J Psycholinguist Res
https://doi.org/10.1007/s10936-017-9544-5

Thai Norms for Name, Image, and Category Agreement, Object Familiarity, Visual Complexity, Manipulability, and Age of Acquisition for 480 Color Photographic Objects

A. J. Benjamin Clarke[1] · **Jason D. Ludington**[2]

Abstract Normative databases containing psycholinguistic variables are commonly used to aid stimulus selection for investigations into language and other cognitive processes. Norms exist for many languages, but not for Thai. The aim of the present research, therefore, was to obtain Thai normative data for the BOSS, a set of 480 high resolution color photographic images of real objects (Brodeur et al. in PLoS ONE 5(5), 2010. https://doi.org/10.1371/journal.pone.0010773). Norms were provided by 584 Thai university students on eight dimensions: name agreement, object familiarity, visual complexity, category agreement, image agreement, two types of manipulability (graspability and mimeability), and age of acquisition. The results revealed comparatively similar levels of name agreement to Brodeur et al. especially when unfamiliar items were factored out. The pattern of intercorrelations among the Thai psycholinguistic norms was comparable to previous studies and our cross-linguistic correlations were robust for the same set of pictures in English and French. Conjointly, the findings extend the relevancy of the BOSS to Thailand, supporting this photographic resource for investigations of language and other cognitive processes in monolingual, multilingual, and brain-impaired populations.

Keywords Thai · Picture norms · Cross-linguistic · Bank of Standardized Stimuli (BOSS)

Introduction

Researchers interested in memory, perception, and language have often used their own visual stimuli in experiments. Unfortunately, pictures tend to be highly idiosyncratic and vary along several peripheral dimensions (e.g., the way they are designed or drawn, how familiar they

✉ A. J. Benjamin Clarke
ajb.clarke@gmail.com; clarke@tu.ac.th

1 Department of English and Linguistics, Thammasat University, Khlong Luang, Pathum Thani, Thailand

2 Faculty of Psychology, Chulalongkorn University, Bangkok, Thailand

Published online: 08 December 2017

are). Similarly, the words that refer to the objects being depicted in the pictures can also vary quite considerably in their characteristics. They may vary in length, how frequently they occur, and the age at which they are acquired. Agreement on whether a picture is a good representation of a particular word is not altogether straightforward either. Pictures may bring to mind a number of different names in much the same way as an individual word can evoke many different images. These idiosyncrasies can potentially have unintended and unaccounted-for effects on measurements of interest. For this reason, the use of standardized stimuli is vital for balanced stimulus selection or accounting for sources of influence on study effects.

In one of the first normative studies, Snodgrass and Vanderwart (1980) obtained data for 260 black-and-white pictures that consisted of everyday common objects. Norms were provided by young American English speaking adults for the variables of name agreement, image agreement, familiarity, and visual complexity. Snodgrass and Vanderwart's picture database has been widely used in several areas of cognitive science, including aphasiology, cognitive psychology, neuropsychology, and psycholinguistics. For example, the line drawings have been used to investigate the psycholinguistic mechanisms underlying normal and impaired speech production in experimental tasks such as picture naming (Cattell 1886), picture-word interference (Rosinski et al. 1975), and the visual-world paradigm (Tanenhaus et al. 1995). These paradigms have helped elucidate the relationships between lexical, phonological and semantic processes involved in the comprehension and production of language—for both monolinguals and bilinguals (e.g., Belke et al. 2005; Damian and Martin 1998; Hanulová et al. 2011). Picture databases and their corresponding normative data, therefore, have been instrumental in deepening our understanding of how the human brain processes words and pictures.

Despite the widespread use of the Snodgrass and Vanderwart (1980) database, it was quickly apparent that normative data were required for different linguistic and cultural contexts. For instance, objects that commonly appear in one culture might appear much less frequently in another culture; similarly, pictures may evoke higher levels of specificity, or elicit higher name agreement from some groups compared to others (e.g., Sanfeliù and Fernandez 1996). These issues were exemplified in a study by Yoon et al. (2004), who had American and Chinese participants generate names to the original 260 Snodgrass and Vanderwart pictures in addition to rating them for familiarity. The findings revealed clear cross-cultural differences for a large proportion of the pictures on the measures of name and concept agreement. In a related study, Sanfeliù and Fernandez (1996) observed language-specific effects on name and image agreement when they compared their Spanish data with data from Japanese and English speakers. Significant differences were found on these measures despite cross-linguistic agreement for familiarity and visual complexity ratings across the same three languages. Taken together, these findings illustrate the importance of standardizing stimuli for the cultural or linguistic group that is under consideration. Indeed, norms for the Snodgrass and Vanderwart (1980) pictures have been collected for other cultures and language communities, including British English (Barry et al. 1997), Chinese (Liu et al. 2011; Weekes et al. 2007), Dutch (Martein 1995), French (Alario and Ferrand 1999; Bonin et al. 2003), Icelandic (Pind et al. 2000), Italian (Dell'Acqua et al. 2000), Japanese (Nishimoto et al. 2005), Portuguese (Pompéia et al. 2001), Spanish (Álvarez and Cuetos 2007; Sanfeliù and Fernandez 1996), Tunisian Arabic (Boukadi et al. 2016), and Bulgarian, German, and Hungarian (Bates et al. 2003).

Notwithstanding this large body of work, some researchers have raised concerns about the ecological validity of black-and-white line drawings (e.g., Brodeur et al. 2010; Moreno-Martínez and Montoro 2012; Viggiano et al. 2004). Undoubtedly, line drawings fail to adequately reflect the rich visual detail and complexity (i.e., color and texture) of real-world objects. It is also fairly well established that color and other visual details can facilitate object recognition in terms of both speed and accuracy (e.g., Ostergaard and Davidoff 1985; Price and Humphreys 1989; Tanaka et al. 2001)—thus some researchers have created databases with more ecologically valid object representations. For example, Rossion and Pourtois (2004) produced a color version of the Snodgrass and Vanderwart (1980) pictures and reported that these yielded more consistent name agreement and faster picture-naming latencies compared to the original black-and-white line drawings for their French-speaking participants (see also Dimitropoulou et al. 2009; Raman et al. 2014; Tsaparina et al. 2011).

High-resolution color photographs offer even more realistic and detailed representations of objects, while capturing other important ecological features such as texture and volume (third-dimension schematics; Brodeur et al. 2010). Improvements in information technology have enabled researchers to employ such images in their experiments and, increasingly, databases of normative photographic stimuli are being constructed and relied upon. Although several such normative databases exist (e.g., Adlington et al. 2009; Brodeur et al. 2010, 2014), only a handful apply to languages other than English (e.g., Brodeur et al. 2012, for Canadian French; Moreno-Martínez and Montoro 2012, for Spanish; Viggiano et al. 2004, for Italian; Shao and Stiegert 2016, for Dutch). Currently, the largest available normative photographic database is the Bank of Standardized Stimuli (BOSS; Brodeur et al. 2010).[1] The BOSS offers a useful set of freely accessible, high-quality, standardized stimuli to facilitate research across a wide variety of disciplines. The BOSS thus allays the ecological concerns that were raised earlier about black-and-white line drawings.

