



Bank of Standardized Stimuli (BOSS): Dutch Names for 1400 Photographs

RESEARCH ARTICLE

C. DECUYPER 

M. BRYSSBAERT 

M. B. BRODEUR

A. S. MEYER

**Author affiliations can be found in the back matter of this article*

 ubiquity press

ABSTRACT

We present written naming norms from 153 young adult Dutch speakers for 1397 photographs (the BOSS set; see Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Brodeur, Guérard, & Bouras, 2014). From the norming study, we report the preferred (modal) name, alternative names, name agreement, and average object agreement. In addition, the data base includes Zipf frequency, word prevalence and Age of Acquisition for the modal picture names collected. Furthermore, we describe a subset of 359 photographs with very good name agreement and a subset of 35 photos with two common names. These sets may be particularly valuable for designing experiments. Though the participants typed the object names, comparisons with other datasets indicate that the collected norms are valuable for spoken naming studies as well.

CORRESPONDING AUTHOR:

C. Decuyper

Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands

Caitlin.Decuyper@mpi.nl

KEYWORDS:

picture naming; Dutch; name agreement; norms

TO CITE THIS ARTICLE:

Decuyper, C., Brysbaert, M., Brodeur, M. B., & Meyer, A. S. (2021). Bank of Standardized Stimuli (BOSS): Dutch Names for 1400 Photographs. *Journal of Cognition*, 4(1): 33, pp. 1–11. DOI: <https://doi.org/10.5334/joc.180>

An important challenge for psycholinguistic and neurobiological studies of speaking is how to elicit specific utterances, for instance the noun “apple”. A commonly used task is picture naming, which is seen as well-suited for studying the core processes of speech planning, including the selection of a concept, the retrieval of the corresponding word representations, and articulatory planning processes (e.g., Levelt, Roelofs, & Meyer, 1999). Numerous studies have been dedicated to unravelling the processes involved in picture naming and their neurological bases (Miozzo, Pulvermüller, & Hauk, 2015; Piai, Roelofs, Rommers, & Maris, 2015).

For any naming experiment, it is crucial to select appropriate stimuli. Items can be selected on the basis of experiment-specific pretests or/and by reference to published norms. In general, it is good to have several item sets to select from, as this allows researchers to measure processes that generalize beyond the specific stimulus set used, a requirement known in test psychology as the multimethod approach (Campbell & Fiske, 1959; Eid & Diener, 2006). In the present paper, we describe the results of a photo norming study carried out with adult native speakers of Dutch. We are convinced that the norms will be of use for researchers working with Dutch speakers, because the available normed stimulus sets consist of drawings only and are smaller in size.

Participants saw a series of photographs and typed the names. Our norming variable of main interest was name agreement (NA), which is the proportion of speakers who use the same name to refer to an image. Previous studies have shown that name agreement is a strong predictor of naming latency. Images with high name agreement are named faster and more accurately than images with low name agreement (Alario et al., 2004; Barry, Morrison, & Ellis, 1997; Bonin, Chalard, Méot, & Fayol, 2002; Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Cheng, Schafer, & Akyürek, 2010; Cuetos, Ellis, & Alvarez, 1999; Dell’acqua, Lotto, & Job, 2000; Ellis & Morrison, 1998; Lachman, Shaffer, & Hennrikus, 1974; Paivio, Clark, Digidon, & Bons, 1989; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995).

A common interpretation of this finding is that images with low name agreement activate multiple competing conceptual and/or lexical representations, and the need to select among them slows down naming and makes it more error-prone compared to naming images with high name agreement (Alario et al., 2004; Levelt et al., 1991). Name agreement tends to affect naming speed and accuracy independently of word frequency and age of acquisition (Lachman et al., 1974; Vitkovitch & Tyrrell, 1995).

Controlling name agreement is relevant for the design of many studies. For instance, researchers may want to generate two or more sets of images differing on one feature and matched on other variables likely to affect naming performance. Name agreement should then be taken into account because of its known effect on naming speed and accuracy. For other studies researchers may want to select images that are highly likely to be named in the same way by all adult native speakers of a language. This can be classic chronometric or neurobiological studies of speech planning (e.g., Jongman, Roelofs, & Lewis, 2020), studies into naming disorders (e.g., Bose & Schafer, 2017), second language testing (Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012) or studies of verbal or visual memory (e.g., Zormpa, Brehm, Hoedemaker, & Meyer, 2019). For other studies, images may be needed that elicit two or more plausible names, such as ‘sofa’, ‘couch’ and ‘settee’ (see e.g., Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998). This holds, for instance, for studies of executive control of picture naming, where naming agreement can be varied to induce more or less competition between lexical items, which needs to be resolved through the recruitment of selective inhibition (e.g., Shao, Roelofs, Acheson, & Meyer, 2014). It also holds for studies of referential communication, where speakers may or may not converge on common names (e.g., Brown-Schmidt, 2009), and for studies interested in the origins of name agreement effects in their own right (Madden, Sale, & Robinson, 2019; Vitkovitch & Tyrrell, 1995). Finally, authors may be interested in pictures that have a low name agreement, because participants may be less likely to encode such pictures verbally (e.g., Nakabayashi, Burton, Brandimonte, & Lloyd-Jones, 2012).

