Affordance Norms for XXXX Concrete Nouns

Nicholas P. Maxwell1, Alen Hajnal2, Mark J. Huff2, Julia Blau3, Brian Day4, Kerry Marsh5, Benjamin Meagher6, Chris Pagano7, Jack Shelley-Tremblay8, & Jeff Wagman9

1Midwestern State University, 2The University of Southern Mississippi, 3Central Connecticut State University, 4Butler University, 5University of Connecticut, 6Hope College, 7Clemson University, 8The University of South Alabama, 9Illinois State University

**Author Note**

Correspondence regarding this article should be addressed to Nicholas P. Maxwell, Department of Psychology, Midwestern State University, 3410 Taft Blvd, Wichita Falls, TX, 76308, United States. Email: nicholas.maxwell@msutexas.edu. The final set of affordance norms is available for download via our OSF page [LINK]. The normed dataset can also be accessed via our interactive Shiny application [LINK]. The authors thank Morgan Ballesteros, Samantha Garcia, and Madisyn Metaxas for their assistance with cleaning the final dataset.

Abstract

[WORDS HERE]

Word Count: XXXX

*Keywords*: Affordances; Body-Object Interaction; Word Norms; Database; *R* Shiny

Affordance Norms for 3000 Highly Concrete Objects

The ability to process and retain concept information is a critical aspect of human cognition. Our collected knowledge of everyday objects allows us to successfully navigate our environment while also providing general information regarding the world around us. For researchers who study the cognitive processes underlying language and memory, having accurate measurements of what words mean and how individuals use them together is critical. Traditionally, researchers have classified concept information into two distinct categories: Lexical properties which describe linguistic characteristics of individual words w(e.g., frequency [CITE]; concreteness, [CITE]; age-of-acquisition [CITE]; etc.) and word-pair properties which describe various aspects of relatedness between cue-target word pairs (e.g., semantic feature similarity, [CITE]; word associations, [CITE]; etc.). While it’s important to account for both types of concept information when designing an experiment, the present study is primarily concerned with meaning-based measures of concepts rather than lexical properties related to a word’s written form.

From an empirical standpoint, an object’s meaning can be operationalized in a variety of ways. In practice, however, cognitive psychologists often make use of two broad classes of concept information when measuring meaning: Semantic representations (which can be further divided into XXX) and word associations. First, semantic features describe an object’s meaning in terms of its constituent features. For example, *dogs* have *tails*, *fur*, and *bark*, while *chairs* are comprised of *legs*, a *back*, and a *seat*, and depending on the design, might be made of *wood* or *metal*. To empirically measure semantic features, researchers commonly use feature production tasks, in which participants are presented with a series of objects and are asked to list the most important facets that comprise each item (see XXX, for review). By employing these tasks in large-scale norming studies, researchers have developed sets of semantic feature production norms (e.g., EXAMPLE; EXAMPLE; EXAMPLE), which can be used to assess the semantic similarity between any two measured concepts (see XXX). The past two decades have seen a proliferation of these norms, with feature production norms now available in a variety of languages (see XXX). Second, researchers can separately assess a concept’s meaning in terms of its connections with other objects. Termed associations, these measures are often described in terms of response probabilities (i.e., the likelihood that *cat* would elicit meow as a response; see XXX) and are assumed to form whenever words repeatedly co-occur with one another in natural language (Fishler, 1977; Nelson, McEvoy, & Dennis, 2000). As a result, associations can capture a variety of knowledge, including semantically unrelated concepts that are frequently used together (e.g., *peanut* and *butter*) as well as semantic knowledge (e.g., *drive* and *car*). Like semantic features, associations can similarly be captured via norming studies, and several of these sets of free association norms are readily available (EXAMPLE, EXAMPLE, EXAMPLE).

While semantic and associative measures are critical for understanding a variety of cognitive processes (e.g., memory, comprehension, perception, etc.), exclusively relying upon these measures results in a limited view of meaning. Specifically, when processing an object’s meaning, individuals also consider it in terms of its perceived use or functionality [SURVIVAL PROCESSING, ONE MORE EXAMPLE?] [TRANSITION TO AFFORDANCES] (i.e., affordances; Gibson, 1977). Unlike semantic-based measures, affordances describe interactive relationships occurring between an actor and an object rather than reflecting a specific object or its relationship to other concepts. For example, *chair* is associated with *table*, *seat*, and *couch* and is also comprised of several semantic features (*legs*, *back*, *armrests*, etc.), it also affords *sitting*, *pushing*, and *standing* upon. However, traditional measures of meaning are not likely to capture these actions. For example, [PULL SOME NUMBERS HERE] Thus, [SUMMARY SENTENCE]

A growing body of research suggests that [embodied stuff here]

While several normed databases are available online which catalogue semantic features and word associations, to date, no norming study has had participants generate affordances for objects. This is surprising, given that Psychology has a rich history of conducting mega-studies to generate standardized stimuli. However, [TRANSITION TO BOI]