Here we sought to standardize the 480-item picture bank for Thai speakers on variables which we felt were most important to research in cognition and psycholinguistics: the Snodgrass and Vanderwart (1980) variables (name agreement, familiarity, visual complexity and image agreement), as well as age of acquisition (Carroll and White 1973), category agreement (Brodeur et al. 2010) and two types of manipulability (how easily items can be grasped and mimed; Magnié et al. 2003; Salmon et al. 2010). We next define these variables and review the empirical findings relating to each one.

Name Agreement

Name agreement is a measure of the extent to which different people agree on the name(s) given to a particular picture. Pictures that elicit a single name have a high name agreement score while those eliciting multiple names have a low name agreement score. It is a robust predictor of naming speed, with lower levels of name agreement corresponding to longer picture-naming latencies (Alario et al. 2004; Barry et al. 1997; Bonin et al. 2003; Ellis and Morrison 1998; Gilhooly and Gilhooly 1979; Liu et al. 2011; Snodgrass and Yuditsky 1996), perhaps because of the additional time needed to select between competing alternative names (e.g., Alario et al. 2004). Name agreement is thought to operate at the level of lexical selection and/or phonological encoding (Barry et al. 1997). Name agreement studies also show that there is considerable semantic variability in the dominant responses when compared cross-linguistically (e.g., Kremin et al. 2003; see also Sanfeliù and Fernandez 1996; Yoon et al. 2004), highlighting the need for studies such as the current one.

[1] Brodeur and colleagues have recently extended the size of the BOSS by providing norms for a further 930 items, increasing the size of the database from 480 items to 1410 items (Brodeur et al. 2014).

Object Familiarity

Object familiarity (or concept familiarity) refers to the familiarity of the concept depicted. It is defined as the extent to which a person is engaged with (or comes into contact with) the concept in daily life. Object familiarity is thought to involve access to semantic representations (Hirsh and Funnell 1995) and has been shown to impact performance on various cognitive processing tasks, including memory (for a review, see Gernsbacher 1984). Object familiarity reliably predicts picture naming latencies, with more familiar objects being named faster than less familiar objects (Cuetos et al. 1999; Ellis and Morrison 1998; Snodgrass and Yuditsky 1996).

Visual Complexity

Visual complexity refers to the subjective quantity of lines and detail used to depict an object in a picture or photograph. It is thought to determine the ease of processing at the structural stage of object recognition and visual complexity has been found to affect tachistoscopic recognition thresholds as well as memorability. Some studies have linked visual complexity to naming latency, reporting longer naming latencies for more visually complex images (Alario et al. 2004; Ellis and Morrison 1998; Humphreys et al. 1988). However, color might be a mitigating factor in determining the influence of visual complexity on object naming (see e.g., Biederman 1987; Zannino et al. 2010) given many studies have failed to obtain such an effect (Barry et al. 1997; Bonin et al. 2003; Cuetos et al. 1999; Snodgrass and Yuditsky 1996; Weekes et al. 2007).

Manipulability

Manipulability is another important variable in studies of object processing, especially from an embodied cognition perspective (e.g., Mahon and Caramazza 2008). Some studies have shown that non-manipulable objects are identified or processed faster than manipulable ones (e.g., Filliter et al. 2005), whereas others have found the opposite (e.g., Guérard et al. 2015; Magnié et al. 2003), an inconsistency which probably stems from definitional and procedural differences. Manipulability has been defined in different ways. Two important ones are how easily objects can be handled (*graspability*; e.g., Salmon et al. 2010), and how easily they can be used (*mimability*; e.g., Magnié et al. 2003). Gibson's (2014) ecological approach to visual perception places object motor affordance in the same category as object perception and knowledge. Recent neurological studies support this: Neural networks for semantic and category knowledge are located along motor and perceptual brain areas (Hauk et al. 2004; Huth et al. 2012). Graspability and mimability roughly correspond to structural and functional levels of object knowledge respectively (Jax and Buxbaum 2010), and, accordingly, need to be taken into account for a more complete understanding of object perception.

Guérard et al. (2015) found that the ease with which objects can be grasped and moved were two highly related measures ($r = .973$). In the current study we combined these measures into a single rating task that was based on ease of grasp to move (previously measured as "grasp to lift"; Guérard et al.). Regarding object mimability, one common measuring technique is to ask how easily an action could be produced for each object to permit onlookers to unambiguously identify it (Magnié et al. 2003). Following Guérard et al., we removed the potentially difficult perspective-taking aspect ("to permit onlookers to unambiguously identify it") from this task and had objects rated on how easily their usage could be mimed.

Category Agreement

Category agreement or typicality (e.g., Dell'Acqua et al. 2000) refers to the extent to which an object is deemed representative of its category. This variable has been shown to influence performance in tasks involving semantic processing and memory; it is also considered an important factor in studies of conceptual development and category learning (e.g., Rosch 1978). It has been often overlooked in normative studies and semantic tests, despite its known influence on naming performance by neurological patients such as aphasics and those with semantic dementia (Moreno-Martínez and Montoro 2012). In both cognitive and neuropsychological studies, there is evidence showing that object processing is influenced by semantic category (for a review, see Capitani et al. 2003). Category-specific deficits in object recognition among brain-damaged patients occur more frequently for objects belonging to biological (e.g., animals, fruits) as opposed to man-made (e.g., tools, furniture) categories (e.g., Silveri et al. 2002). Even normal participants exhibit longer latencies and are less accurate when identifying biological compared to man-made objects (e.g., Gerlach 2001; but see Laws and Neve 1999). Category agreement was therefore included in order to facilitate further cognitive and neuropsychological research on how category membership may affect object knowledge and processing.

Image Agreement

Image agreement refers to the extent to which a mental image formed in response to seeing or hearing an object's name matches its given pictorial representation. Such ratings are assumed to rely on visual processing and reflect access to an object's stored structural representation (Humphreys et al. 1988). Pictures exhibiting a greater degree of overlap with their stored structural representations tend to be recognized or processed faster, ultimately leading to faster and more accurate retrieval of their corresponding names in picture-naming tasks (Alario et al. 2004; Barry et al. 1997; Bonin et al. 2003). Image agreement appears to be language-specific (e.g., Sanfeliù and Fernandez 1996) and is also influenced by color (e.g., Weekes et al. 2007; Zannino et al. 2010).

Age of Acquisition

Age of acquisition is an estimate of the age at which words are learned, typically by asking adults to make ratings according to age-band categories (Gilhooly and Logie 1980). Age of acquisition has also been measured objectively by analyzing the performance of children at different ages in picture naming tasks and, strikingly, subjective age of acquisition ratings provided by adults correlate strongly with objective measures based on children (Álvarez and Cuetos 2007; Chalard et al. 2003; Grigoriev and Oshhepkov 2013; Lotto et al. 2010; Morrison et al. 1997; Pind et al. 2000). Indeed, many researchers have argued that subjective age of acquisition is a valid indicator of the actual age at which words are acquired and can therefore be reliably used as an adequate substitute in the absence of objective age of acquisition data (e.g., Chalard et al. 2003; Ellis and Morrison 1998; Morrison et al. 1997). It is presumably for this reason that many normative studies have favored subjective age of acquisition ratings over objective measures, especially when the aim has been to standardize a large number of variables within the same study.