To select suitable images, researchers can either run pilot and norming studies or make use of existing norms (for an extensive overview of image norming studies in different languages and samples, see Brodeur, Guérard, & Bouras, 2014). Many picture naming studies have used the norms prepared by Snodgrass and Vanderwart (1980). However, the normed picture set is not large (260 pictures) and consists of black-and-white drawings, some of which are now somewhat dated. Moreover, line drawings are not optimal for all research purposes. As Brodeur et al. (2010, 2014) pointed out, black-and-white line drawings are often harder to recognize than coloured images

or photographs as they miss cues related to colour and texture (see also Biederman & Ju, 1988; Brodie, Wallace, & Sharrat, 1991; Moreno-Martinez & Montoro, 2012; Ostergaard & Davidoff, 1985).

Brodeur and colleagues (2010, 2014) generated a large set of photographs, known as BOSS (Bank of Standardized Stimuli) and made them publicly available along with written naming norms collected for Canadian English. A subset of the photographs was also normed for Canadian French (Brodeur et al., 2012) and Thai (Clarke & Ludington, 2018). As in the study by Snodgrass and Vanderwart (1980), pictures were normed for *name*, *object familiarity* (concept rather than picture), *visual complexity* (in terms of the quantity of details and the intricacy of the lines), and *image agreement*. In addition, Brodeur and colleagues collected norms for *category* (participants could make a selection from 18 categories, such as clothing, food, furniture or “other”), *manipulability* (whether or not it is easy to mime the action associated with this object), *object agreement* (reflecting how similar the object is to the one imagine by the participant) and *viewpoint agreement* (whether or not the object is in the same position as the participant imagined the object to be in). Participants were also asked to classify objects as *living or non-living* things.

In the present paper, we describe a norming study of these photographs for speakers of Dutch. We only know of one similar study, with a much smaller set of photographs (327 items; Shao & Stiegert, 2016). We presented 1397 out of the 1468 photographs prepared and described by Brodeur and colleagues (2010, 2014) to young adult speakers of Dutch. As in the original norming studies, we asked the participants to type the names of the objects and, when they could not do so, to indicate whether they did not know the object (DKO) or did not know its name (DKN). These two categories are important as they provide clues to the reasons for naming difficulties. We did not include the category ‘tip-of-the-tongue’ (TOT) of the original study, as TOT states are unlikely to occur for the relatively high-frequent names we expected to be produced and difficult to interpret without follow-up questions. In addition to written names, we obtained ratings of object agreement by asking the participants to indicate on a 5-point scale how well each photograph represented the actual concept. To limit the duration of the study, we did not collect norms for manipulability, object viewpoint, object familiarity, and visual complexity and did not ask participants to categorize stimuli, as had been done in the studies by Brodeur and colleagues. These variables were of lesser interest to our work than NA and can, of course, be obtained from the original studies (or collected in future studies). We provide summary statistics for the items below, as well as correlational analyses, including split-half correlations as a measure of reliability, and we briefly discuss the morphological and semantic relations between competitors for items with multiple plausible names.

As the participants typed the object names, the norms are particularly suitable for studies of written language production (Bonin, Méot, Laroche, Bugajska, & Perret, 2019; Torrance et al., 2018). However, they can also assist researchers in the generation of materials for spoken language studies. For instance, if a photograph elicits only a single written name (e.g. “apple”) it would be highly surprising to see participants use many different names in oral naming. Likewise, if a photograph elicits many different names and/or many DKO and DKN responses, it would be surprising to see participants converge on a single name in oral naming.

METHOD

PARTICIPANTS

All 153 paid participants were recruited from the MPI participant database (mean age = 23 years; 30 males). The composition of the sample matched the composition of the participant pool. All participants were university students, native speakers of Dutch, and reported having normal sight. None of them was color-blind. Ethical approval to conduct the study was given by the Ethics Board of the Social Sciences Faculty of Radboud University.

MATERIALS AND DESIGN

The Canadian English version of BOSS consists of two sets of photographs (see Brodeur et al. 2010, 2014, for details on how the photographs were created). One set of 538 items was described in Brodeur et al. (2010) and a second set of 930 photographs in Brodeur et al. (2014). We tested most of these photographs, divided into three sets. Set 1 consisted of the 467 “useful” photographs in the set of 538 photographs tested by Brodeur et al. (2010). The remaining photographs were discarded as suboptimal by Brodeur et al. (2010) or O’Sullivan,

Lepage, Bouras, Montreuil, and Brodeur (2012). The set of items covers a broad range of object categories, including animals; body parts; building infrastructure; building materials; clothing; decoration and gift accessories; electronic devices and accessories; food; furniture; games, toys, and entertainment; hand labour 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; stationary and school supplies; vehicles; weapons and items related to war.

