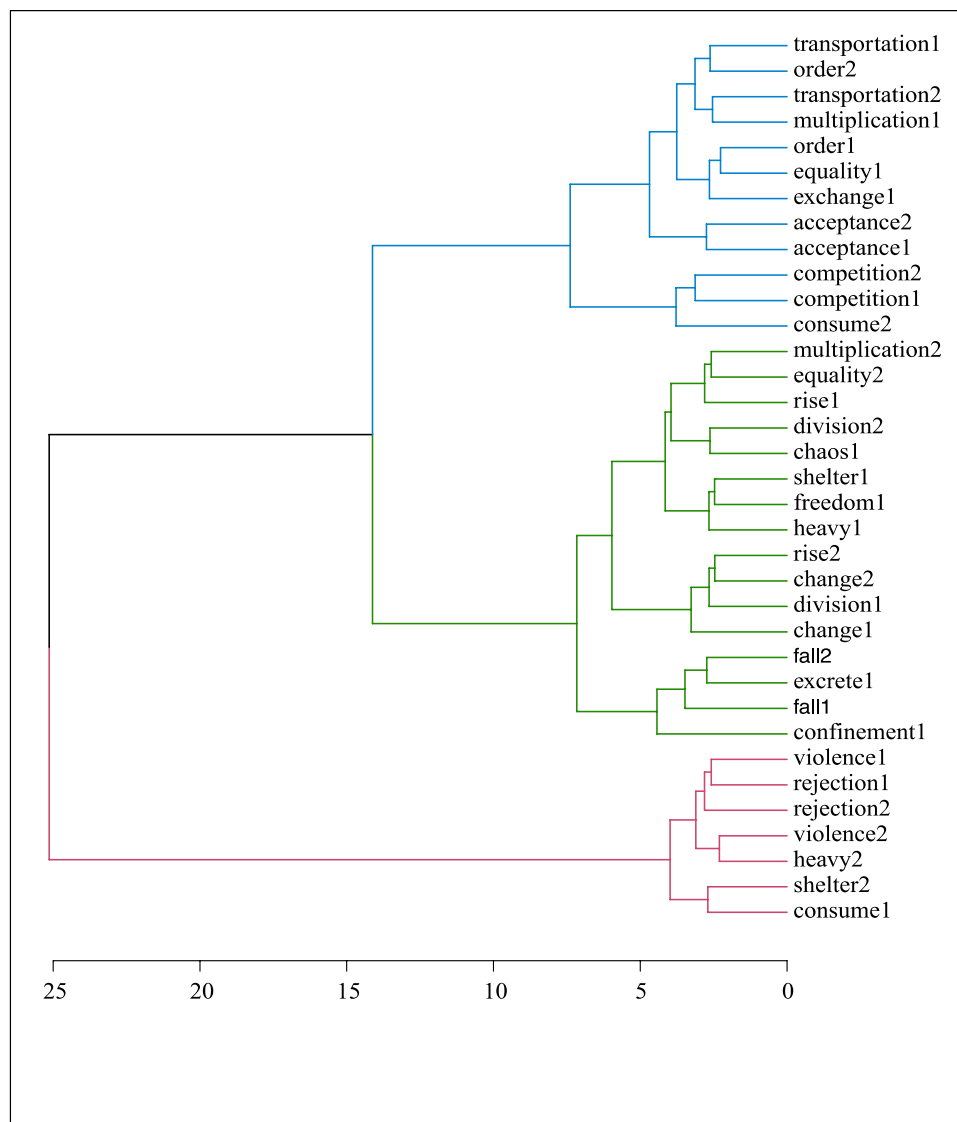**Fig. 4** Semantic dimensions of animations

Fig. 5 Semantic clustering of animations

people might successfully match a football helmet to an oblong ball because they have observed these objects consistently appearing together within the same event schemas. Similarly, a dog who salivates upon hearing a bell in expectation of a food reward does so through classical conditioning (Pavlov, 2010). These examples highlight links between concepts and actions that share some commonalities with symbolic cognition. Yet, these links also substantively differ from symbolic representation in that paired associates do not act as proxies for one another. In the toolkit we have proposed, participants cannot rely on simple familiarity. They are instead compelled to "decode" abstract never-before-seen animations that correspond to broad concepts. This feature of the battery reduces the potential impact of

co-occurrence on the process of mapping symbols (pictures or words) to target animations.