Age of acquisition is clearly an important variable with highly robust effects (for reviews, see Johnston and Barry 2006; Juhasz 2005). Its influence has been widely observed across different types of stimuli (i.e., words, pictures, faces) as well as on a wide variety of lexical and semantic processing tasks (Cortese and Khanna 2007; Morrison and Ellis 1995). In word naming, picture naming, and lexical decision tasks, faster and more accurate performance is typically found with words and concepts acquired earlier in life (Barry et al., 2007; Caroll & White, 1973; Cortese and Khanna 2007; Cuetos et al. 1999; Ellis and Morrison 1998; Gilhooly and Gilhooly 1979; Morrison and Ellis 1995; Snodgrass and Yuditsky 1996; Turner et al. 1998). Importantly, age of acquisition accounts for significant picture naming variance regardless of whether rated or objective age of acquisition measures are used (e.g., Bonin et al. 2003; Ellis and Morrison 1998; Lotto et al. 2010; Liu et al. 2011; Snodgrass and Yuditsky 1996), although objective age of acquisition does appear to be a slightly stronger determinant (e.g., Chalard et al. 2003). Finally, both subjective and objective age of acquisition norms have been published for many languages, but not Thai. We included age of acquisition for this reason and also because of its fundamental interest to cognitive psychology and because age of acquisition data are currently lacking for the BOSS.

The Present Normative Study

Language-specific normative data are currently lacking for Thai. In order to facilitate investigations into language processing with both healthy and impaired participants, Thai language norms are urgently required. The aim of the present research was to standardize the 480-photograph BOSS (Brodeur et al. 2010) for Thai-language speakers on the variables reviewed above. The overarching purpose in creating a Thai normative database is to promote further research in the domains of psycholinguistics, vision, and neuropsychology. The current work represents the first large-scale, formal standardization of pictorial stimuli for the Thai language.

Method

Participants

A total of 584 undergraduate students from two well-known universities in Bangkok, Thailand, took part. From Thammasat University, 460 first year students (188 males and 272 females; mean age, 18.9 years; range 17–22 years; mean years of full-time education, 13.05 years) recruited from a foundation English course either volunteered or received course credit for their participation in the study. From Chulalongkorn University, 124 students (47 males and 77 females; mean age, 20.04 years; range 17–27 years; mean years of full-time education, 13.85 years) were recruited from an introductory psychology course and paid for their participation. All participants were native Thai speakers and had normal or corrected-to-normal vision. Ethical protocols approved by these institutions were followed.

Materials

We used the 480-item BOSS (Brodeur et al. 2010) as well as it associated names in Thai (described in more detail later). The BOSS contains color photographs of common objects which have been standardized for presentation at a pixel resolution of 2000 × 2000 (600 dpi). Each BOSS image was imported into Microsoft PowerPoint and displayed on a white background, with one picture per slide. To reduce task demands, the 480-item set was divided into either two or four subsets depending on the nature of the task. Subset allocations followed Snodgrass and Yuditsky (1996), where the alphabetized list of 480 items was numbered and then each item assigned to one of two or four sets according whether it was odd- or even-numbered. Twenty additional items originally excluded from the BOSS were selected to use in the practice phase.

Procedure

Testing sessions were conducted in large classrooms with groups of between 10 and 30 participants seated in rows in a slightly darkened room. Sample sizes varied by task as a result of the subset allocations described above. In total, there were 60 participants in the name agreement task, 58 in the object familiarity task, 70 in the visual complexity task, 124 in the manipulability task, 120 in the category agreement task, 120 in the image agreement task, and 136 in the age-of-acquisition task. However, a total of 104 participants completed both the image agreement and age-of-acquisition tasks, albeit on separate occasions about a month apart.

The general task procedure (unless otherwise stated) closely followed the procedures previously employed by Snodgrass and Vanderwart (1980) and Alario and Ferrand (1999). First, participants read a set of instructions in Thai for their task. A paraphrased version was then given orally by a research assistant and any questions were answered before testing commenced. Instructions encouraged participants to respond carefully and consistently to each item and to utilize the full range of the scale when rating objects. For the picture rating variables (name agreement, object familiarity, visual complexity, manipulability and category agreement), pictures were displayed sequentially via an overhead projector on a large, white screen situated at the front of a classroom. Presentation was controlled using an IBM-compatible computer running Microsoft PowerPoint. The trial number was displayed in the top right corner of each slide to help the participants align their responses with the test booklet. Item presentation order was fully randomized for each variable. After task completion, participants were asked to provide demographic information in a questionnaire, and then were debriefed and thanked for their participation. Procedures specific to individual variables are discussed further below.

Name Agreement

Participants were asked to generate object names by writing down the name of the first Thai word that came to mind. Objects were presented for six seconds per slide with a two second inter-trial interval. In the case of naming failures, participants were given three different response options: they could tick a box labelled DKO (don't know object) if they considered the picture was of an unknown object; DKN (don't know name) if the object was known but there was difficulty in retrieving the object name; and TOT (tip-of-the-tongue) if the name of the object was known but retrieving it at the time was difficult. These acronyms were defined in Thai.

Object Familiarity

Individual slides were presented at a rate of five seconds per slide, with an inter-trial interval of two seconds during which the screen went blank. Participants judged the familiarity of each concept displayed in the picture based on how often they came into contact with or thought about the concept that was depicted. Instructions emphasized that participants should rate the concept itself rather than the way it was presented in the picture. Ratings were made on a 5-point Likert scale (1 = *very unfamiliar*, 5 = *very familiar*).

Visual Complexity

Individual item timings were identical to those in the object familiarity task. Participants were required to rate the complexity of each image on a 5-point scale, with 1 indicating a *very simple* image and 5 a *very complex* image. Complexity in this task was defined as the amount of detail or intricacy of lines that are used to depict the object in each picture. Instructions emphasized that the complexity of the object in the photograph should be rated rather than the real-life object that it represented.

Manipulability

Two kinds of manipulability ratings were collected, with slide timings identical to those reported above. Rating procedures closely followed those previously employed by Guérard et al. (2015). For graspability, participants were asked to rate the ease with which they could grasp each object for the purpose of moving it, on a 5-point scale (1 = *very difficult to grasp for moving*, 5 = *very easy to grasp for moving*). For the mimability measure, participants were asked to rate each item according to the ease with which they could mime its use when alone. For clarity, we added that such miming was not for the purpose of communicating. This add-on instructional clarification was meant to ensure measurement of functional usage without confounding that with communicability (e.g., Magnié et al. 2003). Mimability was rated on a 5-point scale (1= *very difficult to mime*, 5 = *very easy to mime*).

Category Agreement

In this task, participants were instructed to determine category membership for each object by assigning a single number to it as it was presented. Here we mostly followed the procedure of Brodeur et al. (2010) by using their 19 pre-defined categories: building materials, clothing, decoration and gift accessories, electronic devices and accessories, food, furniture, games, toys and entertainment, hand labor tools and accessories, household articles and cleaners, jewels and money, kitchen utensils, medical instruments and accessories, musical instruments, natural elements and vegetation, outdoor activity and sport items, skin care and bathroom items, stationery and school supplies, weapons and items related to war, and *other*. These categories were translated into Thai by a professional translator, verified for accuracy, numbered, and then finally displayed in alphabetical order in the response booklet. To enable participants to familiarize themselves with the task itself as well as the large number of categories, descending stimulus exposure durations were used during the practice phase. The first set of five practice trials was presented for 20 s per image, followed by five trials at 18 s, then five at 15 s, five at 13 s, five at 12 s, and five at 10 s. After completion of 30 practice trials, the test phase began in which slide duration was kept constant at 10 s throughout. A warning tone was played during the inter-trial interval to alert participants to the onset of the next trial.

