



HAL
open science

IMABASE: A new set of 313 colourised line drawings standardised in French for name agreement, image agreement, conceptual familiarity, age-of-acquisition, and imageability

Patrick Bonin, Bénédicte Poulin-Charronnat, Helle Lukowski Duplessy,
Patrick Bard, Annie Vinter, Ludovic Ferrand, Alain Méot

► **To cite this version:**

Patrick Bonin, Bénédicte Poulin-Charronnat, Helle Lukowski Duplessy, Patrick Bard, Annie Vinter, et al.. IMABASE: A new set of 313 colourised line drawings standardised in French for name agreement, image agreement, conceptual familiarity, age-of-acquisition, and imageability. *Quarterly Journal of Experimental Psychology*, 2020, 73 (11), pp.1862-1878. 10.1177/1747021820932822 . hal-02909374

HAL Id: hal-02909374

<https://hal.science/hal-02909374>

Submitted on 15 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IMABASE: A new set of 313 colored line drawings standardized in French for name agreement, image agreement, conceptual familiarity, age-of-acquisition, and imageability

Journal:	<i>Quarterly Journal of Experimental Psychology</i>
Manuscript ID	QJE-STD-19-433.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	06-May-2020
Complete List of Authors:	Bonin, Patrick; LEAD-CNRS UMR5022, Univ. Bourgogne Franche-Comté Poulin-Charronnat, Bénédicte; LEAD-CNRS UMR5022, Univ. Bourgogne Franche-Comté Lukowski Duplessy, Helle; Cognition Lab, Copenhagen BARD, PATRICK; LEAD-CNRS UMR5022, Univ. Bourgogne Franche-Comté Vinter, Annie; LEAD-CNRS UMR5022, Univ. Bourgogne Franche-Comté Ferrand, Ludovic; Université Clermont Auvergne, CNRS, LAPSCO Méot, Alain; Université Clermont Auvergne, CNRS, LAPSCO
Keywords:	line drawings, psycholinguistic norms

SCHOLARONE™
Manuscripts

Accepted ms: QJE-QJE-STD-19-433.R1 - 9 May 2020

IMABASE: A new set of 313 colorized line drawings standardized in French for name agreement, image agreement, conceptual familiarity, age-of-acquisition, and imageability

Patrick Bonin¹, Bénédicte Poulin-Charronnat¹, Helle Lukowski Duplessy¹, Patrick Bard¹,
Annie Vinter¹, Ludovic Ferrand², and Alain Méot²

¹LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

² LAPSCO - CNRS UMR6024, Université Clermont Auvergne, Clermont-Ferrand, France

Address for correspondence:

Patrick Bonin

LEAD-CNRS (UMR5022)

University of Bourgogne

Pôle AAFE - Esplanade Erasme

BP 26513

21065 Dijon Cedex

France

Email: Patrick.Bonin@u-bourgogne.fr

Contact information:

Patrick Bonin (patrick.bonin@u-bourgogne.fr)

LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

Bénédicte Poulin-Charronnat (benedicte.poulin@u-bourgogne.fr)

LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

Helle Lukowski Duplessy (hedu@psy.ku.dk)

LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

Patrick Bard (Patrick.Bard@u-bourgogne.fr)

LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

Annie Vinter (Annie.Vinter@u-bourgogne.fr)

LEAD - CNRS UMR5022, Univ. Bourgogne Franche-Comté, Dijon, France

Ludovic Ferrand (ludovic.ferrand@uca.fr)

LAPSCO - CNRS UMR6024, Université Clermont Auvergne, Clermont-Ferrand, France

Alain Méot (alain.meot@uca.fr)

LAPSCO - CNRS UMR6024, Université Clermont Auvergne, Clermont-Ferrand, France

Running head:

NORMS FOR A NEW SET OF COLORIZED DRAWINGS

Abstract

We provide normative data for a new set of 313 colorized line drawings. The drawings were standardized on name agreement [$N = 60$ participants], image agreement [$N = 34$], conceptual familiarity [$N = 36$], age of acquisition [$N = 35$], and imageability [$N = 35$]. Objective visual complexity measures are given for the pictures, and objective word frequencies are provided for the modal names of the drawings. Reliability measures for the collected norms are very high. There are high levels of agreement between the names given by the participants and the drawings and comparative analyses indicate that the distribution of name agreement scores is very similar in both our own database and the MultiPic database (Duñabeitia et al., 2018). A novel “picture-choice task” used to assess name-image agreement [$N = 30$] reveals that the great majority of the IMABASE pictures that are also present in MultiPic are rated as providing better pictorial representations of the corresponding concepts. Finally, most of the correlations are comparable with those reported in other normative studies on colorized drawings. The whole set of pictures is freely available from <http://leadserv.u-bourgogne.fr/~lead/openlead> and the norms are available as Supplementary Material.

Keywords: line drawings; psycholinguistic norms

Several normed picture databases are available in the literature. The most famous one is the Snodgrass and Vanderwart (1980) database that provides normative data in American-English on name agreement, image agreement, conceptual familiarity, and visual complexity (see below for a description of these norms) for a set of 260 black-and-white drawings. The goal of Snodgrass and Vanderwart—hereafter SV—was to provide researchers with pictures of drawings and psycholinguistic norms that could be used for the careful design of experiments on perception, memory, and language. The SV database has been impressively adopted by the (mainly psychological) scientific community, and the corresponding paper has been frequently cited (Duñabeitia et al., 2018). SV has encouraged research into the collection of norms for pictured stimuli, and the set of SV pictures has been normed in many different populations and language communities. To name but a few, the SV pictures have been normed for Spanish (Sanfeliù & Fernandez, 1996), British-English (Barry, Morrison, & Ellis, 1997), French (Alario & Ferrand, 1999), Icelandic (Pind, Jónsdóttir, Tryggvadóttir, & Jónsson, 2000), Japanese (Nishimoto, Miyawaki, Ueda, Une, & Takahashi, 2005), Chinese (Weekes, Shu, Hao, Liu, & Tan, 2007), and more recently, for Croatian (Rogić et al., 2013) and Tunisian Arabic (Boukadi, Zouaidi, & Wilson, 2016). This set of pictures has also been normed in English-speaking children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997) and in Portuguese-speaking children (Pompéia, Miranda, & Bueno, 2001).

The SV pictures have been colorized (Rossion & Pourtois, 2004), and norms for these colorized pictures have been collected in Belgian French (Rossion & Pourtois, 2004), in Modern Greek (Dimitropoulou, Duñabeitia, Blitsas, & Carreiras, 2009), in Russian (Tsaparina, Bonin, & Méot, 2011), in Turkish (Raman, Raman, & Mertan, 2014), and in Persian (Bakhtiar, Nilipour, & Weekes, 2013). Other sets of drawings have been constructed and normed (e.g., Bates et al., 2003 [520 black-and-white line drawings of common objects

1
2
3 including 174 pictures from the original SV set]; Bonin, Peereman, Malardier, Méot, &
4 Chalard, 2003 [299 black-and-white drawings of objects]; Khwaileh, Mustafawi, Herbert, &
5 Howard, 2018 [319 black-and-white line drawings of objects]; Martein, 1995 [216 black-and-
6 white line drawings of objects]; Saryazdi, Bannon, Rodrigues, Klammer, & Chambers, 2018
7 [225 everyday objects, each depicted both as a photograph and a matched clipart image];
8 Székely et al., 2004 [520 black-and-white line drawings of common objects]). Finally,
9 photograph databases for objects have also been designed and normed (Brodeur, Dionne-
10 Dostie, Montreuil, & Lepage, 2010 [480 color photographic images of real objects]; Brodeur,
11 Guérard, & Bouras, 2014 [930 color photographic images of real objects]; Clarke &
12 Ludington, 2018 [norms in Thai for the Broder et al. photo set]; Moreno-Martínez &
13 Montoro, 2012[360 high-quality colour images of objects]; Salmon, McMullen, & Filliter,
14 2010 [320 black-and-white photographs of objects]).