This toolkit also leverages the use of multiple symbol modalities, specifically in the Symbol Identification Task, by including both words and images as matching candidates for each concept animation. Each of the 35 concept animations are paired with a matching word and image along with a series of foils. Some clinical populations present with impairments that compromise a particular symbol modality. For example, following neurological injury, persons with aphasia (PWA) demonstrate a focal reduction in the ability to use and communicate with words (Dell et al., 1997; Levelt et al., 1999). PWA are likely to show poor performance in selecting word

Table 6 Animation cluster contrasts

Contrast	Means	Test statistic (*t*)	Effect size (Cohen's *d*)
C1 vs. C2: Dominance	C1 (2.49), C2 (2.14)	1.10	0.44
C1 vs. C3: Dominance**	C1 (2.49), C3 (4.45)	6.81	2.68
C2 vs. C3: Dominance***	C2 (2.14), C3 (4.45)	13.04	4.48
C1 vs. C2: TOM***	C1 (3.82), C2 (2.89)	4.89	1.78
C1 vs. C3: TOM***	C1 (3.82), C3 (1.87)	10.84	5.25
C2 vs. C3: TOM***	C2 (2.89), C3 (1.87)	4.83	1.94
C1 vs. C2: Sociality***	C1 (3.98), C2 (2.67)	6.06	2.33
C1 vs. C3: Sociality	C1 (3.98), C3 (4.11)	0.58	0.27
C2 vs. C3: Sociality***	C2 (2.67), C3 (4.11)	7.46	2.99
C1 vs. C2: Valence**	C1 (3.83), C2 (3.04)	4.34	1.59
C1 vs. C3: Valence***	C1 (3.83), C3 (1.90)	10.38	5.10
C2 vs. C3: Valence***	C2 (3.04), C3 (1.90)	5.35	2.20

***$p < 0.001$, **$p < 0.01$.

Table 7 Concept dimensions and prompts

Dimension	Survey prompt
Familiarity (self)	Have you ever experienced this concept?
Familiarity (others)	How widely experienced do you think this concept is?

targets relative to images. However, poor performance across both symbol modalities might support a central impairment in symbolic cognition (i.e., asymbolia) vs. a focal lexical impairment.

The Symbol Identification Task can also bolster investigation into symbolic cognition by exploiting error types. There are two foil types (semantically associated, visually similar) with unique relationships to animations that can be used to characterize mapping errors. Low matching accuracy with high selection of visually-similar-to-animation symbols could illustrate poor mapping of stimuli to internal representations. In this case, an individual is relying on low-level visual similarity to select matches. Similarly, low matching accuracy defined by high selection of semantic-associates could indicate some degree of preserved semantic processing without successful symbol mapping.

The use of varied concept categories is another important consideration for evaluating symbolic thinking. This toolkit is constructed to support exploration of categorical conceptual impairments by including a range of abstract concepts. Our concepts tended to cluster into three primary groups based on four dimensions (Dominance, Theory of Mind, Sociality, Valence). The three clusters were defined as Cluster 1: Highly Social and Positive, Cluster 2: Minimally Social, and Cluster 3: Highly Social and Negative. One possibility is that clinical populations who experience limited Theory of Mind would show typical performance when matching Cluster 2 animations and symbols, but poor performance when matching within the other two highly social concepts clusters.

There are few, if any, resources designed to support the evaluation of concept knowledge with labeled and unlabeled stimuli. This toolkit can be used to investigate concept knowledge without reliance on verbal labels using the concepts-without-words definitions where 80% of participants experienced the final set of 51 concept definitions.