To create sets 2 and 3, the 930 items from Brodeur et al. (2014) were randomized and split into two sets of 465 items each. The photographs were sized to 400 by 400 pixels (10.58 cm on the screen), corresponding to approximately 10 degrees of visual angle at a distance of 60 cm. Each list was seen by 50 to 52 participants. Two of the 153 participants were accidentally invited twice to participate in the study and named two sets. For each participant, a new random sequence of items was created.

PROCEDURE

Participants were tested individually in a quiet room. They were asked to type the names of a set of photographs and rate on a 5-point scale how well each photograph depicted the object. Photographs were presented one by one on a computer screen, together with the scale. Participants typed their response underneath the photograph and pressed Enter to end the keyboard input. They used the digits 1 to 5 (top of the keyboard) to indicate how well the photograph represented the object (“Indicate on a scale from 1 to 5 how well the photograph depicts the object”, 1 = very poor representation; 5 = perfect representation). The selected number was highlighted on the scale.

When a participant could not name an object, they were asked to press Enter to skip the keyboard input and then press “a” or “b” to indicate that they did not recognize the object (button a, “DKO”) or that they recognized the object but did not know its name (button b, “DKN”). The instructions were available to the participants on screen on every trial. The manual response (pressing 1-5 or a-b) triggered the next trial with a delay of 500 ms. Participants could not go back and change their responses. The study was run on desktop computers and laptops making use of Presentation software (NeuroBehavioral Systems Inc., 2017).

DATA CODING AND NORMS

We first identified DKN and DKO responses and excluded them from further analyses. In some trials (141 responses, i.e. only 0.20% of the data), participants did not name the picture, but did press 1-5 instead of selecting “a” or “b”. These responses were manually coded as “c” and excluded as well (but not labelled as DKN or DKO). To prepare the name agreement norms, obvious spelling errors and typos were corrected to allow for aggregation of responses. Where multiple spellings were possible (e.g. “giraf” and “giraffe”), the more common spelling was selected. Furthermore, we aggregated across all responses pertaining to the same lemma, i.e. across singular and plural forms (e.g. “amandel” and “amandelen”), and simple forms and diminutives (“hoed” and “hoedje”). Object names including adjectives counted as separate responses, as long as they defined specific types of objects (e.g. “plastic beker” (plastic cup). Other adjectives (as in “blauwe beker” (blue cup)) were ignored.

We computed Modal Name Agreement (NA) as the percentage of participants who gave the most common name. We also computed the H-value for name agreement, defined as:

$$H = \sum_{i=1}^k P_i \log_2 \left(\frac{1}{P_i} \right)$$

The H-value is sensitive to the number of different names (k) that were given to an object and the proportion of participants that used each of these names (P_i) after excluding DKN and DKO responses. This statistic was also used by Snodgrass and Vanderwart (1980) and Brodeur et al. (2010; 2014). An object with a unique name (naming agreement = 100%), has an H-value of 0. H increases with the number of alternative names given. For Object Agreement we report the mean across participants.

To facilitate the use of the norms, we added indices of Word Frequency, Word Prevalence (WP), and Age of Acquisition (AoA) from other sources to our dataset. We report the frequency per million words (SUBTLEXWF) and its log (Lg10WF) from the commonly used SUBTLEX-NL database (Keuleers, Brysbaert, & New, 2010), as well as Zipf values on a 7-point logarithmic scale, calculated as log₁₀(fpmw*1000) (see Van Heuven, Mandera, Keuleers, & Brysbaert,

2014 for more info; <https://osf.io/3d8cx/wiki/home/>). Word Prevalence scores, i.e. the percentage of a population knowing a word, were taken from a large online study by Keuleers, Stevens, Mandera, and Brysbaert (2015). Norms for Age of Acquisition, referring to the age at which a word was acquired, were collected and aggregated with data from a study by Brysbaert, Stevens, De Deyne, Voorspoels, and Storms (2014).

RESULTS AND DISCUSSION

Due to a technical error, only 465 trials were presented in all sessions (instead of 467 trials for sessions with set 1 and 465 for sets 2 and 3). However, because the items were randomized differently for each participant, two different photographs from set 1 were omitted from each participant’s list. For each photograph, a maximum of 52 responses could be obtained. Across all items, the percentages of DKO and DKN responses were 5% and 8%, respectively. For 31 out of 1397 items, fewer than 70% of participants produced a name, either because they did not know the object or because they did not know its name.

Norms for all 1397 photographs are provided in OSF Table B. This table includes the modal Dutch name for each item (modal_Name; name that was used by most participants, including inaccurate names), followed by the English (file)name (English_Name) as per Brodeur et al. (2010, 2014), the number of participants that saw this photograph (Nparts), and the number of DKO and DKN responses for this item. To calculate modal name agreement (modal_NA) and H-value, DKO and DKN (and “c”) responses were excluded first, but incorrect names were included. We also list modal name agreement as a proportion of the preferred name out of all responses, including DKO and DKN (and “c”) responses (modal_NA_all). Modal NA can be high because of one popular name or because a lot of participants did not respond, but there was consensus among those who did. (This is also helpful in the comparison with Table A2). Average object agreement (OA) was calculated over OA scores of participants using the modal name. Word Frequency (SUBTLEXWF, lg10WF, Zipf), Word Prevalence (WP) and Age of Acquisition (AoA) were added for the modal name of each item (Keuleers et al., 2010; Van Heuven et al., 2014; Brysbaert et al. 2014).