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 The availability of normed pictures has helped in the design of studies investigating
32 the factors influencing naming speed in healthy adults (e.g., Alario et al., 2004; Bonin,
33 Chalard, Méot, & Fayol, 2002; Cuetos, Ellis, & Alvarez, 1999; Snodgrass & Yuditsky, 1996),
34 as well as naming accuracy in patients (e.g., Cuetos, Aguado, Izura, & Ellis, 2002). Likewise,
35 the different processes and representations that are involved in object naming have been
36 delineated. It is generally agreed that object naming involves three main processing stages: (1)
37 object comprehension, which entails the perceptual analysis of the object, access to stored
38 structural representations, and, finally, access to semantic information; (2) name retrieval,
39 which according to certain views of speech production (Levelt, Roelofs, & Meyer, 1999), is
40 conceptualized as a two-stage process consisting of (a) lexical selection, i.e., activation of
41 abstract lexical entries that are not phonologically specified: lemmas, and (b) word-form
42 encoding; and (3) articulation. In adults, different factors have been found to influence the
43 various stages of picture naming (Bonin, 2013; Roelofs & Ferreira, 2019, for reviews). As
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 claimed by Alario et al. (2004), the most important predictors of naming speed include name
4 agreement and age of acquisition (AoA) of the object names, followed by the objective
5 frequency of the names (objective lexical frequency norms are available from different
6 computerized databases in different languages; for the French language, Lexique.org provides
7 different word frequency measures, New, Pallier, Brysbaert, & Ferrand, 2004) and, finally,
8 image agreement. Other predictors (e.g., imageability, conceptual familiarity, visual
9 complexity) have been found to exert a reliable influence on naming speed, but in a less
10 consistent manner (Bonin, 2013; see Perret & Bonin, 2019, for a Bayesian meta-analysis of
11 the predictors of naming speed).

Why collect norms on a new set of colorized pictures?

26
27 Given the availability of pictures and psycholinguistic norms for a large number of
28 items in the literature, readers may wonder why we decided to design a totally new set of
29 pictures, corresponding largely to the SV concepts, and collect norms in adults on these new
30 pictures and their names.

31
32 One important issue relating to the SV pictures, even in their colorized version
33 (Rossion & Pourtois, 2004), is that they are at present not fully optimized for their use in
34 experiments on perception, memory, or language processes for the following reasons. First of
35 all, many of the SV pictures need to be updated. There are many instances of old-fashioned
36 representations of objects. For example, the way a doll, a phone, a television, a doorknob, or a
37 train is represented in the SV database no longer corresponds to the typical doll, phone,
38 television, doorknob or train that children or adults see in modern life. Certain items that were
39 very familiar in the past (e.g., tape recorder, spinning wheel, rolling pin, dresser, salt shaker,
40 thimble) are no longer so. Also, many objects present in our environment today are absent in
41 the set of 260 black-and-white drawings published in 1980. It is important that norms
42 collected for a set of stimuli at a given period of time within a given language and/or
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 community be updated, because objects represented in an unusual manner will be difficult to
4
5 identify. This is especially damaging if one's aim is to investigate the processes underpinning
6
7 picture naming. Indeed, in adults, Johnston, Dent, Humphreys, and Barry (2010) found that
8
9 their new and more up-to-date ratings accounted for a larger amount of variance in naming
10
11 times than did those taken from an earlier study (e.g., Barry et al., 1997). Of course, having
12
13 norms for objects that are typical of the current environment is also very useful to all
14
15 researchers who use pictures as stimuli in experiments with children. Second, the different
16
17 items in the SV database are taken from different semantic categories (15 categories) but, for
18
19 certain categories, there are too few exemplars (e.g., vehicles: $n = 10$; birds: $n = 8$; or insects:
20
21 $n = 8$). As far as possible, we therefore added some more items for these semantic categories
22
23 in order to achieve a better representation of each category in the new database. Third, certain
24
25 SV pictures are typical of the American culture (see, in particular, Duñabeitia et al., 2018
26
27 footnote 2; Yoon, et al., 2004), and do not correspond to the way most adults and children
28
29 represent the corresponding concepts in other cultures. To give a few examples, the drawings
30
31 of plug, bread, barn, key, fridge, sandwich, and bus are all clearly typical of American culture.
32
33 Fourth, many pictures in SV are difficult to recognize, even for adults, as assessed by name-
34
35 agreement or image-agreement scores. The French normative study by Alario and Ferrand
36
37 (1999) revealed that there were 39 out of 260 pictures of objects that had name-agreement
38
39 scores lower than 75% even for very familiar concepts, such as pig, train, frog, toe, pen,
40
41 camel, bread, or leopard. Many pictures are rated as being "not typical" of the objects they
42
43 represent: 68 pictures out of 260 have a rating lower than 3 (on a five-point scale with 5 =
44
45 very high agreement between the drawing and the mental representation of the underlying
46
47 concept). This represents a clear limitation when it comes to selecting items for the design of
48
49 experiments involving both children and adults.
50
51
52
53
54
55
56
57
58
59
60

1
2
3 It is worth noting that we are not the only authors to point to the need to provide
4 researchers with new databases of colorized line drawings or new sets of photographs. Russo
5 et al. (2018) have provided a database of 640 colorized images of animals (normed on name
6 agreement, picture name agreement, and AoA) because the number of exemplars per category
7 of animals is generally low in the available databases. As claimed by Duñabeitia et al. (2018),
8 there is in fact a limited choice for researchers who want to use (colorized) drawings in their
9 experiments. Fortunately, they have provided a set of 750 drawings that have been normed in
10 six European languages: the MultiPic database. We became aware of this work while we were
11 designing our own database. However, the items in MultiPic have been normed on only two
12 variables—name agreement and visual complexity—which represents a serious limitation, in
13 particular for the design of naming experiments in which a number of influential factors need
14 to be controlled for methodological or statistical purposes (e.g., image agreement, AoA; see
15 Perret & Bonin, 2019).
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 To design the present picture database, we built on the existing list of depicted SV
34 concepts because these concepts are generally related to everyday common objects. It is
35 indeed useful to have pictures corresponding to these concepts in order to design naming
36 experiments, especially in order to permit comparisons of the results with those of previous
37 naming studies based on such concepts. We decided to use *colorized* line drawings instead of
38 more traditional black-and-white drawings. The choice of colorized drawings was motivated
39 by the fact that faster and/or more accurate naming performances have been obtained with
40 colorized than with black-and-white drawings (Bonin, Méot, Laroche, Bugajska, & Perret,
41 2019; Rossion & Pourtois, 2004). Norms on photographs are undoubtedly useful since they
42 represent real instances of objects, i.e., idiosyncratic (token) representations of objects.
43
44
45
46
47
48
49
50
51
52
53
54
55
56 However, we wanted to have generic (type) representations of objects because our database
57 was also designed to be useful for researchers investigating naming in children; generic
58
59
60

1
2
3 representations seem to be more suitable to this end. More importantly, some findings in the
4
5 literature indicate that naming accuracy is lower in response to photographs than to
6
7 corresponding drawings of the same objects (Clarke & Ludington, 2018). For instance, the
8
9 overall name agreement is higher in the SV database (88%), as well as in the Bonin et al.
10
11 (2003) database that was designed to complement the SV set (77%), than in the BOSS
12
13 database (64% in the Brodeur et al. study [2010] and 59.45% with the BOSS pictures normed
14
15 in Thai, Clarke, & Ludington [2018]), which uses photographs of objects. As claimed by
16
17 O'Sullivan, Lepage, Bouras, Montreuil, and Brodeur (2012), photographs include details and
18
19 colors that prompt participants to rely on idiosyncratic features and, as a result, they elicit a
20
21 number of different names. This can be a problem when designing experiments in which
22
23 naming times are the dependent variable (but see Saryazdi et al., 2018).
24
25
26
27