Image Agreement

E-Prime (Version 2.0) was used to present stimuli in this task. Participants made judgments on the object names generated in the naming task (i.e., the modal names; see name agreement task procedure above). On each trial, participants were presented with a modal word in 40-point Calibri typeface, followed by its corresponding picture. Task instructions emphasized that participants should form a mental image when they saw each word and then rate how closely their mental image matched the picture that followed on a 5-point scale. A rating of 1 indicated *low agreement* (i.e., the picture provided a poor match to the generated image), and a rating of 5 indicated *high agreement*. If they could not form an image or thought of a different image to the one pictured, they were instructed to respond NI (no image) or DO (different object). Each trial began with a fixation cross (500 ms), blank screen (500 ms), and a modal word (2500 ms). The screen then went blank for 3500 ms and participants were instructed to focus on the screen (or close their eyes) while forming a mental image of the object corresponding to the name they had seen. After 3500 ms the blank screen was replaced by its associated picture, which remained on screen for 3000 ms. During the 2000 ms inter-trial interval, a warning tone was played for 750 ms prior to the onset of the next trial.

Age of Acquisition

In this task, participants had to estimate the age at which they thought they had first learned each of the modal names (see above), whether in written or oral form (Carroll and White 1973). Four unique booklets were prepared, each containing a subset of 120 names presented in random order. One booklet was completed by a group of participants in each experimental session. Following Gilhooly and Logie (1980), age-of-acquisition ratings were made using a 7-point scale, with each point representing an age range of two years (1 = *0–2 years old*; 2 = *3–4 years old*; 3 = *5–6 years old*; 4 = *7–8 years old*; 5 = *9–10 years old*; 6 = *11–12 years old*; 7 = *13 years old or older*). These age bands were presented at the top of each page of the response booklet, and participants answered by filling in a square corresponding to each age band.

Results and Discussion

For object familiarity, visual complexity, manipulability, image agreement, and age of acquisition, scores were computed by averaging responses across items. For name agreement, we followed guidelines based on previous research (e.g., Brodeur et al. 2010; Snodgrass and Vanderwart 1980) by compiling written name responses for each object and correcting any obvious misspellings in order to regularize the raw naming responses. Second, in cases where two or more names were used for a particular object, only the first name was taken as a valid response. Third, all classifiers, quantifiers, and any other descriptors that brought no additional meaning to the response were removed (e.g., in the context of *pot to boil hot water* a transliteration of the Thai phrase for coffee pot, the word *hot* was a redundant descriptor given for boiled water, so was removed). Responses with added elaborations (descriptions which add specificity, precision, or further meaning; e.g., box in *Tupperware box*) were treated as different responses. Commonly acceptable abbreviations (e.g., *TV* and *television*) were treated as different responses because they could develop subtle semantic variances from parent words over time (see Snodgrass and Vanderwart 1980). Word order differences (e.g., *horse doll* and *doll horse*) were only regularized if their meanings were judged as identical by a Thai semanticist. Finally, some pairs of responses seemed highly similar in meaning and these were presented to two different native Thai speakers independently. The words judged equivalent in meaning by both judges were further discussed with the same Thai semanticist to determine how they should be coded (e.g., *vegetable broccoli* vs. *broccoli*; *almond nut* vs. *almond*). As a result of these discussions, some semantically identical responses were regularized. These particular regularizations (those that required Thai native speaker judgments:

alternate word order pairs and highly similar word pairs) resulted in the modification of 68 out of 7197 responses classified as modal (cumulatively across 480 items and 30 participants), so less than .01% total adjustment was made to modal word counts.

Two indices of name agreement were computed. The first measure, percentage name agreement, was calculated as the percentage of modal names (the most frequently occurring name for each object) as a proportion of all name responses for each object, excluding naming failures (DKO, DKN & TOT) and non-responses (i.e., no answer). Higher scores indicate greater name agreement. In cases where two names held identical percentages, the more precise name was chosen. Although our computation deviated slightly from previous norming studies (i.e., those that included naming failures in the denominator), here we followed Brodeur et al. (2010) for greater comparability. The second measure, known as the information statistic H, captures more information about the distribution of names across participants than about modal naming frequency (see Snodgrass and Vanderwart 1980). The H statistic was computed for each object using the standard formula:

$$H = \sum_{i=1}^{k} P_i \log_2(1/P_i),$$

where k refers to the number of different names given to each object, and P_i is the proportion of participants giving each name. As with the percentage name agreement measure described above, naming failures and non-responses were excluded in computing H values. An object with an H value of 0 indicates that the picture elicited the same naming response from all participants who provided a name for it. Increasing H values indicate decreasing levels of name agreement because of an increase in the frequency of alternative responses and their relative weightings. For example, an H value of 1 indicates that two names were used equally often to name that particular object. For the category agreement task, the modal category, percentage category agreement, and corresponding H value (H_{cat}) were computed using the same procedures described above. Naming failure rates (DKO, DKN, and TOT) were calculated as total number of naming errors divided by total number of responses possible. There were 14,400 responses possible, given 30 participants were presented 480 pictures each.[2]

Description of the Normative Data

A full summary of the rating data from the Thai-speaking participants is provided online as supplementary material (Clarke and Ludington 2017). For easy reference and comparison purposes, items are ordered alphabetically by their English modal name and item number as per the BOSS database (Brodeur et al. 2010). The following additional information is provided for each photograph: (1) the most frequent name given in Thai as well as the intended Thai name (where different) and its transliteration; (2) two measures of name and category agreement: the percentage of participants producing the most common name or assigning the most common category, and the corresponding H statistics; (3) the means and standard deviations for object familiarity, visual complexity, manipulability, image agreement, and age of acquisition; (4) the percentage of DKO, DKN, and TOT responses as well as the number of alternative names for each item are also provided. Table 1 reports the summary statistics for all variables in the database.[3] The 25th (Q1) and 75th (Q3) percentiles are provided to aid researchers who wish to select or exclude items by quartiles.

[2] There were only 14,368 actual responses because some participants did not respond on some items.

[3] The object *peeler* had a DKO score of 87% and not one single participant was able to name it. This object was therefore excluded from all analyses involving the name agreement data.

Table 1 Summary statistics for the main study variables

	NA%	*H*	Ofam	VC	Grasp	Mime	CA%	H_{cat}	IA	AoA
M	59.45	1.65	3.33	2.59	4.00	3.96	68.11	1.18	4.00	4.00
SD	24.89	0.95	0.90	0.77	0.58	0.69	21.26	0.72	0.63	0.98
Median	57.89	1.68	3.27	2.51	4.16	4.13	66.67	1.16	4.10	3.96
Range	92.00	3.87	3.71	3.71	3.03	3.30	80.00	3.25	3.00	4.82
Min	8.00	0	1.29	1.06	1.84	1.70	20.00	0	2.00	1.88
Max	100.00	3.87	5.00	4.77	4.88	5.00	100	3.25	5.00	6.71
Q1	38.10	0.89	2.71	2.01	3.75	3.53	50.00	0.57	3.53	3.26
Q3	82.76	2.43	4.06	3.14	4.44	4.53	90.00	1.78	4.52	4.71
IQR	44.66	1.54	1.35	1.13	0.69	1.00	40.00	1.21	0.98	1.44
Skew[a]	1.26	0.95	1.42	1.23	0.69	0.67	1.40	1.05	0.73	1.09