This freely available toolkit (visit https://osf.io/9jr6h) should support a more refined and comprehensive evaluation of symbolic knowledge and impairments in various populations. Much remains unclear about symbolic cognition despite its key role in human communication. In order to more effectively investigate and characterize symbolic cognition we must consider the dynamic, fluid nature of symbolic representation and the relationship between its essential components: concepts, symbols, and symbol-mapping. Here we offer an open-source toolkit designed to evaluate symbolic cognition that may be used to improve our understanding of symbolic processing and the nature of asymbolia across the life span.
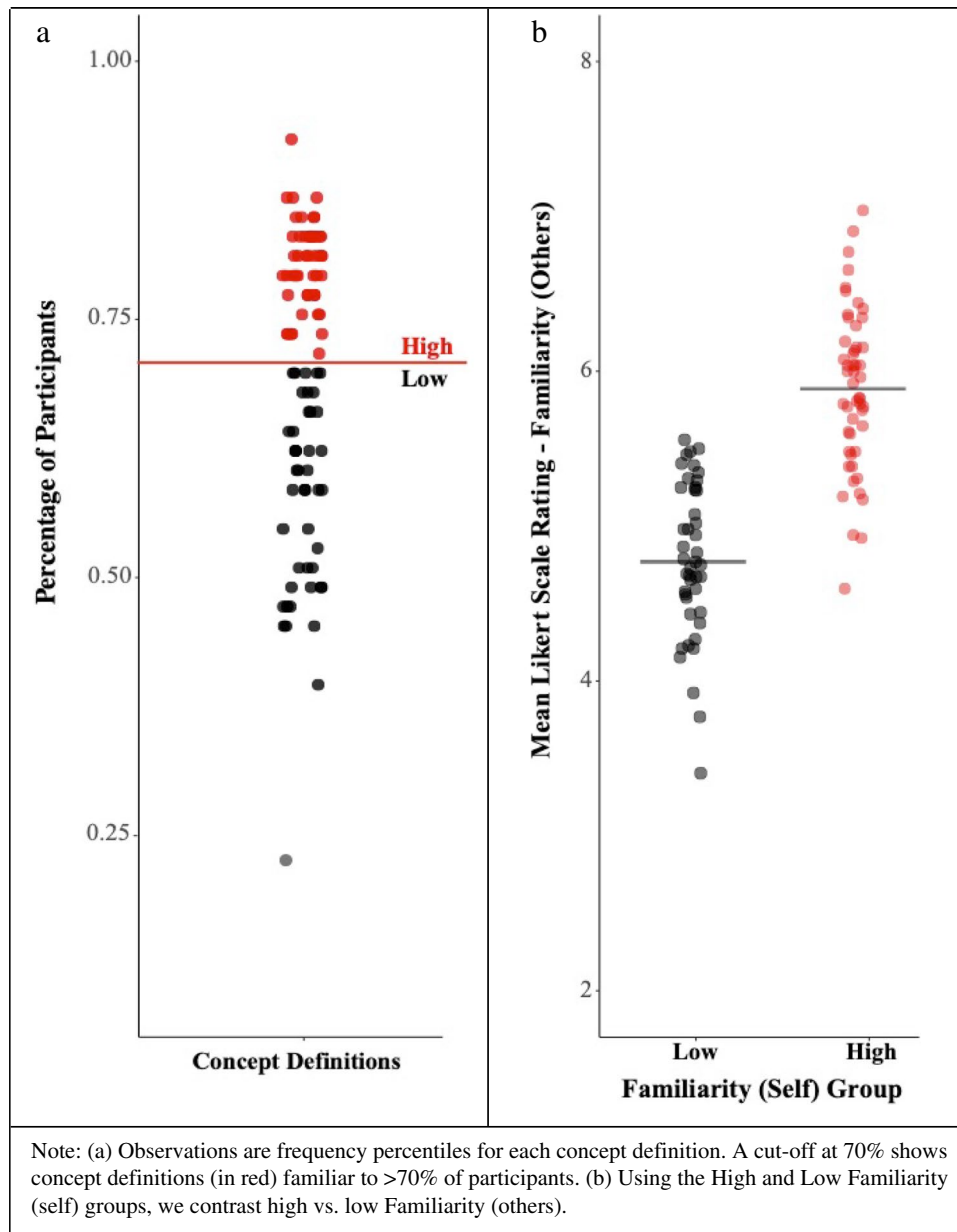
Note: (a) Observations are frequency percentiles for each concept definition. A cut-off at 70% shows concept definitions (in red) familiar to >70% of participants. (b) Using the High and Low Familiarity (self) groups, we contrast high vs. low Familiarity (others).

Fig. 6 High and low Familiarity concepts

Declarations

Conflicts of Interest The authors have declared no conflicts of interest.

Consent to Participate Informed consent was obtained from all individuals included in this study.

References

Antonucci, S., Beeson, P., & Rapcsak, S. (2004). Anomia in patients with left inferior temporal lobe lesions. *Aphasiology, 18*(5–7), 543–554. https://doi.org/10.1080/02687030444000219

Barron, F. (1988). Putting creativity to work. In *The nature of creativity: Contemporary psychological perspectives* (76–98). Cambridge University Press.

Bobbitt, D. (2005). The Essential Peirce: Selected Philosophical Writings, Volume 2 (1893–1913). *American Journal of Semiotics, 21*, 77–78.

Bozeat, S., Lambon Ralph, M., Graham, K., Patterson, K., Wilkin, H., Rowland, J., Rogers, T., & Hodges, J. (2003). A duck with four legs: Investigating the structure of conceptual knowledge using picture drawing in semantic dementia. *Cognitive Neuropsychology, 20*(1), 27–47.