28
29 Turning to the investigation of memory processes, certain findings suggest that the
30
31 degree of visual iconicity, i.e., the degree to which images resemble the real-world objects
32
33 they depict (Saryazdi et al., 2018), does not play a significant role in memory performance
34
35 since similar levels of recall and recognition have been found with the use of colored
36
37 photographs and black-and-white line drawings of objects (Snow, Skiba, Coleman, &
38
39 Berryhill, 2014).
40
41
42

43 **The present normative study: norms collected from pictures and their names**

44
45 Norms correspond either to characteristics of the pictures or to their names. Among
46
47 the norms collected from pictures are the visual complexity of the pictures, the 'typicality' of
48
49 the depicted drawings as assessed by image agreement, the degree of agreement of the visual
50
51 representation of objects with their names (i.e., name agreement), the familiarity of the
52
53 depicted concepts. Ratings of AoA are also often collected from object names. To collect
54
55 name agreement, image agreement, visual complexity as well as conceptual familiarity ratings
56
57 from pictures, adults are generally tested collectively in tasks in which pictures (or words) are
58
59
60

1
2
3 projected onto a large white screen by means of an overhead projector. However, these norms
4
5 can also be collected individually by presenting pictures on a computer screen. AoA from
6
7 written words is also collected individually or collectively. Likert scales are used for the
8
9 ratings. In the present study, we collected norms on name agreement, image agreement, and
10
11 conceptual familiarity from the pictures; and on AoA and imageability from the written
12
13 names (in the Procedure subsection, we provide detailed information about how these norms
14
15 were collected).

16
17
18
19 Name agreement corresponds to the degree of agreement among individuals on a
20
21 specific name to be used to refer to a pictured object, whereas image agreement corresponds
22
23 to the degree to which the mental images formed by participants in response to an object
24
25 name match the object's appearance. Name agreement is one of the strongest factors
26
27 influencing naming performance, and in particular naming speed (e.g., Barry et al., 1997;
28
29 Bonin et al., 2002; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996; Valente, Burki, &
30
31 Laganaro, 2014; Vitkovitch & Tyrrell, 1995). In word production, name agreement acts either
32
33 at the comprehension level or at the lexical level (Barry et al., 1997). Image agreement has its
34
35 effect at the level of structural representations, which correspond to the canonical perceptual
36
37 representations of objects (Humphreys, Riddoch, & Quinlan, 1988).

38
39
40
41
42 Because different colorized drawings are used to depict the same concepts common to
43
44 MultiPic (Duñabeitia et al., 2018), Rossion and Pourtois' (2004) database, and the current
45
46 database, we also included a picture-choice task in order to assess which drawings best
47
48 depicted the common concepts. In this task, participants were first presented with a written
49
50 name and had to form a mental representation of the concept referred to by the name.
51
52 Depending on the trial, two or three pictures were displayed on a computer screen and the
53
54 participants had to choose the one which, according to them, was the best depiction of the
55
56 object name.
57
58
59
60

1
2
3 Conceptual familiarity is a measure of the acquaintance with the concept depicted by
4 the picture (or referred to by the object name. When pictures are used for these ratings, care is
5 taken to explain that the ratings have to be attributed to the concept itself, and not to the way
6
7
8
9
10 it is represented. This variable is thought to index the conceptual level in lexical processing
11 tasks. Mixed findings have been reported concerning the impact of this variable in spoken or
12
13
14
15 written naming studies (see for a review Perret & Bonin, 2019).

16
17 When evaluating the visual complexity of pictures, participants are required to take
18 account of the number of lines and details in the drawing. Evidence for the influence of
19 this variable is mixed at best, with certain studies reporting that visually complex drawings
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

 Peereman, et al., 2003; and see Perret & Bonin, 2019). Székely and Bates (2000) have
proposed other measures of visual complexity that do not rely on subjective ratings such as
the size of the digitized picture file, i.e., objective visual complexity (see also Székely et
al., 2004). They claim that using objective visual complexity scores avoids the problem
that subjective ratings of visual complexity may be influenced by psycholinguistic
variables (e.g., familiarity with the object) that are not directly related to visual complexity.
As in Tsaparina et al.'s (2011) study, we decided to report only objective visual
complexity scores (the number of bytes in JPEG format) for our set of colorized pictures.¹

 In order to estimate the AoA of the words, participants use Likert scales. For instance,
they have to estimate the age at which they think they learned the word *cow* in its written or
oral form. In general, the values of the scale correspond to x-year age bands, for instance 2-
year age bands. AoA is an important determinant of lexical processing and one of the most

important determinants of object naming speed (see Johnston & Barry, 2006; Juhasz, 2005 for reviews).

Imageability corresponds to ratings of the ease with which mental images can be generated in response to words. Imageability is used as an index of semantic code activation in lexical processing (e.g., Cortese, Simpson, & Woolsey, 1997; Strain, Patterson, & Seidenberg, 1995, 2002). Finally, we also reported objective word frequency (subtitle frequency and book frequency measures taken from Lexique.org, New et al., 2004) for the modal names provided for each object since word frequency is certainly the most popular variable in psycholinguistics (see Brysbaert, Mandera, & Keuleers, 2019, for a recent review).

In the following, we report several analyses that were performed on the collected data and on the data taken from the Multipic and RP databases. First of all, we describe the reliabilities that were computed for the different collected norms from the new set of items. Second, we report several comparative analyses on name agreement scores in order to investigate precisely how the naming outcomes from IMABASE differ from those obtained with the Multipic and RP databases. In addition, as indicated above, we used a “novel” task—a “picture-choice task” —to assess further name-image agreement. Here also, we compared the different scores obtained for our pictures to those obtained for the pictures taken from Multipic and RP. Third, descriptive statistics are reported together with the distributions of the norms. Fourth, we present the bivariate correlations that were performed on the collected variables. These analyses were conducted to analyze and summarize the correlational structure of the norms, and to compare it to other normative studies on colorized drawings. Finally, we report multiple regression analyses that were performed in order to analyze more deeply the relations between each subjective norm and the other dimensions.

Method

Participants. The participants ($N = 230$; 39 males, 190 females and one participant who did

not specify sex; mean age: 20.35 years old; range: 18 to 33 years old) were all native speakers of French and had normal or corrected-to-normal vision. Six subgroups were established corresponding to the six collected norms, with the mean ages and standard deviations of the participants being as follows: for name agreement ($N = 60$ because two sublists were used, each involving 30 participants), $M = 20.48$, $SD = 1.76$; for the picture-choice task ($N = 30$), $M = 20.17$, $SD = 1.12$; for image agreement ($N = 34$), $M = 20.88$, $SD = 3.09$; for familiarity ($N = 36$), $M = 20.44$, $SD = 1.92$; for AoA ($N = 35$), $M = 20.2$, $SD = 1.97$; and for imageability ($N = 35$), $M = 19.91$, $SD = 1.76$.

The different rating tasks were performed collectively with each participant being randomly assigned to one of the rating tasks. All participants provided fully informed consent to participate and received course credits for their participation. The study was approved by the Statutory Ethics Committee from the University Clermont Auvergne.