NA% = percentage name agreement; *H* = name agreement; Ofam = object familiarity; VC = visual complexity; Grasp = graspability; Mime = mimability; CA% = percentage category agreement; *H*cat = category agreement; IA = image agreement; AoA = age of acquisition; Q1 = 25th percentile; Q3 = 75th percentile; IQR = interquartile range

[a] Skew = (Q3 - Median)/(Median - Q1); > 1 is positively skewed

The mean percentage name agreement for the whole photographic dataset was 59.45%, with a corresponding *H* value of 1.65. These values suggest less overall consensus on the object names (or greater sources of name disagreement) compared to name agreement values routinely observed with the Snodgrass and Vanderwart (1980) stimuli. With line drawings, name agreement typically exceeds 80% and corresponding *H* values tend to be within the range of .2 to .8 (e.g., Alario and Ferrand 1999; Bates et al. 2003; Bonin et al. 2003; Dimitropoulou et al. 2009; Pompéia et al. 2001; Raman et al. 2014; Sanfeliù and Fernandez 1996; Snodgrass and Vanderwart 1980; Tsaparina et al. 2011). Notably, our name agreement measures are more compatible with those previously obtained with the BOSS (Brodeur et al. 2010). We discuss these similarities in greater detail later.

The norms for each category are displayed in Table 2. The category that was assigned the greatest number of objects was kitchen utensils (55), followed by food (47). Other categories containing relatively high numbers of objects were hand labor tools and accessories (42), and natural elements and vegetation (41). The number of objects assigned to some categories was rather low, although most categories assignments were in line with those previously observed by Brodeur et al. (2010), with the main exceptions of food (47 vs. 78 items, for ours and Brodeur et al.'s studies, respectively), natural elements and vegetation (41 vs. 11, respectively), and skin care and bathroom items (19 vs. 32, respectively).

Correlations Among the Measured Variables

We examined the relationships between the variables using correlational analyses. Table 3 shows the matrix of correlation coefficients for all objects with significant correlations marked with an asterisk. The strongest relationships were obtained between the two indices of name agreement (percentage and *H* statistic) and between the two indices of category agreement. The correlations between name agreement (*H*) and image agreement (−.435) and name agreement and object familiarity (−.314) were moderate and significant, suggesting that objects which elicit greater name agreement tend to be more familiar to participants and also evoke fewer different pictorial representations. These findings are generally consistent with

Table 2 Summary statistics for all variables for each category

Category	N	NA (%)	H	DKO (%)	DKN (%)	TOT (%)	Ofam	VC	Grasp	Mime	CA (%)	H_{cat}	IA	AoA
Building materials	3	70	1.53	3	1	1	3.3	1.6	2.6	2.7	81	0.96	4.4	4.4
Clothing	29	63	1.47	2	3	1	3.9	2.7	4.4	4.6	79	0.93	3.9	3.6
Decoration and gift accessories	24	61	1.58	8	10	3	3.0	2.8	4.1	3.8	59	1.49	3.8	3.8
Electronic devices and accessories	37	62	1.63	4	6	3	3.7	3.2	3.8	3.9	76	0.95	4.1	4.4
Food	47	61	1.56	10	8	4	3.4	2.3	4.1	4.0	71	0.84	4.1	3.9
Furniture	6	67	1.35	14	4	0	3.5	2.7	3.4	4.1	49	2.03	3.4	3.7
Games, toys, and entertainment	25	56	1.83	8	11	5	2.9	3.0	3.8	3.7	76	0.93	3.8	3.4
Hand labor tools and accessories	42	61	1.53	12	16	6	2.6	2.8	3.8	3.6	72	1.12	3.9	4.6
Household articles and cleaners	31	53	1.93	5	9	4	3.5	2.5	4.2	4.2	63	1.43	3.9	3.7
Jewels and money	7	61	1.49	12	6	3	3.2	2.9	4.2	3.8	53	1.63	3.7	3.6
Kitchen utensils	55	50	2.05	10	13	5	3.3	2.4	4.0	4.1	72	1.13	3.8	4.2
Medical instruments and accessories	14	58	1.62	7	10	5	3.1	2.8	3.8	3.9	66	1.31	4.2	4.4
Musical instruments	3	76	0.89	0	24	7	2.8	2.8	3.3	4.3	84	0.69	3.8	4.8
Natural elements and vegetation	41	65	1.39	10	10	5	3.3	2.5	3.9	3.5	61	1.12	4.1	3.9
Outdoor activity and sport items	14	64	1.39	2	5	2	3.2	2.9	4.0	4.2	77	0.95	4.2	4.5
Skin care and bathroom items	19	61	1.62	8	8	3	3.7	2.4	4.3	4.2	62	1.41	4.2	3.9
Stationery and school supplies	40	64	1.56	4	6	4	3.8	2.4	4.2	4.3	80	0.83	4.2	3.8
Weapons and items associated to war	0	–	–	–	–	–	–	–	–	–	–	–	–	–
Other	42	54	1.82	8	9	4	3.2	2.5	3.9	3.9	45	2.05	4.0	4.2

NA = percentage name agreement; H = name agreement; DKO = don't know object; DKN = don't know name; TOT = tip of the tongue; Ofam = object familiarity; VC = visual complexity; grasp = graspability; Mime = mimability; CA = percentage category agreement; H_{cat} = category agreement; IA = image agreement; AoA = age of acquisition

Table 3 Correlation matrix for all study variables

	H	NA%	Ofam	VC	Grasp	Mime	CA%	H_{cat}	IA
NA%	−.952**								
Ofam	−.314**	.353**							
VC	.032	−.022	−.208**						
Grasp	.038	− .035	.287**	−.312**					
Mime	−.195**	.253**	.621**	−.208**	.459**				
CA%	−.031	.078	.224**	.042	.059	.283**			
H_{cat}	.069	−.115*	−.283**	.008	−.110*	−.321*	−.933**		
IA	−.435**	.456**	.456**	−.188**	.096*	.349**	.141**	−.218**	
AoA	.198**	−.255**	−.370**	.095*	−.120**	−.202**	.034	−.020	.067

H = name agreement; NA% = percentage name agreement; Ofam = object familiarity; VC = visual complexity; Grasp = graspability; Mime = mimability; CA% = percentage category agreement; H_{cat} = category agreement; IA = image agreement; AoA = age of acquisition; Name agreement $N = 479$; all other cases $N = 480$.
* $p. < .05$, ** $p. < .01$

those observed in previous studies (Adlington et al. 2009; Alario and Ferrand 1999; Bonin et al. 2003; Brodeur et al. 2010; Liu et al. 2011; Moreno-Martínez and Montoro 2012; Raman et al. 2014; Snodgrass and Vanderwart 1980; Tsaparina et al. 2011).