As we were also interested in alternatives for the modal names, the number of unique names per item (Names) and the two most frequent competitors (Alt1 and Alt2) to the modal name are reported, together with their Naming Agreement scores (Alt1_NA and Alt2_NA). The last two columns in the table indicate whether the modal name was a correct label for the photograph (modal_Valid; correct = 1; data were coded by multiple trained native speakers; hyper- and hyponyms, diminutives, plural-singular forms and non-standard names for the object were categorised as correct as well) and whether this item was included in the set of photographs used for further analysis (selected = 1; see below for exclusion criteria). As in Brodeur et al. (2010), photographs that were not recognized by many participants (DKO score over 20%), photographs that were named incorrectly by the majority of participants (invalid modal name), and photographs for which the most common name was used by fewer than 20% of the participants, were removed from the set of useful stimuli and put in a separate file of difficult to name pictures. In total, 208 photographs (15%) were removed from the analyses reported below.

Further analysis was carried out in R (R Core Team, 2018), with the remaining 1189 photographs (OSF TABLE C; note that invalid alternative names were excluded for Table C as well; modal NA per item is the same in Tables B and C). **Table 1** summarizes the norms for this set of items; **Figure 1** shows the frequency distribution of the norms.

VARIABLE	MEAN	SD	MIN	MAX
Modal name agreement (%)	71	23	21	100
H-value name	1.2	0.9	0	3.7
DKO (%)	2	4	0	20
DKN (%)	6	10	0	72
Object agreement	4.2	0.5	2	5
Frequency (Zipf)	3.5	0.9	1.7	6.6
Prevalence	1.84	0.13	0.88	1.96
Age of acquisition	7	2	4	14

Table 1 Norms of the 1189 photographs. Modal name agreement is the percentage of people that gave the most common name (calculated over the modal_NA, not modal_NA_all column in OSF Table C). H is a measure of entropy describing name agreement. DKO refers to the percentage of participant who indicated they did not recognize the object. DKN is the percentage of participant who indicated they did not know the name of the object. Object agreement is how well the photographs resembled the object shown (rated on a 5-point scale). Zipf scores for word frequency, word prevalence, and Age of Acquisition were added from different databases.

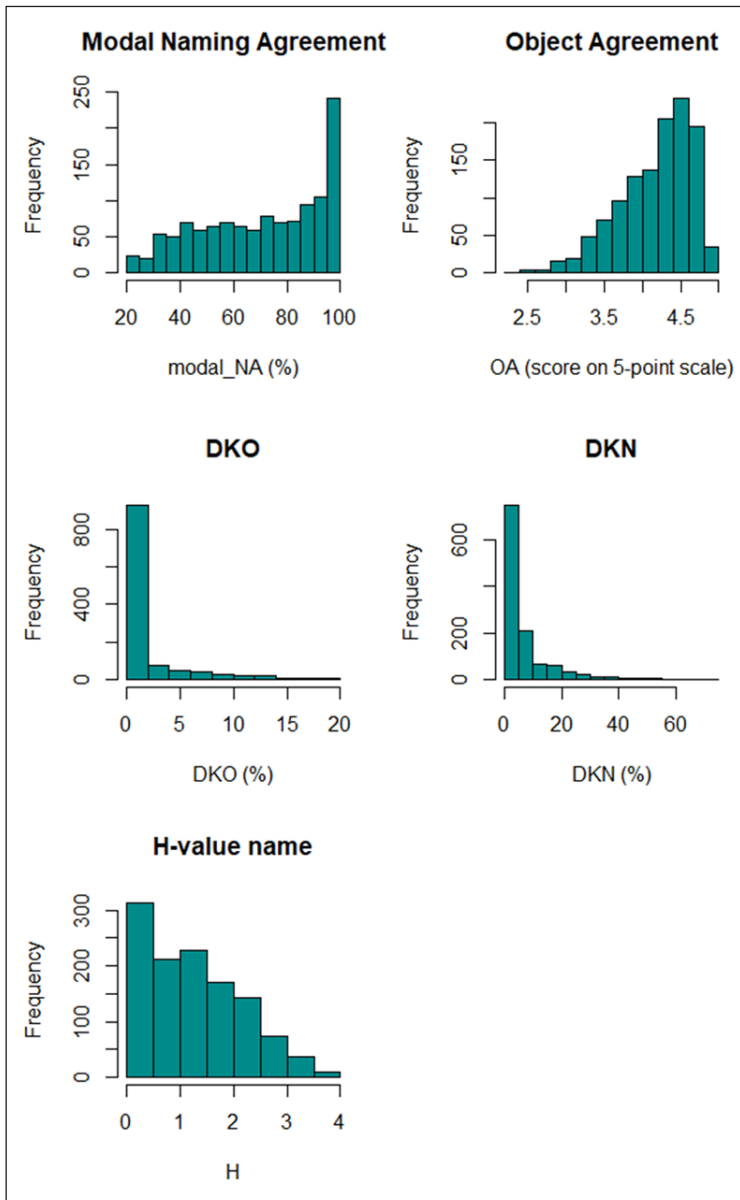


Figure 1 Frequency distributions of the five dependent variables in the current study.