Stimuli. The entire set of colorized pictures is available for free download at <http://leadserv.u-bourgogne.fr/~lead/openlead>. SV categorized their concepts into 15 semantic categories. We used these SV semantic categories and added more items for certain semantic categories, in order to have a better representation of each category in the new database. We also decided to add exemplars that did not belong to the semantic categories provided by SV. We therefore decided a priori on the different semantic categories that we thought useful to add (e.g., weapons, sports) as well as on the number of exemplars to include in each category (which varies from 4 to 21). Likewise, we had 313 items that were classified into 24 (a priori) semantic categories (see Table 1 for the different semantic categories and the corresponding numbers of items). This information about how we classified our items into these semantic categories (see Supplementary Material A) should help researchers to find specific items more easily when designing experiments with drawings. However, in order to establish whether our a priori classification of the items was reliable, we asked ten independent adults to classify the

1
2
3 picture names among the twenty-four semantic categories (see Table 1). Given the relatively
4
5 large number of categories, an acceptable interrater reliability was observed as indicated by
6
7 the value of .77 obtained for Krippendorff's alpha (this index was computed using the SPSS
8
9 macro developed by Hayes & Krippendorff, 2007). It should be noted that two categories
10
11 were chosen at an equal level for only five words and that the modal category was chosen by
12
13 less than 6 participants for 14 items (see details in the Supplementary Material A).²
14
15

16
17 -----
18
19 Table 1 about here
20
21 -----
22

23
24 The selected SV concepts (216 concepts in common with SV, overlap = 83%),
25
26 together with the added concepts, can be seen in the Supplementary Material A. The drawings
27
28 were designed by a draftsman, graphic artist and illustrator (Bonzai Studio:
29
30 <http://www.bonzaistudio.com/bonzai-studio-dijon-animatin-3d/>), who was instructed to use
31
32 the same graphic style throughout. The drawings were designed with a digital tablet and pen
33
34 set. The same space was used for the whole set of drawings (dimensions 2362 × 2362 pixels).
35
36 Two authors of the paper (PB and BPC) told the artist exactly what they wanted to be
37
38 represented for each concept, assessed the quality of the drawings, and improvements were
39
40 made when needed.
41
42
43

44
45 **Procedure.** To collect norms for name agreement, image agreement, conceptual familiarity,
46
47 and AoA, we closely followed the procedures used in the Alario and Ferrand (1999), Bonin,
48
49 Peereman, et al. (2003), and Tsaparina et al. (2011) studies. The instructions for each norm
50
51 are provided in full in the Supplementary Material B.
52

53
54 For name agreement and image agreement, the pictures were projected on a large
55
56 white screen by means of an overhead projector. For each rating task, the set of pictures was
57
58 presented in a different random order to the different groups of participants. The participants
59
60

1
2
3 were tested collectively in small groups (no more than 8 individuals). In the name-agreement
4 task, the participants were instructed to identify each picture by writing down the first name
5 that came to mind. Whenever they were not able to provide a name for a given picture, they
6 had to indicate the reason, that is to say, (1) they did not know the object (DKO), (2) they did
7 not know the name of the object (DKN), or (3) they were in a tip-of-the-tongue state (TOT).
8 The name-agreement task was the first task to be conducted in order to compute name
9 agreement scores that then allowed us to select the modal names to be used in the normative
10 measures made on the basis of the object labels (e.g., AoA).
11
12
13
14
15
16
17
18
19
20

21 To collect image-agreement ratings, the participants were asked to first generate a
22 mental image from a name that was presented for one second. The modal names
23 corresponding to the pictures were displayed in Times New Roman, font size 40, on a large
24 white screen prior to the presentation of the pictures. There was then a 5-second interval
25 during which the adults had to generate a mental image corresponding to the name either,
26 while keeping their eyes closed or while staring at the screen. As soon as the picture was
27 displayed on the screen, the participants had to rate the degree of agreement between the
28 presented picture and their generated mental image on a 5-point Likert scale. A rating of 1
29 indicated "low agreement" and a rating of 5 indicated "high agreement".
30
31
32
33
34
35
36
37
38
39
40
41

42 For the conceptual familiarity ratings, the participants had to rate on a 5-point Likert
43 scale (1 = low familiarity; 5 = high familiarity) "how usual or unusual the depicted object was
44 in their realm of experience". More precisely, they were told that familiarity corresponds to
45 "the degree to which you come into contact with or think about the concept". The participants
46 were instructed to rate the concept itself, and not the way it was drawn.
47
48
49
50
51
52

53 The ratings for AoA and imageability were collected using LimeSurvey from the
54 names assigned to the pictures. In the AoA task, the participants had to estimate the age at
55 which they thought they had learned each of the names in its written or oral form. A 5-point
56
57
58
59
60

Likert scale was used with 1 = “learned between 0-3 years” and 5 = “learned at age 12 or after”. Concerning imageability, based on the instructions provided by Paivio, Yuille, and Madigan (1968) and their French translation (see Bonin, Méot, Ferrand, & Roux, 2011), adults had to rate on a 5-point Likert scale the ease of generating a mental image from the picture names (from 1 = difficult to 5 = easy).

Objective measures

Digitized images for all 313 pictures were used to calculate the objective visual complexity scores, using image file size metrics. Following Székely and Bates (2000), the size of the image files in JPEG format was used as an objective measure of visual complexity. To avoid report errors when assessing the number of bytes of each picture, the procedure was automated using a batch script running on Windows.

Objective word frequency was taken from Lexique.org (New et al., 2004) and, more precisely, we used subtitle frequency and book frequency measures of the modal names. Length was measured as the number of letters.

Results

In the Supplementary Material C, the two measures of name agreement corresponding to the percentage of participants giving the most common name (the %NA) and the H statistic (as described in SV, 1980) are provided. In the Supplementary Material D, the different nonmodal names are listed for each item together with their corresponding frequencies of occurrence. We also indicate DNO, DKN, and TOT responses.

In the Supplementary Material C, means and standard deviations for image agreement, conceptual familiarity, imageability, and rated AoA are provided for the items listed alphabetically according to the French names of the pictures (English translations are provided). The values corresponding to the visual complexity of the pictures are reported.

Objective word frequency values and the number of letters of the modal names are also provided.

Reliability

One participant provided unusual details when naming the pictures and her/his data was therefore discarded from the name-agreement task. Another participant, who rated 88% of the pictures as unfamiliar, i.e., score of 1, was excluded from the familiarity-rating task.

Table 2 about here

Table 2 shows the intraclass correlation coefficients (random effects of both participants and items - ICC(2, k) in Shrout and Fleiss [1979]'s terminology) for the different norms obtained with the use of Likert scales. Table 2 also reports the correlations between the percentages of name agreement scores—the details concerning the response coding are given below—obtained from even and odd participants for each list of words³. The reliability scores turned out to be high for name agreement. With ICC values above .80, reliability was high for *all* the collected norms.

In order to further gauge the reliability of our ratings for AoA and imageability, we correlated these ratings with the ratings taken from previous French studies (Alario & Ferrand, 1999; Bonin, Méot et al., 2003; Bonin et al., 2011; Bonin, Peereman et al., 2003; Ferrand et al., 2008). The AoA ratings in the present study were highly correlated with the ratings taken from other French studies (see Table 3). However, the correlations were lower for imageability. These lower correlations are possibly due to the fact that a lot of words are highly imageable.

 Table 3 about here

Name agreement comparisons with the MultiPic and Rossion and Pourtois databases

Two authors of the paper (HLD and AM) checked the answers for spelling errors, corrected them, and recoded initial responses by merging basic variants of the same name (e.g., hyphenated or pluralized forms). For any given expected name, abbreviations (e.g., *télé* for *télévision*), word additions (e.g., *chapiteau de cirque* for *chapiteau* or *nœud rose* for *nœud*), deletions (e.g., *raquette* for *raquette de tennis*), and synonyms (e.g., *autocar* or *car* for *bus*) were counted as different responses.

 Figures 1A and 1B about here

Two measures of name agreement were computed: percentage name agreement and the H statistic. For each item, the percentage name agreement (%NA) was calculated as the proportion of the modal name across all naming outcomes, including naming failures (DKO, DKN, and TOT). The H statistic⁴ captures more information than %NA about the distribution of names across participants. An H value of 0 indicates that the drawing elicited the same naming outcome from all participants. Increasing H values indicate decreasing levels of name agreement (H can vary from 0 in the case of perfect agreement to $\log_2(N)$ when all participants produce a different name, where N corresponds to the total number of participants giving name responses to the picture).