Object familiarity correlated significantly with all other variables in the present study, although some of these correlations were fairly weak. Overall, the pattern of correlations for object familiarity generally mirrors the pattern observed in normative studies using both line drawings (Alario and Ferrand 1999; Dell'Acqua et al. 2000; Nishimoto et al. 2005; Raman et al. 2014; Sanfeliù and Fernandez 1996; Snodgrass and Vanderwart 1980; Weekes et al. 2007) and photographic stimuli (Adlington et al. 2009; Brodeur et al. 2010, 2014; Moreno-Martínez and Montoro 2012; Shao and Stiegert 2016). The moderate negative correlation between object familiarity and age of acquisition (−.370) indicates that words that are acquired at an earlier age are also more familiar (Álvarez and Cuetos 2007; Barry et al. 1997; Bonin et al. 2003; Liu et al. 2011). Positive correlations between object familiarity and manipulability are also in line with a number of previous studies (Brodeur et al. 2010, 2014; Magnié et al. 2003; Salmon et al. 2010). Notably, in our study, object familiarity correlated more strongly with one manipulability index (mimability), while visual complexity correlated more strongly with the other (graspability). This makes some intuitive sense: Knowledge about an object's function is retrieved from semantic memory, which may explain why miming was related to object familiarity. Grasping, however, requires volumetric analysis of object structure, which may explain its relationship to visual complexity (e.g., Guérard et al. 2015; Jax and Buxbaum 2010).

Comparison with Other Bank of Standardized Stimuli (BOSS) Norms

In order to demonstrate a measure of comparability across studies, we compared our item means with item means from the French and English norms previously obtained with the BOSS (Brodeur et al. 2010, 2012). Naturally, we expected some differences in average normative values across studies, but we also expected strong correlations to emerge from the pattern of item means across languages because we used the same stimuli and methodologies.

Table 4 Comparison of Thai norms with English and French norms for 416 shared items from the BOSS

Variables	Thai				English				French			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
NA (%)	62	24	10	100	65	23	20	100	63	24	14	100
H	1.58	0.95	0.00	3.87	1.61	1.00	0.00	4.03	1.53	0.97	0.00	3.83
DKO (%)	3	5	0	20	1	3	0	18	1	3	0	17
DKN (%)	8	10	0	70	5	7	0	47	3	6	0	37
TOT (%)	4	5	0	27	2	3	0	18	2	4	0	24
Ofam	3.51	0.80	1.43	5.00	4.07	0.39	3.00	4.80	4.13	0.51	2.60	5.00
VC	2.59	0.76	1.06	4.51	2.42	0.44	1.40	4.10	2.08	0.38	1.30	3.70
Mime	4.10	0.59	1.70	5.00	2.61	0.77	1.20	4.50	2.90	0.79	1.30	4.80
CA (%)	70	21	23	100	73	20	26	100	76	19	18	100
H_{cat}	1.10	0.70	0.00	2.66	1.15	0.67	0.00	2.70	1.03	0.66	0.00	2.84
IA	4.10	0.57	2.63	5.00	3.92	0.53	2.50	4.90	4.04	0.56	1.80	5.00
NMI (%)	1	4	0	47	1	3	0	21	2	4	0	35

NA = percentage name agreement; H = name agreement; DKO = don't know object; DKN = don't know name; TOT = tip-of-the-tongue; Ofam = object familiarity; VC = visual complexity; Mime = mimability; CA = percentage category agreement; H_{cat} = category agreement; IA = image agreement; NMI = no mental image. BOSS = Bank of Standardized Stimuli; English normative data from Brodeur et al. (2010); French normative data from Brodeur et al. (2012)

Initial observations revealed that naming failures were quite high in our sample, averaging 20.63% across the DKO, DKN and TOT categories, which is substantially higher than the values reported by Brodeur et al. (2010, 2012) for English and French speakers. The overall higher rate of naming failure among Thai participants probably reflects some degree of cultural unfamiliarity with the stimuli (the BOSS contains pictures of objects that are commonly found in North America). To enhance comparability of our data, we followed Brodeur et al. (2010) and identified items that were not recognized (DKO) by greater than 20% of the participants. Brodeur et al. removed 58 items from their original 538-item dataset, resulting in a final database of 480 items. In the present study, 64 items (13.33%) had DKO rates of greater than 20%. Removing these from our dataset resulted in a subset of 416 shared items.

Summary statistics for the overlapping variables across studies are displayed in Table 4. For this subset of shared items, it is evident that the Thai normative values closely resemble the English and French norms, especially on the three agreement measures (i.e., name, category and image agreement). Nevertheless, some cross-linguistic differences remained. Object familiarity remained lower in comparison to ratings from the French and English samples, suggesting that some of the objects in the database are less culturally familiar to the Thai sample. The differences in mimability ratings probably resulted from differing definitions between our study and Brodeur et al.'s (2010, 2012). We simplified our tool so participants would only rate the ease of miming rather than miming to communicate, which was our aim at the outset. Additional research is needed to determine whether these task-simplified ratings prove useful.

Finally, we computed Pearson correlations for the 416 item means using the original data from Brodeur et al. (2010, 2012) in order to compare Thai, English, and French norms for the overlapping variables (see Table 5). As expected, all correlations were positive, moderate to strong, and highly significant (all $ps < .001$). Despite lower ratings of object familiarity

Table 5 Cross-linguistic correlations between Thai, English and French norms for the 416 shared items from the BOSS

Variables	Th–En	Th–Fr	En–Fr
Name agreement	.456**	.406**	.618**
H	.549**	.498**	.693**
Object familiarity	.673**	.598**	.795**
Visual complexity	.843**	.811**	.747**
Mimability[a]	.352**	.405**	.876**
Category agreement	.654**	.459**	.773**
H_{cat}	.699**	.527**	.719**
Image agreement[b]	.487**	.425**	.602**

English normative data from Brodeur et al. (2010). French normative data from Brodeur et al. (2012). *Th* Thai, *En* English, *Fr* French.

$^{**}p. < .01$

[a] Brodeur et al. (2010, 2012) measured ease of miming an object such that an observer could identify the object being mimed rather than the ease of actually miming the object. We feel that these two measures are conceptually similar enough to warrant their comparison

[b] Although traditional instructions were used for the image agreement task, Brodeur et al. (2010) used a slightly modified version that required their participants to ignore the object's position when making their image agreement ratings.

in the present sample (3.51 vs. 4.07 vs. 4.13 for Thai, English, and French, respectively), the reasonably strong cross-linguistic correlations observed between Thai–English (.67), and between Thai–French (.59) suggest similar *patterns* of object familiarity across studies. However, it is important to note that our cross-linguistic correlations are a little weaker in magnitude compared to that observed for English and French speakers (.79) in Brodeur et al. (2012; Table 5). Such a difference is not unexpected. Brodeur et al.'s familiarity ratings were obtained from English and French speakers who share the same general environment and are likely to have had more frequent exposure to the actual objects depicted in the BOSS, at least compared to the Thai sample (a similar point was raised by Brodeur et al. 2012). Brodeur et al. (2012) also point out that all of their mean item ratings for familiarity were greater than 3 for both English and French. In the present study, a total of 297 out of 416 images (71%) received mean item ratings of 3 or above suggesting a greater range of unfamiliarity with these stimuli. We believe that this is a strength rather than a weakness of the current database as it means that future researchers can select items from a greater range of familiarity values.

For visual complexity, we expected strong correlations across studies given that ratings for this variable are thought to be largely language independent. Our ratings of .84 for Thai–English and .81 for Thai–French support these expectations. Robust cross-study correlations have also been demonstrated for this visual complexity in other studies norming color pictorial datasets: for Greek, French, Spanish and American English (Dimitropoulou et al. 2009), for Turkish, Russian, Greek and French (Raman et al. 2014), and for French and English (Brodeur et al. 2012; also see Table 5). Taken together, strong cross-study correlations for visual complexity, including those observed in the current study for Thai, English, and French, suggest that visual complexity is largely culturally independent.