As can be seen in **Table 1**, on average 2% of the participants indicated they did not know the object ($sd = 4$), and 6% did not know the name of the object ($sd = 10$). Modal naming agreement for the remaining items was 71% ($sd = 23$), and the average H-value for naming agreement was 1.2 ($sd = 0.9$). The average object agreement score was high (4.2, $sd = 0.5$, on a 5-point scale), showing that the photographs originally selected for Canadian participants are suitable for use in the Netherlands. The object names had an average Zipf score (word frequency) of 3.5 ($sd = 0.9$) corresponding to 3.16/million word. The average word prevalence was 1.84 ($sd = 0.13$), meaning that on average, the words were known by 97 % of the population. The AoA was 7 ($sd = 2$). All in all, the item set can be characterized as an easy set for Dutch university students.

To determine the reliability of our norms, we computed split-half correlations. We split the dataset into two sets based on participant number (even or odd). Valid items were selected for each group in the same way as for the entire dataset (modal name should be a correct name; $DKO < 20\%$; $modal_NA > 20\%$). The *even* and *odd* set each contained 1175 items. Because of the per-group filtering 90 photographs occurred in only one of the sets. Hence 1130 items were included in the correlations. We found high positive correlations for modal Name Agreement ($r_s = 0.89$, $p < .001$) and Object Agreement ($r_s = 0.79$, $p < .001$). We also compared the dominant names for each item across groups. This was the same for 1038 out of the 1130 items (92%). All in all, the NA norms show good reliability.

Table 2 shows the Spearman correlations among the variables assessed in the study and garnered from other sources. As expected, we found a strong negative correlation ($r_s = -0.97$, $p < .001$) between NA and H-value, as lower H-values imply stronger agreement. Thus we only comment on the correlations of the remaining variables with NA. Name agreement correlated

significantly with object agreement, showing that items rated as better representations of objects (high score on the 5-point scale) were named more consistently (higher NA) than items rated as poorer representations. As expected, name agreement correlated negatively with the proportions of DKO and DKN responses. This indicates that some objects were harder to recognize or name than others, yielding more varied and more missing responses. NA did not correlate with Zipf frequency, word prevalence, or AoA. Thus, participants were not more, or less, likely to agree on “easy” than on “harder” names. Finally, there were the well-documented significant correlations between AoA, Zipf frequency and word prevalence.

	NA	H	OA	%DKO	%DKN	ZIPF	WP	AOA
NA								
H	-0.97**							
OA	0.40**	-0.45**						
%DKO	-0.24**	0.28**	-0.51**					
%DKN	-0.42**	0.49**	-0.44**	0.47**				
Zipf	0.04	-0.03	-0.22**	-0.05	-0.12**			
	$p = .20$	$p = .25$		$p = .14$				
WP	0.08	-0.08	-0.08	-0.05	-0.12**	0.24**		
	$p = .02$	$p = .01$	$p = .02$	$p = .12$				
AoA	-0.08	0.09	0.13**	0.11**	0.21**	-0.63**	-0.30**	
	$p = .01$	$p = .006$						

As mentioned in the Introduction, researchers often need to select items with high name agreement. To facilitate such a selection, we categorized the items in OSF Table C as “good” (NA above 90%), “fairly good” (NA between 89 and 75%) and “poor” (the remaining items). There were 359 “good” and 226 “fairly good” items. For both categories the proportions of DKO and DKN responses were low (below 6%).

Since some studies might require sets of photographs with multiple plausible names, we determined how many items there were with two frequently given names. Note that we used the raw data (checked for spelling, but not aggregated for the same lemma, see OSF Table D) for all 1397 photographs in the analyses that are described next (e.g. *vogelhuis* and *vogelhuisje* are two strong names for the item birdhouse). We found a set of 35 *strong* pairs, where each of two names was used by between 40 and 60% of the participants, and a set of 187 *plausible* pairs, where each of two names was used by 25 to 50% of the participants. (Note that there is some overlap between these sets. There are 204 unique strong/plausible pairs; See OSF Tables D1 and D2. Of course researchers can select different types of pairs on the basis of the data given in OSF Table D, or Table B if one would like to look at aggregated names).

We were interested in the nature of the pairs in the *strong* and *plausible* sets. We established whether or not the two names were morphologically related, i.e. had at least one morpheme in common, and conducted a broad assessment of semantic relatedness. We found that 48% of the pairs (97 out of 204 pairs) had morphologically unrelated names (e.g. *kopje-mok*; *cup-mug*). These items might be most suitable for studies where different competing names are required. The remaining pairs showed various types of morphological relatedness. Most frequently, they consisted of a noun and a compound (28%; e.g. *stoel-tuinstoel*; *chair-garden chair*), or two compounds (12%; e.g. *kerstboom-denneboom*; *christmas tree-pine tree*), or a noun and its diminutive form (7%; e.g. *kaars-kaarsje*; *candle*).