Figure 1A shows the distributions for the 200 words common to the current database (IMABASE) and the Rossion and Pourtois (2004) picture database (see Table 1A for the descriptive statistics linked to Figure 1A in the Supplementary Material E). It is important to

1
2
3 note that, given the serious concerns about the name agreement scores provided by Rossion
4 and Pourtois, as set out in Tsaparina et al.'s (2011) study, the comparisons were made with the
5 name agreement scores that we (Bonin, Guillemard-Tsaparina, & Méot, 2013) collected in
6 French for the Rossion and Pourtois set of colorized SV pictures. The distribution of the name
7 agreement scores for the IMABASE items is somewhat more shifted to the right than the
8 Rossion and Pourtois items (difference in confidence interval of the 95% bootstrapped means:
9 [1.29; 4.82]; Wilcoxon signed rank test: $Z = 4.05, p < .001$).

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1B (see Table 1B for the descriptive statistics linked to Figure 1B in the Supplementary Material E) shows the same information for the 263 words common to MultiPic (Duñabeitia et al., 2018). It is worth noting that for 19 modal names in the MultiPic database, which corresponded to two or even three pictures from this database, we decided to select the picture having the highest %NA. The differences in the distribution of name agreement scores between IMABASE and MultiPic are very tenuous (difference in confidence interval of the 95% bootstrapped means: [-1.67; 1.35]; Wilcoxon signed rank test: $Z = -.39, p = .696$).

It must be stressed, however, that the criteria used by Duñabeitia et al. (2018) to classify the naming responses were not exactly the same as those used in the present work. Duñabeitia et al. noted that the initial responses were recoded “by [...] discarding parts of speech that were not nouns or verbs (e.g., determiners and adjectives)” and that “trials where participants did not know the name of the concept (2.4% of the data across languages) and idiosyncratic responses (i.e., responses provided by a single participant, corresponding to 2.7% of the data) were excluded”⁵.

As far as the items from our database that are common to the other two databases are concerned, the percentages of items with a 100% name agreement score were 40.2% (IMABASE) vs. 41.3% (MultiPic), and 42% (IMABASE) vs. 30% (Rossion & Pourtois).

Naming failures (including DKO, DKN, TOT, and non-responses) in our database were somewhat lower than those found in MultiPic when the entire databases were taken into account (1.1% [IMABASE] vs. 2.7% [MultiPic], but they were virtually identical for the shared modal names (1% [IMABASE] vs. 1.1% [MultiPic]) when only taking account of MultiPic pictures yielding the modal names with the highest name-agreement scores, see above). Idiosyncratic responses were comparable in both cases (2.3% [IMABASE] vs. 2.8% [MultiPic] for the entire databases, and 2.2% [IMABASE] vs. 1.6% [MultiPic] for the shared modal names). Naming failures and idiosyncratic responses were also very similar between the current database and the data provided by Bonin, Guillemard-Tsaparina, et al. (2013) for the Rossion and Pourtois drawings: naming failures (1.6%) and idiosyncratic responses (3.6%) for the entire databases; 1.2% and 2% respectively, for the words in common.

Picture-choice task

As explained above, in order to explore more deeply certain differences between the three picture databases, we included a picture-choice task. For each (modal) name in our database, the participants were presented with the different pictures that were available in the three databases, and they had to choose the one they thought best matched the object referred to by the name.

 Table 4 about here

As can be seen from Table 4, a large majority of the pictures taken from our database were chosen as the best match (about 75% of choices), with the MultiPic pictures being the second preference (around 20% of all choices). The Rossion and Pourtois pictures were rarely chosen.

Descriptive statistics

Table 5 about here

Table 5 provides descriptive statistics for the collected ratings as well as additional characteristics of the items. The distributions corresponding to the collected norms are shown in Figure 2.

Figure 2 about here

Percentage name agreement and H were highly negatively and positively skewed, respectively, indicating high levels of agreement between the names given by the participants to the pictures. Even if it was less pronounced, a negative skew was also found for the image agreement variable, which conversely shows that the participants generally agreed with the pictures proposed for the modal names.

With both mean and median roughly equal to 4.5, most of the words were highly imageable, resulting in a high negative skew for imageability ratings. As said earlier, this property is probably the reason why there are relatively low levels of correlations between the current imageability scores and other scores taken from previous published databases.

Most of the words were also judged to be acquired early in life (more than 50% before 6 years of age), whereas only very few (6.1%, $N = 19$) were rated to be acquired after 10 years of age. Finally, the familiarity ratings were distributed along the entire range of the scale except at the extremes. There were two modes: one for very familiar concepts (ratings around 4) and one for relatively unfamiliar concepts (ratings around 2).

Bivariate correlations and multiple regressions with behavioral variables as dependent variables

Table 6 about here

As shown by the bivariate correlations (Table 6), the correlations between the two name agreement variables and the other psycholinguistic variables were generally low, with higher agreement being obtained for more imageable and familiar concepts and early acquired names. Name agreement also increased with image agreement. Low correlations were found for image agreement, which—in addition to name agreement—was significantly correlated only with objective word frequency, and, to a lesser extent, with imageability, with the result, that image agreement was higher for less frequent and more imageable words. Imageability was related to all the other variables, and more particularly to familiarity and AoA, with highly imageable items being judged to be more familiar and acquired earlier in life. Highly-imageable items also had higher name agreement scores and, to a lesser extent, were more objectively frequent and longer. Early acquired words were found to be more imageable, conceptually familiar, frequent, and shorter than late acquired words. Finally, the more imageable, frequent, and early acquired the items were, the more conceptually familiar they were judged to be. Moreover, but to a lesser extent, objectively complex pictures and shorter names were also judged to be more familiar.

Overall, most of the correlations are comparable with those reported in other normative studies on colorized drawings (Bakhtiar et al., 2013; Dimitropoulou et al., 2009; Duñabeitia et al., 2018; Raman et al., 2014; Rossion & Pourtois, 2004 with the name agreement scores taken from Bonin, Guillemard-Tsaparina, et al., 2013; Tsaparina et al., 2011). One notable exception is found in the visual complexity variable, for which the correlation with the other subjective ratings were systematically lower (and generally not significant) in the present database than those found in most normative studies on pictures.

1
2
3 This finding is most probably due to the fact that we used objective visual complexity scores
4
5 in the present work, and not subjective measures of visual complexity. Another noticeable
6
7 discrepancy concerns the significant negative correlation found between image agreement and
8
9 objective word frequency. No such correlation was found in the three studies (Bakhtiar et al.,
10
11 2013; Raman et al., 2014; Rossion & Pourtois, 2004 with the name agreement scores taken
12
13 from Bonin, Guillemard-Tsaparina, et al., 2013; Tsaparina et al., 2011) that reported this
14
15 correlation (with, however, nonsignificant results in two of them). Finally, we did not find a
16
17 correlation between name agreement and familiarity, whereas such a correlation has generally
18
19 been found in other studies (Bakhtiar et al., 2013; Raman et al., 2014; Rossion & Pourtois,
20
21 2004).

22
23
24
25
26 In order to study if the above relationships persisted when other dimensions were
27
28 controlled for, we included each psycholinguistic variable in turn as the dependent variable
29
30 and the other variables as independent variables in multiple regression analyses. Given the
31
32 high correlation found between the two name-agreement measures, only percentage name
33
34 agreement was retained for the analyses. In order to make comparisons easier, all the
35
36 independent variables (IVs) were transformed into Z-scores. Nonlinearities were introduced
37
38 by including restricted cubic splines with three knots for some independent variables using a
39
40 forward approach: At each step, nonlinear terms were included for the independent variables
41
42 whose inclusion led to the greatest increase in variance compared with the model that did not
43
44 include such terms. We opted for a minimum of 1% of added variance, and nonlinearity was
45
46 not included if the change between the two models was not significant. The results of these
47
48 analyses are given in Table 7 and Figure 3.