On the other hand, the correlation between Thai name agreement and name agreement in English and French was marginally weaker compared to that observed between English and French (Brodeur et al. 2012). This suggests that, for some objects at least, high levels of name agreement in Thai do not necessarily correspond to objects that have high levels of name agreement in English or French. Thus despite similar name agreement percentages and *H* values across studies, there are likely some linguistic differences in how objects are named. Comparisons of name agreement for the Snodgrass and Vanderwart pictures reveal similar patterns. Boukadi et al. (2016), for example, compared name agreement scores for Tunisian Arabic with overlapping pictures from French, American, and Spanish samples revealing correlations of .32, .40, and .15, respectively. Similarly, Alario and Ferrand (1999) obtained cross-linguistic correlations of .42 for French and American English and .50 for French and Spanish using identical pictures (see also, Dell'Acqua et al. 2000; Dimitropoulou et al. 2009; Nishimoto et al. 2005; Sanfeliù and Fernandez 1996).

The weakest correlation observed in the current study concerns mimability, which at .35 for Thai–English and .40 for Thai–French is substantially lower than that obtained by Brodeur et al. (2012) for English–French (.88). Recall that Brodeur et al.'s (2012) definition of manipulability concerned how easily participants could pantomime an object's use unambiguously to an observer, while in the present study we removed this potentially difficult perspective-taking aspect from the instructions. As such participants rated only how easy it was to mime an object's use rather than miming usage for the purposes of communicating. We believe that these definitional differences, rather than intergroup influences of mother tongue or culture, is at the root of the low mimability association between the current results and those of Brodeur et al. (2010, 2012). Further empirical work is clearly needed in this domain.

Conclusions

It is clear that Snodgrass and Vanderwart's (1980) prescient work has had a massive impact on the research community, enabling scientists to conduct well-controlled studies isolating various cognitive factors without interference from a host of otherwise confounding psycholinguistic and structural variables. Given the mushrooming of normative studies over the past couple of decades, their need and usefulness for research in psycholinguistics and the cognitive sciences has grown strongly. It is also clear that advances in information technology have prompted a shift towards the use of more ecological, photographic stimuli.

Here, we endeavored to extend the existing BOSS norms to a different cultural and linguistic context, Thailand, a fast-developing economy ripe for future research in these areas. Based on our survey of the literature, we selected eight dimensions which we carefully considered as some of the most important variables to cognitive and psycholinguistics research. We largely followed Brodeur et al. (2010, 2012), but divided their manipulability variable into two variables, graspability (Salmon et al. 2010) and mimability (Magnié et al. 2003), to capture structural and functional dimensions, respectively. Rated age of acquisition was also measured in the current study and these appear to be the first such norms available for Thai, as well as for the BOSS. Objective measures, currently lacking for both, would be a welcomed addition in future research.

Our results show that the Thai norms compared favorably with norms found in other populations using the BOSS (Brodeur et al. 2010, 2012). In addition, our cross-linguistic correlational analyses further supported the Thai normative database in terms of validating its measurement methodologies and constructs. We noted where numerical differences arose across studies, and interpreted these differences in terms of linguistic and cultural differences. The current norms add to the value of the BOSS as a resource by providing researchers with opportunities to match and compare existing and future norms between languages and cultures. We anticipate that our work should encourage researchers operating within different linguistic communities to take advantage of the rich and complex stimuli provided by the BOSS, enabling new investigations into behavioral and brain processes governing language and cognition in healthy controls and in other populations such as children and those with brain injury.

Acknowledgements This research was supported by a grant from the Language Institute, Thammasat University, B.E. 2557, awarded to the first author. We would like to thank Mathieu Brodeur for providing access to the Bank of Standardized Stimuli and Upsorn Tawilapakul for valuable advice concerning Thai linguistics. We also extend our gratitude to Jutarat Anansalung, Suthasinee Jaismith, Jirayu Loetcharoenwanich, Patitta Malathong, Natcha Meemanan, Morakot Radomkit, Luxsamee Tanviboolaya, and Paisarn Yamwong for their research assistance, and to all of the student volunteers of Thammasat University and Chulalongkorn University for their participation.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

Adlington, R. L., Laws, K. R., & Gale, T. M. (2009). The Hatfield Image Test (HIT): A new picture test and norms for experimental and clinical use. *Journal of Clinical & Experimental Neuropsychology*, *31*(6), 731–753. https://doi.org/10.1080/13803390802488103.

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 531–552.

Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155.

Álvarez, B., & Cuetos, F. (2007). Objective age of acquisition norms for a set of 328 words in Spanish. *Behavior Research Methods*, *39*(3), 377–383.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *The Quarterly Journal of Experimental Psychology, 50A*(3), 560–585. https://doi.org/10.1080/027249897392026.

Belke, E., Brysbaert, M., Meyer, A.S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition*, *96*(2), B45–B54. https://doi.org/10.1016/j.cognition.2004.11.006.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 158–167.

Boukadi, M., Zouaidi, C., & Wilson, M. A. (2016). Norms for name agreement, familiarity, subjective frequency, and imageability for 348 object names in Tunisian Arabic. *Behavior Research Methods*, *48*(2), 585–599. https://doi.org/10.3758/s13428-015-0602-3.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE*, *5*(5). https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 new normative photos. *PloS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0106953

Brodeur, M. B., Kehayia, E., Dion-Lessard, G., Chauret, M., Montreuil, T., Dionne-Dostie, E., et al. (2012). The bank of standardized stimuli (BOSS): Comparison between French and English norms. *Behavior Research Methods*, *44*(4), 961–970. https://doi.org/10.3758/s13428-011-0184-7.

Capitani, E., Laiacona, M., Mahon, G., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, *20*, 213–261. https://doi.org/10.1080/02643290244000266.

Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning & Verbal Behavior*, *12*(5), 563–576. https://doi.org/10.1016/S0022-5371(73)80036-2.

Cattell, J. (1886). The time to see and name objects. *Mind*, *11*(41), 63–65.

Chalard, M., Bonin, P., Méot, A., Boyer, B., & Fayol, M. (2003). Objective age-of-acquisition (AoA) norms for a set of 230 object names in French: Relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *European Journal of Cognitive Psychology*, *15*(2), 209–245. https://doi.org/10.1080/09541440244000076

Clarke, A. J. B., & Ludington, J. D. (2017). *A set of 480 colour photographic pictures standardised for Thai: A new resource for cognitive and psycholinguistic research*. [Data Collection]. Colchester, Essex: UK Data Archive. 10.5255/UKDA-SN-852500.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2342 words. *The Quarterly Journal of Experimental Psychology*, *60*(8), 1072–1082. https://doi.org/10.1080/17470210701315467.

Cuetos, F., Ellis, A. W., & Álvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 650–658.

Damian, M. F., & Martin, R. C. (1998). Is visual lexical access based on phonological codes? Evidence from a picture-word interference task. *Psychonomic Bulletin & Review*, *5*(1), 91–95. https://doi.org/10.3758/BF03209461.

Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, *32*(4), 588–615.

Dimitropoulou, M., Duñabeitia, J. A., Blitsas, P., & Carreiras, M. (2009). A standardized set of 260 pictures for Modern Greek: Norms for name agreement, age of acquisition, and visual complexity. *Behavior Research Methods*, *41*(2), 584–589.

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*(2), 515–523. https://doi.org/10.1037/0278-7393.24.2.515.