[BOI HERE – DESCRIBE + SHORTCOMINGS]

[ALTERNATIVE APPROACH – OPEN ENDED W/ MULTIPLE RESPONSE]

[EXPAND HERE] [WHY IS UNDERSTANDING INTERACTIVE PROPERTIES IMPORTANT?] [SURVIVAL PROCESSING MAYBE?] [DIFFERENT “TYPES” OF SEMANTIC REPRESENTATIONS?] [BOI FACILITATES SEMANTIC PROCESSING]

**The Present Study**

[GOAL OF THE PRESENT STUDY] [FIRST DETAIL THE CREATION OF THE DATASET, THEN THE ONLINE INTERFACE, FINALLY DESCRIBE A SERIES OF VALIDATION STUDIES]

**Method**

**Participants**

Participants were recruited across two broad settings. First, XX undergraduate students were recruited from XX universities within the eastern and southern United States (see Table x for individual *n*s for each testing site). Second, an additional XX participants completed the study online via Prolific (www.Prolific.co). All university students completed the study in exchange for partial course credit, while Prolific participants were compensated at a rate of $3.00 per 20-minute session. Prolific participants were required to have completed at least a high-school level degree or equivalent and to be native English speakers. For completeness, demographic information is reported in the Appendix (Table Ax).

**Materials**

To generate the stimuli, we selected 3005 nouns from the MRC psycholinguistic database (CITE). Words were initially selected based on concreteness, such that only high concreteness words were included (concreteness ≥ xx). Of the 3005 words that were initially generated, five were randomly selected to serve as practice items. The remaining 3000 items were then once randomized before being equally split into 100 list containing 30 items each. Overall, the final set of 3000 words had a mean concreteness rating of XX, mean SUBTLEX frequency rating of XX (Brysbaert & New, 2009), and a mean BOI rating of XX (Pexman et al., 2019). Full descriptive statistics, including the percentage of stimuli overlapping with existing lexical datasets, are displayed in Table X.

**Procedure**

For all testing sites, data collection occurred online via Collector, an open-source platform for conducting web-based psychological experiments (Garcia & Kornell, 2015). Prior to beginning the norming task, particpants were informed that they would be presented with a series of object words and that they would be required to list as many uses for each object as they could reasonably generate. Participants were reminded that a single object typically has multiple uses and were encouraged to list multiple object uses when possible. To illustrate this point, the word *ball* was provided as an example, with *throw*, *bounce*, and *step on* all provided as examples of possible affordances. The complete set of instructions can be viewed at [OSF link].

Following the initial set of instructions, participants completed a five-item practice set, which familiarized them with the norming task. For all trials, words were presented individually in the center of the screen, and participants were instructed to type each affordance into a textbox located directly below the word. To maximize potential affordances, participants were not given specific instructions on how to format their responses (i.e., tense, single words vs. phrases, etc.) except that they were instructed to separate each unique affordance with a comma. After completing the practice trials, participants immediately began the full study, which randomly presented them with a list of 30 items randomly selected from one of the 100 item lists. Responses for each list were self-paced, and above each trial was a prompt which reminded participants to list as many uses for each object as they could generate. Following completion of this task, participants were debriefed. The total study took approximately 20 minutes to complete.

**Data Processing**

Data processing was conducted in *R* following Buchanan, De Deyne, & Montefinese’s (2020) guidelines for processing lexical output from norming studies. Below, we detail each step used to create the final dataset before describing the calculation of two affordance measures: Affordance Strength (AFS) and Affordance Set Size (AFFS). Given both the predicted size of the final dataset and because data collection was split across multiple institutions, the data processing steps listed below were conducted across several batches of data, which ranged from approximately 25 to 500 participants collected at a single testing site.

***Missing Data and Parsing Affordances.*** First, all blank responses were removed from the dataset as well as any responses suggesting that participants were unfamiliar with a specific object (e.g., “I don’t know”, “unknown”, “unsure”, “?”, etc.). [WHAT PERCENT OF THE DATA WAS REMOVED?]. Second, because participants typically provided multiple affordances to each object, the XXX package was used to parse out individual affordances for each object. This resulted in a long-format dataset, with each individual affordance having its own row in the dataset (i.e., “to drink from, throw it, pencil holder” become “to drink from”, “throw it”, and “pencil holder”).

***Spelling Corrections and Tokenization.*** [WORDS HERE]

***Lemmatization and Part of Speech Tagging.*** [WORDS HERE]

[STARTED W/ 3000 ITEMS, ANY ITEMS REMOVED?]

[HOW MANY PARTICIPANTS WERE REMOVED?]

[POSSIBLY A FIGURE?]

[HOW WERE AFS AND AFFS COMPUTED?]

**Results**

[WORDS HERE]

**Research Questions**

[WORDS HERE]

**Descriptive Data**

[WORDS HERE]

**Validity**

[WORDS HERE]

**Discussion**

[WORDS HERE]

**Conclusion**

[WORDS HERE]

**References**

[FIRST ONE HERE]