53
54 -----
55
56 Table 7 and Figure 3 about here
57
58 -----
59
60

1
2
3 The partial relationships were generally similar to the bivariate relationships.
4
5 Noticeable differences were as follows. First of all, three relatively important relations that
6
7 were found in the bivariate correlations vanished when other IVs were controlled for: (a) high
8
9 name agreement was no longer related to early acquisition; (b) length did not influence AoA
10
11 ratings (the same was observed for familiarity with, however, a lower bivariate correlation);
12
13 and (c) word frequency did not have a positive effect on imageability. Second, other relations
14
15 turned out to be significant: (a) name agreement decreased with higher familiarity (and the
16
17 reverse); (b) image agreement was influenced by AoA in a nonlinear relationship in such way
18
19 that it decreased slowly, with higher AoA ratings being observed below the mean, before
20
21 increasing considerably; (c) the positive relation between image agreement and imageability
22
23 was more pronounced. Third, nonlinear relations indicated that (a) name agreement increased
24
25 with imageability ratings below the mean, and then tended to stabilize; (b) familiarity
26
27 increased only for values situated at the right of the word frequency values, whereas the
28
29 reverse was true for its relation with visual complexity; (c) imageability was stable for early
30
31 acquired words and then decreased greatly.
32
33
34
35
36
37
38
39

40 General Discussion

41
42 The collection of psycholinguistic norms for various types of stimuli (e.g., faces of
43
44 celebrities: Bonin, Perret, Méot, Ferrand, & Mermillod, 2008; Marful, Díez-Álamo, Plaza-
45
46 Navas, & Fernandez, 2018; acronyms: Bonin, Méot, Millotte, & Bugajska, 2015; Izura, &
47
48 Playfoot, 2012; idioms: Bonin, Méot, Boucheix, & Bugajska, 2018; Bonin, Méot, &
49
50 Bugajska, 2013) is now a frequent practice in psychology. In the present research work, we
51
52 provided a new set of 313 colorized drawings, which have been normed on five important
53
54 psycholinguistic variables: name agreement, image agreement, conceptual familiarity, AoA,
55
56 and imageability. Importantly, all the collected norms turned out to be highly reliable.
57
58
59
60

1
2
3 Overall, the pictures in our database are consensual since there are high levels of agreement
4
5 between the names given by the participants and the drawings. Most of the object names are
6
7 highly imageable and acquired early in life, and the concepts are mostly familiar. Finally, the
8
9 pattern of correlations among the norms is generally similar to those found in other similar
10
11 studies (Bakhtiar et al., 2013; Dimitropoulou et al., 2009; Duñabeitia et al., 2018; Raman et
12
13 al., 2014; Rossion & Pourtois, 2004 with the name agreement scores taken from Bonin,
14
15 Guillemard-Tsaparina, et al., 2013; Tsaparina et al., 2011).

16
17 From a general standpoint, having norms for stimuli (1) permits the
18
19 control—methodological or statistical—of the potential influence of confounding variables
20
21 when investigating a specific factor in a given lexical processing task, such as object naming;
22
23 (2) helps to establish the underlying structure of the norms and to better understand what
24
25 exactly they measure, and (3) last but not least, permits the investigation of their relationships
26
27 with on-line (or off-line) measures of word processing such as object naming (e.g., Bonin et
28
29 al., 2002; Bonin, Guillemard-Tsaparina, et al., 2013; Bonin, Peereman, et al., 2003). Many
30
31 studies have attempted to identify the determinants of naming speed in adults, and this has
32
33 been made possible thanks to the availability of norms for pictures in different languages.
34
35
36
37
38
39

40
41 As claimed earlier, there is a limited range of options for researchers who want to use
42
43 *colorized drawings* in their experiments (Duñabeitia et al., 2018). Even though photographs
44
45 of objects have been thought to be more ecological than line drawings (Moreno-Martínez &
46
47 Montoro, 2012), they generally elicit less consensus on naming responses than photographs
48
49 (Clarke & Ludington, 2018). Consequently, line drawings are to be preferred in the field of
50
51 speeded object naming because they limit the number of trials that have to be discarded when
52
53 analyzing naming times. As pointed out by Clarke and Ludington, name agreement with line
54
55 drawings is generally above 80%, with H values located between .20 and .80. It is worth
56
57 remembering that for our picture set, the means of name agreement and H were, respectively,
58
59
60

1
2
3 89.64% and .40. Given the finding that the distribution of name agreement scores was very
4 similar between our database and MultiPic, researchers will be able to select pictures from
5 these two databases and combined them. However, as the picture-choice task suggests,
6
7 whenever researchers have to choose between two pictures in order to represent one concept,
8 they should select the one from IMABASE rather than the corresponding picture from
9
10 MultiPic given that the great majority of the IMABASE pictures that were common to
11
12 MultiPic were rated as providing better pictorial representations of the corresponding
13
14 concepts. Turning to memory experiments, the format of pictures seems to be a lesser concern
15
16 (Snow et al., 2014). However, our set of drawings will also be useful for the design of
17
18 memory experiments because, as evaluated by the picture-choice task, the concepts depicted
19
20 by the drawings were very easy to recognize. For example, pictures are used in memory
21
22 experiments in children. To give an example, Fitamen et al. (2019) used pictures from the
23
24 Rossion and Pourtois database to investigate memory processes in 4- and 5-year-old
25
26 preschoolers. For this type of population, it is especially important to have drawings that are
27
28 easy to recognize (and indeed in Fitamen et al.'s study, any given picture had to have more
29
30 than 80% correct denominations at 4 years of age to be selected). Thus, our drawings could be
31
32 very helpful for memory researchers working on children.
33
34
35
36
37
38
39
40
41

42 The colorized version of SV has long been a good option for the design of experiments
43 aimed at exploring perceptual, language or memory processes in both children and adults
44 (e.g., Bonin, Guillemard-Tsaparina, & Méot, 2013; Bonin, Méot, Laroche, Bugaiska, &
45 Perret, 2019; Dimitropoulou et al., 2009; Fitamen, Blaye, & Camos, 2019; Hachmann,
46 Cashdollar, Postiglione, & Job, 2020; Reis, Faísca, Ingvar, Petersson, 2006). Thus, this new
47 set of colorized line drawings and the corresponding collected norms should help investigate
48 further the determinants of naming speed and accuracy in adults but also in children,
49 teenagers, and older adults. There have been many studies on picture naming in adults but
50
51
52
53
54
55
56
57
58
59
60

1
2
3 fewer in children (e.g., Cychowicz et al., 1997; D'Amico, Devescovi, & Bates, 2001) or in
4
5 older adults (e.g., Au, Joung, Nicholas, Obler, Kass, & Albert, 1995), and we are not aware of
6
7 any study that has specifically focused on the naming speed and accuracy in teenagers.
8
9

10 As we also pointed out in the Introduction, with the passing of time, the SV pictures
11
12 have come to suffer from several limitations: Certain pictures were found not to be “good”
13
14 representations of familiar concepts; certain pictures were obviously linked to American
15
16 culture and were not very suitable for other cultures; certain objects needed to be updated.
17
18 The current database, together with the recent MultiPic database of colorized line drawings,
19
20 will greatly help researchers to select items when designing experiments that use pictures as
21
22 stimuli (since its publication, the MultiPic database has already been used to study lexical
23
24 selection in spoken word production, e.g., Gauvin, Jonen, Choi, McMahon, & de Zubicaray,
25
26 2018). Our database is somewhat reduced in terms of the number of pictures it offers for
27
28 selection compared to MultiPic, but we think that is likely sufficient for many experimental
29
30 designs (remember that the SV database contains only 260 pictures, and it has been widely
31
32 used to design many experiments in several psychology domains). However, and more
33
34 importantly, our database is normed on many more psycholinguistic variables than MultiPic,
35
36 which is normed only on two variables: name agreement and visual complexity (the latter of
37
38 which has been found to be a factor of lesser importance in picture naming than, for instance,
39
40 conceptual familiarity, see Perret and Bonin, 2019). In the future, we hope to be able to
41
42 increase the number of pictures proposed in IMABASE. Recently, “new” psycholinguistic
43
44 variables have been collected for items, such as object manipulability and interestingly, this
45
46 variable has been found to be a significant predictor of naming latency (Guérard, Lagacé, &
47
48 Brodeur, 2015; Lorenzoni, Peressotti, & Navarrete, 2018). Thus, we also hope in the future to
49
50 collect additional psycholinguistic variables for our material. Another limitation of the current
51
52 work is that the norms were collected from French-speaking adults who live in France, and it
53
54
55
56
57
58
59
60