E-Prime (Version 2.0) [Computer software]. Sharpsburg, PA: Psychology Software Tools, Inc.

Filliter, J. H., McMullen, P. A., & Westwood, D. (2005). Manipulability and living/non-living category effects on object identification. *Brain & Cognition*, *57*(1), 61–65. https://doi.org/10.1016/j.bandc.2004.08.022.

Gerlach, C. (2001). Structural similarity causes different category-effects depending on task characteristics. *Neuropsychologia*, *39*(9), 895–900.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281.

Gibson, J. J. (2014). *The ecological approach to visual perception: Classic edition*, 1st edn. Hove: Psychology Press.

Gilhooly, K. J., & Gilhooly, M. L. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition*, *7*(3), 214–223. https://doi.org/10.3758/BF03197541.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427.

Grigoriev, A., & Oshhepkov, I. (2013). Objective age of acquisition norms for a set of 286 words in Russian: Relationships with other psycholinguistic variables. *Behavior Research Methods*, *45*(4), 1208–1217. https://doi.org/10.3758/s13428-013-0319-0.

Guérard, K., Lagacé, S., & Brodeur, M. B. (2015). Four types of manipulability ratings and naming latencies for a set of 560 photographs of objects. *Behavior Research Methods*, *47*(2), 443–470. https://doi.org/10.3758/s13428-014-0488-5.

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language & Cognitive Processes*, *26*(7), 902–934. https://doi.org/10.1080/01690965.2010.509946.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307.

Hirsh, K. W., & Funnell, E. (1995). Those old, familiar things: Age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of Neurolinguistics*, *9*(1), 23–32. https://doi.org/10.1016/0911-6044(95)00003-8.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*(1), 67–104. https://doi.org/10.1080/02643298808252927.

Huth, A. G., Nishimoto, S., Vo, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224.

Jax, S. A., & Buxbaum, L. J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition*, *115*(2), 350–355. https://doi.org/10.1016/j.cognition.2010.01.004.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition*, *13*(7/8), 789–845. https://doi.org/10.1080/13506280544000066.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*(5), 684–712. https://doi.org/10.1037/0033-2909.131.5.684.

Kremin, H., Akhutina, T., Basso, A., Davidoff, J., De Wilde, M., Kitzing, P., et al. (2003). A cross-linguistic data bank for oral picture naming in Dutch, English, German, French, Italian, Russian, Spanish, and Swedish (PEDOI). *Brain & Cognition*, *53*(2), 243–246. https://doi.org/10.1016/S0278-2626(03)00119-2

Laws, K. R., & Neve, C. (1999). A normal category-specific advantage for naming living things. *Neuropsychologia*, *37*(11), 1263–1269.

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0016505

Lotto, L., Surian, L., & Job, R. (2010). Objective age of acquisition for 223 Italian words: Norms and effects on picture naming speed. *Behavior Research Methods*, *42*(1), 126–133. https://doi.org/10.3758/BRM.42.1.126.

Magnié, M. N., Besson, M., Poncet, M., & Dolisi, C. (2003). The Snodgrass and Vanderwart set revisited: Norms for object manipulability and for pictorial ambiguity of objects, chimeric objects, and nonobjects. *Journal of Clinical & Experimental Neuropsychology*, *25*(4), 521–560. https://doi.org/10.1076/jcen.25.4.521.13873.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1–3), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004.

Martein, R. (1995). Norms for name and concept agreement, familiarity, visual complexity and image agreement on a set of 216 pictures. *Psychologica Belgica*, *35*(4), 205–225.

Moreno-Martínez, F.J., & Montoro, P.R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0037527

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology*, *50A*(3), 528–559.

Morrison, C. M., & Ellis, A. W. (1995). The role of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*(1), 116–133.

Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods*, *37*(3), 398–416.

Ostergaard, A. L., & Davidoff, J. B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*(3), 579–587.

Pind, J., Jónsdóttir, H., Gissurardóttir, H., & Jónsson, F. (2000). Icelandic norms for the Snodgrass and Vanderwart (1980) pictures: Name and image agreement, familiarity, and age of acquisition. *Scandinavian Journal of Psychology*, *41*(1), 41–48.

Pompéia, S., Miranda, M.C., & Bueno, O.F.A. (2001). A set of 400 pictures standardised for Portuguese: Norms for name agreement, familiarity and visual complexity for children and adults. *Arquivos de Neuro-Psiquiatria*, *59*(2-B), 330–337. https://doi.org/10.1590/S0004-282X2001000300004

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *The Quarterly Journal of Experimental Psychology*, *41A*(4), 797–828.

Raman, I., Raman, E., & Mertan, B. (2014). A standardized set of 260 pictures for Turkish: Norms of name and image agreement, age of acquisition, visual complexity, and conceptual familiarity. *Behavior Research Methods*, *46*(2), 588–595. https://doi.org/10.3758/s13428-013-0376-4.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, *46*(1), 247–253.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236.

Salmon, J. P., McMullen, P. A., & Filliter, J. H. (2010). Norms for two types of manipulability (graspability and functional usage), familiarity, and age of acquisition for 320 photographs of objects. *Behavior Research Methods*, *42*(1), 82–95. https://doi.org/10.3758/BRM.42.1.82.

Sanfeliù, M. C., & Fernandez, A. (1996). A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behavior Research Methods, Instruments, & Computers*, *28*(4), 537–555.

Shao, Z., & Stiegert, J. (2016). Predictors of photo naming: Dutch norms for 327 photos. *Behavior Research Methods*, *48*(2), 577–584. https://doi.org/10.3758/s13428-015-0613-0.

Silveri, M. C., Cappa, A., Mariotti, P., & Puopolo, M. (2002). Naming in patients with Alzheimer's disease: Influence of age of acquisition and categorical effects. *Journal of Clinical & Experimental Neuropsychology*, *24*(6), 755–764.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*(2), 174–215.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, *28*(4), 516–536. https://doi.org/10.3758/BF03200540.

Tanaka, J., Weiskopf, D., & Williams, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, *5*(5), 211–215.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Tsaparina, D., Bonin, P., & Méot, A. (2011). Russian norms for name agreement, image agreement for the colorized version of the Snodgrass and Vanderwart pictures and age of acquisition, conceptual familiarity, and imageability scores for modal object names. *Behavior Research Methods*, *43*(4), https://doi.org/10.3758/s13428-011-0121-9.

Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory & Cognition*, *26*(6), 1282–1291.

Viggiano, M. P., Vannucci, M., & Righi, S. (2004). A new standardized set of ecological pictures for experimental and clinical research on visual object processing. *Cortex*, *40*(3), 491–509.

Weekes, B. S., Shu, H., Hao, M., Liu, Y., & Tan, L. H. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods*, *39*(2), 335–342.

Yoon, C., Feinberg, F., Luo, T., Hedden, T., Gutchess, A. H., Chen, H. Y. M., et al. (2004). A cross-culturally standardized set of pictures for younger and older adults: American and Chinese norms for name agreement, concept agreement, and familiarity. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 639–649.

Zannino, G. D., Perri, R., Salamone, G., Di Lorenzo, C., Caltagirone, C., & Carlesimo, G. A. (2010). Manipulating color and other visual information influences picture naming at different levels of processing: Evidence from Alzheimer subjects and normal controls. *Neuropsychologia*, *48*(9), 2571–2578. https://doi.org/10.1016/j.neuropsychologia.2010.05.003.