Bruner, J. (1964). The course of cognitive growth. *American Psychologist, 19*(1), 1–15. https://doi.org/10.1037/h0044160

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Cicone, M., Wapner, W., Foldi, N., Zurif, E., & Gardner, H. (1979). The relation between gesture and language in aphasic communication. *Brain and Language, 8*(3), 324–349. https://doi.org/10.1016/0093-934X(79)90060-9

Coelho, C., & Duffy, R. (1987). The relationship of the acquisition of manual signs to severity of aphasia: A training study. *Brain and Language, 31*(2), 328–345. https://doi.org/10.1016/0093-934X(87)90078-2

Deacon, T. (1998). *The Symbolic Species*. W.W. Norten & Company.

Deacon, T. (2011). The symbol concept. *Oxford University Press*. https://doi.org/10.1093/oxfordhb/9780199541119.013.0043

Deacon, T., Velichkovsky, B., & Rumbaugh, D. (1996). Prefrontal cortex and symbol learning: Why a brain capable of language evolved only once. In *Communicating meaning: The evolution and development of language*. (pp. 103–138). Lawrence Erlbaum Associates, Inc. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1996-98786-004&site=ehost-live

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science (New York, NY), 311*(5759), 381–384. https://doi.org/10.1126/science.1121739

Dell, G., Schwartz, M., Martin, N., Saffran, E., & Gagnon, D. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801–838.

DeLoache, J. (1995). Early Understanding and Use of Symbols: The Model. *Current Directions in Psychological Science, 4*(4), 109–113. https://doi.org/10.1111/1467-8721.ep10772408

DeLoache, J. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*(2), 66–70. https://doi.org/10.1016/j.tics.2003.12.004

Duffy, R., & Buck, R. (1979). A Study of the Relationship between Propositional (Pantomime) and Subpropositional (Facial Expression) Extraverbal Behaviors in Aphasics. *Folia Phoniatrica Et Logopaedica, 31*(2), 129–136. https://doi.org/10.1159/000264160

Duffy, R., & Duffy, J. (1981). Three Studies of Deficits in Pantomimic Expression and Pantomimic Recognition in Aphasia. *Journal of Speech, Language, and Hearing Research, 24*(1), 70–84. https://doi.org/10.1044/jshr.2401.70

Duffy, R., Duffy, J., & Pearson, K. (1975). Pantomime Recognition in Aphasics. *Journal of Speech and Hearing Research, 18*(1), 115–132. https://doi.org/10.1044/jshr.1801.115

Duffy, R., & Liles, B. (1979). A Translation of Finkelnburg's (1870) Lecture on Aphasia as "Asymbolia" with Commentary. *Journal of Speech and Hearing Disorders, 44*(2), 156–168. https://doi.org/10.1044/jshd.4402.156

Garcez, A. d'Avila, Gori, M., Lamb, L. C., Serafini, L., Spranger, M., & Tran, S. N. (2019). Neural-Symbolic Computing: An Effective Methodology for Principled Integration of Machine Learning and Reasoning. ArXiv:1905.06088 [Cs]. http://arxiv.org/abs/1905.06088

Geraghty, K., Waxman, S., & Gelman, S. (2014). Learning words from pictures: 15- and 17-month-old infants appreciate the referential and symbolic links among words, pictures, and objects. *Cognitive Development, 32*, 1–11. https://doi.org/10.1016/j.cogdev.2014.04.003

Goldenberg, G., & Randerath, J. (2015). Shared neural substrates of apraxia and aphasia. *Neuropsychologia, 75*, 40–49. https://doi.org/10.1016/j.neuropsychologia.2015.05.017

Goldstein, K. (1936). The Significance of the Frontal Lobes for Mental Performances. *Journal of Neurology, Neurosurgery & Psychiatry, s1-17*(65), 27–40. https://doi.org/10.1136/jnnp.s1-17.65.27

Goldstein, K. (1948). *Language and language disturbances; aphasic symptom complexes and their significance for medicine and theory of language.* http://psycnet.apa.org/psycinfo/1949-03819-000

Hayes, K., & Hayes, C. (1953). Picture perception in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology, 46*(6), 470–474. https://doi.org/10.1037/h0053704

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology, 57*(2), 243. https://doi.org/10.2307/1416950

Humphreys, G., & Riddoch, M. (2003). A case series analysis of "category-specific" deficits of living things: The HIT account. *Cognitive Neuropsychology Special Issue: The Organisation of Conceptual Knowledge in the Brain: Neuropsychological and Neuroimaging Perspectives, 20*(36), 263–306.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75.