1
2
3 is possible that the set of norms may not be suitable for use with adults speaking French in
4
5 other parts of the world such as Belgium, Switzerland, parts of Africa (e.g., Senegal) or
6
7 Oceania (e.g., New Caledonia). It would be interesting in the future to collect French norms
8
9 on this set of pictures but with such French speakers and to compare them with the current
10
11 norms.
12
13

14
15 Hopefully, as has been the case for the SV picture database, the scientific community
16
17 will adopt this new database and foreign colleagues will think about collecting norms in their
18
19 own native languages. To this end, the pictures are open-access and free from copyright
20
21 restrictions for noncommercial purposes. It is important to point out that IMABASE should
22
23 not be restricted to the investigation of picture naming. Indeed, all fields of research that use
24
25 picture naming as an experimental task should be impacted by the publication of such
26
27 standardized pictures. Likewise, as already said, the database should help researchers who
28
29 investigate learning and memory processes in children and in both young adults and the
30
31 elderly. In these areas of research, pictures are often used as stimuli (Bonin, Gelin, &
32
33 Bugajska, 2014). Standardized pictures are, for instance, used to investigate the development
34
35 of short-term memory coding strategies in children (e.g., Henry, Messer, Luger-Klein, &
36
37 Crane, 2012), or to address the question of the semantic impairment in early dementia of
38
39 Alzheimer type (Garrad, Lambon Ralph, Patterson, Pratt, & Hodges, 2005). In addition, the
40
41 present database could be useful for the investigation of perceptual processes involved in
42
43 visual object recognition which was also the case for the SV database (e.g., Snodgrass &
44
45 Corwin, 1988).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary Material

The Supplementary Material is available at: qjep.sagepub.com

Availability and Open Practices Statement

The IMABASE whole set of pictures is available for public download at <http://leadserv.u-bourgogne.fr/~lead/openlead>. The different files corresponding to the norms are formatted as .xls. None of the studies were preregistered.

References

- 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*, 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*, 140–155.
- Au, R., Joung, P., Nicholas, M., Obler, L. K., Kass, R., & Albert, M. L. (1995). Naming ability across the adult life span. *Aging & Cognition, 2*, 300–311.
- Bakhtiar, M., Nilipour, R., & Weekes, B. S. (2013). Predictors of timed picture naming in Persian. *Behavior Research Methods, 45*, 834–841.
- Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29*, 639–647.
- 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*, 560–585.
- Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10*, 344–380.
- Bonin, P. (2013). *Psychologie du langage. La fabrique des mots*. Bruxelles: DeBoeck.
- Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology, 93*, 89–114.

- 1
2
3 Bonin, P., Gelin, M., & Bugaiska, A. (2014). Animates are better remembered than
4
5 inanimates: Further evidence from word and picture stimuli. *Memory & Cognition*, *42*,
6
7 370–382.
8
9
- 10 Bonin, P., Guillemard-Tsaparina, D., & Méot, A. (2013). Determinants of naming latencies,
11
12 object comprehension times, and new norms for the Russian standardized set of the
13
14 colorized version of the Snodgrass and Vanderwart pictures. *Behavior Research*
15
16 *Methods*, *45*, 731–745.
17
18
- 19 Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M.-C.
20
21 (2003). Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de
22
23 valence émotionnelle pour 866 mots [Concreteness, imageability, subjective frequency
24
25 and emotional valence norms for 866 words]. *L'Année Psychologique*, *104*, 655–694.
26
27
- 28 Bonin, P., Méot, A., Boucheix, J. M., & Bugaiska, A. (2018). Psycholinguistic norms for 320
29
30 fixed expressions (idioms and proverbs) in French. *The Quarterly Journal of*
31
32 *Experimental Psychology*, *71*, 1057–1069.
33
34
- 35 Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French
36
37 idiomatic expression. *Behavior Research Methods*, *45*, 1259–1271.
38
39
- 40 Bonin, P., Méot, A., Ferrand, L., & Roux, S. (2011). L'imageabilité: Normes et relations avec
41
42 d'autres variables psycholinguistiques [Imageability: Norms and relationships with
43
44 other psycholinguistic variables]. *L'Année Psychologique*, *111*, 327–357.
45
46
- 47 Bonin, P., Méot, A., Laroche, B., Bugaiska, A., & Perret, C. (2019). The impact of image
48
49 characteristics on written naming in adults. *Reading & Writing*, *32*, 13-31.
50
51
- 52 Bonin, P., Méot, A., Millotte, S., & Bugaiska, A. (2015). Norms and reading times for
53
54 acronyms in French. *Behavior Research Methods*, *47*, 251–267.
55
56
- 57 Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299
58
59 pictures for psycholinguistic studies: French norms for name agreement, image
60

- 1
2
3 agreement, conceptual familiarity, visual complexity, image variability, age of
4 acquisition, and naming latencies. *Behavior Research Methods, Instruments, &*
5
6
7
8
9
10 Bonin, P., Perret, C., Méot, A., Ferrand, L., & Mermillod, M. (2008). Psycholinguistic norms
11 and face naming times for photographs of celebrities in French. *Behavior Research*
12
13
14
15
16
17 Boukadi, M., Zouaidi, C., & Wilson, M. A. (2016). Norms for name agreement, familiarity,
18 subjective frequency, and imageability for 348 object names in Tunisian Arabic.
19
20
21
22
23
24 Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of
25 standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used
26
27
28
29
30 as visual stimuli in cognitive research. *PLoS ONE*, 5, e10773.
31
32
33
34
35 Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS)
36 Phase II: 930 new normative photos. *PLoS ONE*, 9, e106953.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology*, 64, 545–559.
- Brysbaert, M., Mandra, P., & Keuleers, E. (2018). The word frequency effect in word processing: A review update. *Current Directions in Psychological Science*, 27, 45-50.
- Bueno, S., & Megherbi, H. (2009). French categorization norms for 70 semantic categories and comparison with Van Overschelde et al.'s (2004) English norms. *Behavior Research Methods*, 41, 1018–1028.
- Clarke, A. J. B., & Ludington, J. D. (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*,

- 1
2
3 47, 607–626.
4
5 Cortese, M. J., Simpson, G. B., & Woolsey, S. (1997). Effects of association and imageability
6
7 on phonological mapping. *Psychonomic Bulletin & Review*, 4, 226–231.
8
9
10 Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish:
11
12 Predictors and errors. *Brain and Language*, 82, 344–365.
13
14 Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and
15
16 Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, &*
17
18 *Computers*, 31, 650–658.
19
20
21 Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by
22
23 young children: Norms for name agreement, familiarity, and visual complexity.
24
25 *Journal of Experimental Child Psychology*, 65, 171–237.
26
27
28 D'Amico, S., Devescovi, A., & Bates, E. (2001). Picture-naming and lexical access in Italian
29
30 children and adults. *Journal of Cognition and Development*, 2, 71–105.
31
32
33 Dimitropoulou, M., Duñabeitia, J. A., Blitsas, P., & Carreiras, M. (2009). A standardized set
34
35 of 260 pictures for Modern Greek: Norms for name agreement, age of acquisition, and
36
37 visual complexity. *Behavior Research Methods*, 41, 584–589.
38
39
40 Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., &
41
42 Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six
43
44 European languages. *The Quarterly Journal of Experimental Psychology*, 71, 808–
45
46 816.
47
48
49 Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval.
50
51 *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 515–523.
52
53
54 Ferrand, L., Bonin, P., Méot, A., Augustinova, M., New, B., Pallier, C., & Brysbaert, M.
55
56 (2008). Age of acquisition and subjective frequency estimates for all generally known
57
58
59
60