Concerning the semantic relation between the members of the pairs, we found that most commonly (46% of the pairs) they were (1) hypo-hypernym pairs, as in *gitaar-elektrische gitaar* (*guitar-electric guitar*), or *aap-gorilla* (*ape-gorilla*). Synonyms (2), broadly defined (as in *wereldbol-globe*; *globe*), occurred less often (31% of the pairs). Pairs could also be (3) closely related concepts (18%), e.g. *salamander-hagedis* (*salamander-lizard*); or (4) could refer to very different concepts (4%), e.g. *koelbox-papierversnipperaar* (*cooler-shredder*). These categories need to be validated, but give an initial impression of the kinds of names competing. Most useful for research might be (1) and (4).

Table 2 Overview of Spearman’s rank correlation coefficients (r_s , rho) for correlations between all variables.

* Bonferroni correction: 28 pairwise comparisons, so significant at p-value smaller than .002 (.05/28).

** $p < .001$.

NA = Name Agreement; H = H-value; OA = Object Agreement; %DKO = percentage of trials in which participants did not recognize the object; %DKN = percentage of trials in which participants did not know the name of the object; Zipf = Zipf score for word frequency; WP = Word Prevalence (z-scores); AoA = Age of Acquisition.

Finally, although a cross-linguistic comparison was not the main goal of this work, we compared our results to those in the norming studies conducted by Brodeur and colleagues. Compared to the Canadian data set, Dutch modal NA (71%, *sd* = 23) was slightly higher (64% in Brodeur et al., 2010; 58% and 61% in Brodeur et al., 2014) and thus the average H-value (1.2, *sd* = 0.9) slightly lower (1.65 in Brodeur et al. 2010; 1.89 and 1.53 in Brodeur et al. 2014). Thus, the Dutch participants used, on average, the modal name more frequently and used fewer alternative names per photograph than the Canadian participants. The object agreement scores in the two studies were similar (average 4.2, *sd* = 0.5 in the present study, compared to 3.90, *sd* = 0.50 in Brodeur et al. (2010) and 3.69, *sd* = 0.52 and 3.57, *sd* = 0.57 in Brodeur et al. (2014)). Correlations between variables (NA, H, and OA), were very similar too (see **Table 3**).

	NAMING AGREEMENT			OBJECT AGREEMENT		
	DUTCH (<i>r_s</i>)	2010 (<i>r</i>)	2014(<i>r</i>)	DUTCH (<i>r_s</i>)	2010 (<i>r</i>)	2014 (<i>r</i>)
Object Agreement	.40**	.33*	.29*			
H-value	-.97**	-.96*	-.95*	-.45**	-.38*	-.35*

Table 3 Correlations between the various dependent variables in the present study (Dutch) and the two studies of Brodeur et al. (2010, 2014).

In sum, we present norms for written naming for 1397 photographs provided by young adult speakers of Dutch. The set includes 359 items with good name agreement, and it includes 35 items with two strong competing names. These sets might be most useful for designing studies. We highlight again that participants typed their responses. Spoken name agreement may deviate somewhat from the norms reported here because, for instance, some names are used more often in spoken than written language. Nonetheless, we hope that the current norms can assist researchers in their item selection.

DATA ACCESSIBILITY STATEMENT

All photographs, together with the norms (see Table B for all 1397 pictures, and Table C for norms for the useful set of 1185 pictures) and additional information on lexical competitors (Table D, D1, and D2) are available at <https://osf.io/kwu87/>.

In addition to the tables referred to in the manuscript, we uploaded a table (Table A1) with the raw data and a table (Table A2) with naming agreement for every used name for a certain picture after aggregating. Table A1 contains all individual responses (before aggregating; includes spelling mistakes and typo's) and reports picture set (1–3), (anonymized) participant no., trial no., English picture name, used Dutch name, and object agreement score for each trial. This file will be of interest to researchers who want to try out different ways of summarizing the data, who are investigating the details of names given to pictures, or who are interested in spelling errors. Table A2 provides an overview of all different names used for each picture (after spelling check and aggregating), together with the proportion of participants that used this name (NA_all), calculated as the number of occurrences of this name (Used) divided by the number of participants that saw this picture (Nparts), i.e. including DKO and DKN responses. The OSF project also contains the R script used for all analyses (incl. all necessary input files; one of these files is a .txt file with the raw data (responses in column Input), note that in some trials there is an inconsistency between Button and Score. This is because some participants typed 'a' or 'b' instead of using the buttons. We removed the Input and manually changed the Score so these responses would be treated as a NO RESPONSE and could be categorized as DKO or DKN. However, we did not change the value in the Button column to keep the participant's original response in the dataset.).