Lyn, H. (2007). Mental representation of symbols as revealed by vocabulary errors in two bonobos (Pan paniscus). *Animal Cognition, 10*(4), 461–475. https://doi.org/10.1007/s10071-007-0086-3

McCarthy, R., & Kartsounis, L. (2000). Wobbly words: Refractory anomia with preserved semantics. *Neurocase, 6*(6), 487–497. https://doi.org/10.1080/13554790008402719

Namy, L., Campbell, A., & Tomasello, M. (2004). The Changing Role of Iconicity in Non-Verbal Symbol Learning: A U-Shaped Trajectory in the Acquisition of Arbitrary Gestures. *Journal of Cognition and Development, 5*(1), 37–57. https://doi.org/10.1207/s15327647jcd0501_3

Nelson, D., McEvoy, C., & Schreiber, T. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments & Computers, 36*(3), 402–407.

Noun Project. (n.d.). Noun Project. Retrieved July 12, 2020, from https://thenounproject.com

Parmentier, R. (2016). Peirce and Saussure on Signs and Ideas in Language. In *Signs and Society* (42–60). Indiana University Press. https://doi.org/10.2307/j.ctt2005sns.7

Pavlov, I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences, 17*(3), 136–141. https://doi.org/10.5214/ans.0972-7531.1017309

Pedersen, T., Patwardhan, S., & Michelizzi, J. (2004). WordNet::Similarity: Measuring the relatedness of concepts. *Demonstration Papers at HLT-NAACL 2004 on XX - HLT-NAACL '04*, 38–41. https://doi.org/10.3115/1614025.1614037

Pickett, L. (1974). An assessment of gestural and pantomimic deficit in aphasic patients. *Acta Symbolica, 5*, 69–86.

Rakoczy, H., Tomasello, M., & Striano, T. (n.d.). How Children Turn Objects Into Symbols: A Cultural Learning Account. *Symbol Use and Symbol Representation*, 67–97.

Reilly, J., Flurie, M., Ungrady, M. (2020). Eyetracking during picture naming predicts future vocabulary dropout in progressive anomia. *Neuropsychological Rehabilitation*, 1–19 https://doi.org/10.1080/09602011.2020.1835676

Reilly, J., Peelle, J., Antonucci, S., & Grossman, M. (2011). Anomia as a marker of distinct semantic memory impairments in Alzheimer's disease and semantic dementia. *Neuropsychology, 25*(4), 413–426. https://doi.org/10.1037/a0022738

Saussure, F. (1916). *Cours de linguistique generale (1907)* (C. Bally & A. Sechehaye, Eds.). Philosophical Library.

Savage-Rumbaugh, E. (1986). *Ape language: From conditioned response to symbol*. Columbia University Press. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=psyh&AN=1986-98215-000&site=ehost-live

Savchenko, A., Alekseev, A., Kwon, S., Tutubalina, E., Myasnikov, E., & Nikolenko, S. (2020). Ad Lingua: Text Classification Improves Symbolism Prediction in Image Advertisements. *Proceedings of the 28th International Conference on Computational Linguistics*, 1886–1892. https://doi.org/10.18653/v1/2020.coling-main.171

Seyfarth, R., Cheney, D., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*(4), 1070–1094. https://doi.org/10.1016/S0003-3472(80)80097-2

Shusterman, A., Ah Lee, S., & Spelke, E. S. (2008). Young children's spontaneous use of geometry in maps. *Developmental Science, 11*(2), F1–F7. https://doi.org/10.1111/j.1467-7687.2007.00670.x

Ungrady, M., Flurie, M., Zuckerman, B., Mirman, D., & Reilly, J. (2019). Naming and knowing revisited: Eyetracking correlates of anomia in progressive aphasia. *Frontiers in Human Neuroscience, 13*, 354. https://doi.org/10.3389/fnhum.2019.00354

Varney, N. (1978). Linguistic correlates of pantomime recognition in aphasic patients. *Journal of Neurology, Neurosurgery & Psychiatry, 41*(6), 564–568. https://doi.org/10.1136/jnnp.41.6.564

Vygotsky, L. (1967). Play and Its Role in the Mental Development of the Child. *Soviet Psychology, 5*(3), 6–18. https://doi.org/10.2753/RPO1061-040505036

Werner, H., & Kaplan, B. (1963). *Symbol formation*. Wiley.