ETHICS AND CONSENT

Ethical approval to conduct the study was given by the Ethics Board of the Social Sciences Faculty of Radboud University. Written informed consent was obtained from all subjects in this study.

COMPETING INTERESTS

The authors have no competing interests to declare.

C. Decuyper  orcid.org/0000-0002-8623-4398

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

M. Brysbaert  orcid.org/0000-0002-3645-3189

Department of Experimental Psychology, Ghent University, Gent, Belgium

M. B. Brodeur

Douglas Mental Health University Institute, Montréal, Canada

A. S. Meyer

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

REFERENCES

- 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. <https://doi.org/10.3758/BF03195559>
- 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 Section A*, 50(3), 560–585. DOI: <https://doi.org/10.1080/783663595>
- Biederman, I., & Ju, G.** (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20(1), 38–64. DOI: [https://doi.org/10.1016/0010-0285\(88\)90024-2](https://doi.org/10.1016/0010-0285(88)90024-2)
- Bonin, P., Chalard, M., Méot, A., & Fayol, M.** (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, 93(1), 89–114. DOI: <https://doi.org/10.1348/000712602162463>
- Bonin, P., Méot, A., Laroche, B., Bugaiska, A., & Perret, C.** (2019). The impact of image characteristics on written naming in adults. *Reading and Writing*, 32(1), 13–31. DOI: <https://doi.org/10.1007/s11145-017-9727-3>
- 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. <https://doi.org/10.3758/BF03195507>
- Bose, A., & Schafer, G.** (2017). Name agreement in aphasia. *Aphasiology*, 31(10), 1143–1165. DOI: <https://doi.org/10.1080/02687038.2016.1254148>
- 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), e10773. DOI: <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), e106953. DOI: <https://doi.org/10.1371/journal.pone.0106953>
- Brodeur, M. B., Kehayia, E., Dion-Lessard, G., Chauret, M., Montreuil, T., Dionne-Dostie, E., & Lepage, M.** (2012). The bank of standardized stimuli (BOSS): Comparison between French and English norms. *Behavior Research Methods*, 44(4), 961–970. DOI: <https://doi.org/10.3758/s13428-011-0184-7>
- Brodie, E. E., Wallace, A. M., & Sharrat, B.** (1991). Effect of Surface Characteristics and Style of Production on Naming and Verification of Pictorial Stimuli. *The American Journal of Psychology*, 104(4), 517. DOI: <https://doi.org/10.2307/1422939>
- Brown-Schmidt, S.** (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61(2), 171–190. DOI: <https://doi.org/10.1016/j.jml.2009.04.003>
- Brysbaert, M., Stevens, M., Deyne, S. D., Voorspoels, W., & Storms, G.** (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80–84. DOI: <https://doi.org/10.1016/j.actpsy.2014.04.010>
- Campbell, D. T., & Fiske, D. W.** (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105. DOI: <https://doi.org/10.1037/h0046016>
- Cheng, X., Schafer, G., & Akyürek, E. G.** (2010). Name agreement in picture naming: An ERP study. *International Journal of Psychophysiology*, 76(3), 130–141. DOI: <https://doi.org/10.1016/j.ijpsycho.2010.03.003>
- Clarke, B., & Ludington, J.** (2018). Thai Norms for Name, Image, and Category Agreement, Object Familiarity, Visual Complexity, Manipulability, and Age of Acquisition for 480 Color Photographic Objects. *Journal of Psycholinguistic Research*, 47, 607–626. DOI: <https://doi.org/10.1007/s10936-017-9544-5>
- Cuetos, F., Ellis, A. W., & Alvarez, B.** (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, 31(4), 650–658. DOI: <https://doi.org/10.3758/BF03200741>

- 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. DOI: <https://doi.org/10.3758/BF03200832>
- Eid, M., & Diener, E.** (Series Eds.). (2006). *Handbook of multimethod measurement in psychology*, xiv, 553. Washington, DC, US: American Psychological Association. DOI: <https://doi.org/10.1037/11383-000>
- Ellis, A. W., & Morrison, C. M.** (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 515–523. DOI: <https://doi.org/10.1037/0278-7393.24.2.515>
- Gollan, T. H., Weissberger, G. H., Runqvist, E., Montoya, R. I., & Cera, C. M.** (2012). Self-ratings of Spoken Language Dominance: A Multi-Lingual Naming Test (MINT) and Preliminary Norms for Young and Aging Spanish-English Bilinguals. *Bilingualism (Cambridge, England)*, 15(3), 594–615. DOI: <https://doi.org/10.1017/S1366728911000332>
- van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M.** (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, 67(6), 1176–1190. DOI: <https://doi.org/10.1080/17470218.2013.850521>
- Jescheniak, J. D., & Schriefers, H.** (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(5), 1256–1274. DOI: <https://doi.org/10.1037/0278-7393.24.5.1256>
- Jongman, S. R., Roelofs, A., & Lewis, A. G.** (2020). Attention for Speaking: Prestimulus Motor-cortical Alpha Power Predicts Picture Naming Latencies. *Journal of Cognitive Neuroscience*, 32(5), 747–761. DOI: https://doi.org/10.1162/jocn_a_01513
- Keuleers, E., Brysbaert, M., & New, B.** (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3), 643–650. DOI: <https://doi.org/10.3758/BRM.42.3.643>
- Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M.** (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology (2006)*, 68, 1–62. DOI: <https://doi.org/10.1080/17470218.2015.1022560>
- Lachman, R., Shaffer, J. P., & Hennrikus, D.** (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior*, 13(6), 613–625. DOI: [https://doi.org/10.1016/S0022-5371\(74\)80049-6](https://doi.org/10.1016/S0022-5371(74)80049-6)
- Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J.** (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122–142. DOI: <https://doi.org/10.1037/0033-295X.98.1.122>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S.** (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(01), 1–75. DOI: <https://doi.org/10.1017/S0140525X99001776>
- Madden Daniel, L., Sale Martin, V., & Robinson Gail, A.** (2019). Differentiating Beyond Name Agreement for Picture Naming: Insight From Age-Related Selection Deficits. *Journal of Speech, Language, and Hearing Research*, 62(5), 1373–1380. DOI: https://doi.org/10.1044/2018_JSLHR-L-18-0095
- Miozzo, M., Pulvermüller, F., & Hauk, O.** (2015). Early Parallel Activation of Semantics and Phonology in Picture Naming: Evidence from a Multiple Linear Regression MEG Study. *Cerebral Cortex (New York, N.Y. : 1991)*, 25(10), 3343–3355. PubMed (25005037). DOI: <https://doi.org/10.1093/cercor/bhu137>
- 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), e37527. DOI: <https://doi.org/10.1371/journal.pone.0037527>
- Nakabayashi, K., Burton, A. M., Brandimonte, M. A., & Lloyd-Jones, T. J.** (2012). Dissociating positive and negative influences of verbal processing on the recognition of pictures of faces and objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(2), 376–390. DOI: <https://doi.org/10.1037/a0025782>
- Ostergaard, A. L., & Davidoff, J. B.** (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(3), 579–587. DOI: <https://doi.org/10.1037/0278-7393.11.3.579>
- O'Sullivan, M., Lepage, M., Bouras, M., Montreuil, T., & Brodeur, M. B.** (2012). North-American Norms for Name Disagreement: Pictorial Stimuli Naming Discrepancies. *PLOS ONE*, 7(10), e47802. DOI: <https://doi.org/10.1371/journal.pone.0047802>
- Paivio, A., Clark, J. M., Digdon, N., & Bons, T.** (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition*, 17(2), 163–174. DOI: <https://doi.org/10.3758/BF03197066>
- Peterson, R. R., & Savoy, P.** (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(3), 539–557. DOI: <https://doi.org/10.1037/0278-7393.24.3.539>
- Piai, V., Roelofs, A., Rommers, J., & Maris, E.** (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping*, 36(7), 2767–2780. DOI: <https://doi.org/10.1002/hbm.22806>

- Shao, Z., Roelofs, A., Acheson, D. J., & Meyer, A. S.** (2014). Electrophysiological evidence that inhibition supports lexical selection in picture naming. *Brain Research*, 1586, 130–142. DOI: <https://doi.org/10.1016/j.brainres.2014.07.009>
- Shao, Z., & Stiegert, J.** (2016). Predictors of photo naming: Dutch norms for 327 photos. *Behavior Research Methods*, 48(2), 577–584. DOI: <https://doi.org/10.3758/s13428-015-0613-0>
- 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 and Memory*, 6(2), 174–215. DOI: <https://doi.org/10.1037/0278-7393.6.2.174>
- Snodgrass, J. G., & Yuditsky, T.** (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, 28(4), 516–536. DOI: <https://doi.org/10.3758/BF03200540>
- Torrance, M., Nottbusch, G., Alves, R. A., Arfé, B., Chanquoy, L., Chukharev-Hudilainen, E., ... Wengelin, Å.** (2018). Timed written picture naming in 14 European languages. *Behavior Research Methods*, 50(2), 744–758. DOI: <https://doi.org/10.3758/s13428-017-0902-x>
- Vitkovitch, M., & Tyrrell, L.** (1995). Sources of Disagreement in Object Naming. *The Quarterly Journal of Experimental Psychology Section A*, 48(4), 822–848. DOI: <https://doi.org/10.1080/14640749508401419>
- Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S.** (2019). The production effect and the generation effect improve memory in picture naming. *Memory*, 27(3), 340–352. DOI: <https://doi.org/10.1080/09658211.2018.1510966>

TO CITE THIS ARTICLE:

Decuyper, C., Brysbaert, M., Brodeur, M. B., & Meyer, A. S. (2021). Bank of Standardized Stimuli (BOSS): Dutch Names for 1400 Photographs. *Journal of Cognition*, 4(1): 33, pp. 1–11. DOI: <https://doi.org/10.5334/joc.180>

Submitted: 14 October 2020

Accepted: 09 July 2021

Published: 23 July 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Cognition is a peer-reviewed open access journal published by Ubiquity Press.