- monosyllabic French words and their relation with other psycholinguistic variables. *Behavior Research Methods*, *40*, 1049–1054.
- Fitamen, C., Blaye, A., & Camos, V. (2019). Five-year-old children's working memory can be improved when children act on a transparent goal cue. *Scientific Reports*, *9*:15342.
- Forsythe, A. (2009). Visual complexity: Is that all there is? In D. Harris (Ed.), *Proceedings of the 8th International Conference on Engineering Psychology and Cognitive Ergonomics*. Springer: Heidelberg, 158–166.
- Forsythe, A., Street, N., & Helmy, M. (2017). Revisiting Rossion and Pourtois with new ratings for automated complexity, familiarity, beauty, and encounter. *Behavior Research Methods*, *49*, 1484–1493.
- Garrad, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language*, *93*, 79–94.
- Gauvin, H. S., Jonen, M. K., Choi, J., McMahan, K., de Zubicaray, G. I. (2018). No lexical competition without priming: Evidence from the picture–word interference paradigm. *The Quarterly Journal of Experimental Psychology*, *71*, 2562–2570.
- 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*, 443–470.
- Hachmann, W. M., Cashdollar, N., Postiglione, F., & Job, R. (2020). The relationship of domain-general serial order memory and reading ability in school children with and without dyslexia. *Journal of Experimental Child Psychology*, *193*, 104789.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*, 77–89.

- 1
2
3 Henry, L. A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual, and
4 semantic coding strategies and children's short-term picture memory span. *The*
5 *Quarterly Journal of Experimental Psychology*, *65*, 2033–2053.
6
7
8 Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture
9 identification. *Cognitive Neuropsychology*, *5*, 67–103.
10
11
12 Izura, C., & Playfoot, D. (2012). A normative study of acronyms and acronym naming.
13 *Behavior Research Methods*, *44*, 862–889.
14
15
16 Johnston, R., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual*
17 *Cognition*, *13*, 789–845.
18
19
20 Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and
21 naming times for a set of 539 pictures: The role of age of acquisition. *Behavior*
22 *Research Methods*, *42*, 461–469.
23
24
25 Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification.
26 *Psychological Bulletin*, *131*, 684–712.
27
28
29 Khwaileh, T., Mustafawi, E., Herbert, R., & Howard, D. (2018). Gulf Arabic nouns and
30 verbs: A standardized set of 319 object pictures and 141 action pictures, with
31 predictors of naming latencies. *Behavior Research Methods*, *50*, 2408-2425.
32
33
34 Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech
35 production. *Behavioral and Brain Sciences*, *22*, 1–75.
36
37
38 Lorenzoni, A., Peressotti, F., & Navarrete, E. (2018). The manipulability effect in object
39 naming. *Journal of Cognition*, *1*, Article 30.
40
41
42 Marful, A., Díez-Álamo, A. M., Plaza-Navas, S., & Fernandez, A. (2018). A normative study
43 for photographs of celebrities in Spain. *PLoS ONE*, *13*, e0197554.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Martein, R. (1995). Norms for name agreement and concept agreement, familiarity, visual
4 complexity and image agreement on a set of 216 pictures. *Psychologica Belgica*, *35*,
5 205–225.
6
7
8
9
- 10 Moreno-Martínez, F. J., & Montoro, P. R. (2012). An ecological alternative to Snodgrass &
11 Vanderwart: 360 high quality colour images with norms for seven psycholinguistic
12 variables. *PLoS ONE*, *7*, e37527.
13
14
15
- 16 New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical
17 database. *Behavior Research Methods, Instruments, & Computers*, *36*, 516–524.
18
19
- 20 Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative
21 set of 359 pictures. *Behavior Research Methods*, *37*, 398–416.
22
23
24
- 25 O’Sullivan, M., Lepage, M., Bouras, M., Montreuil, T., & Brodeur, M. B. (2012). North-
26 American norms for name disagreement: Pictorial stimuli naming discrepancies. *PLoS*
27 *ONE*, *7*, e47802.
28
29
30
31
- 32 Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and
33 meaningfulness values for 925 nouns. *Journal of Experimental Psychology*
34 *Monographs*, *76*, 1–25.
35
36
37
38
- 39 Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture
40 naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*, 2533-
41 2545.
42
43
44
45
46
- 47 Pind, J., Jónsdóttir, H., Tryggvadóttir, H. B., & Jónsson, F. (2000). Icelandic norms for the
48 Snodgrass and Vanderwart (1980) pictures: Name and image agreement, familiarity,
49 and age of acquisition. *Scandinavian Journal of Psychology*, *41*, 41–48.
50
51
52
- 53 Pompéia, S., Miranda, M. C., & Bueno, O. F. A. (2001). A set of 400 pictures standardised
54 for Portuguese: Norms for name agreement, familiarity and visual complexity for
55 children and adults. *Arquivos de Neuro-Psiquiatria*, *59(2-B)*, 330–337.
56
57
58
59
60

- 1
2
3 Raman, I., Raman, E., & Mertan, B. (2014). A standardized set of 260 pictures for Turkish:
4
5 Norms of name and image agreement, age of acquisition, visual complexity, and
6
7 conceptual familiarity. *Behavior Research Methods*, *46*, 588–595.
8
9
- 10 Reis, A., Fáisca, L., Ingvar, M., & Petersson, K. M. (2006). Color makes a difference: Two-
11
12 dimensional object naming in literate and illiterate subjects. *Brain and Cognition*, *60*,
13
14 49–54.
15
16
- 17 Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. In P. Hagoort (Ed.),
18
19 *Human language: From genes and brains to behavior* (pp. 35-50). MIT Press.
20
21
- 22 Rogić, M., Jerončić, A., Bošnjak, M., Sedlar, A., Hren, D., & Deletis, V. (2013). A visual
23
24 object naming task standardized for the Croatian language: A tool for research and
25
26 clinical practice. *Behavior Research Methods*, *45*, 1144–1158.
27
28
- 29 Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart 's object pictorial
30
31 set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–
32
33 236.
34
35
- 36 Russo, N., Hagmann, C. E., Andrews, R., Black, C., Silberman, M., & Shea, N. (2018).
37
38 Validation of the C.A.R.E. stimulus set of 640 animal pictures: Name agreement and
39
40 quality ratings. *PLoS ONE*, *13*, e0192906.
41
42
- 43 Salmon, J. P., McMullen, P. A., & Filliter, J. H. (2010). Norms for two types of
44
45 manipulability (graspability and functional usage), familiarity, and age of acquisition
46
47 for 320 photographs of objects. *Behavior Research Methods*, *42*, 82–95.
48
49
- 50 Sanfeliù, M. C., & Fernandez, A. (1996). A set of 254 Snodgrass-Vanderwart pictures
51
52 standardized for Spanish: Norms for name agreement, image agreement, familiarity,
53
54 and visual complexity. *Behavior Research Methods, Instruments, & Computers*, *28*,
55
56 537–555.
57
58
59
60

- 1
2
3 Saryazdi, R., Bannon, J., Rodrigues, A., Klammer, C., & Chambers C. G. (2018). Picture
4
5 perfect: A stimulus set of 225 pairs of matched clipart and photographic images
6
7 normed by Mechanical Turk and laboratory participants. *Behavior Research Methods*,
8
9
10 50, 2498–2510.
11
- 12 Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability.
13
14 *Psychological Bulletin*, 86, 420–428.
15
- 16 Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory:
17
18 Applications to dementia and amnesia. *Journal of Experimental Psychology: General*,
19
20 117, 34–50.
21
22
- 23 Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for
24
25 name agreement, image agreement, familiarity, and visual complexity. *Journal of*
26
27 *Experimental Psychology: Human Learning and Memory*, 6, 174–215.
28
29
- 30 Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart
31
32 pictures. *Behavior Research Methods, Instruments, & Computers*, 28, 516–536.
33
34
- 35 Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are
36
37 more memorable than photographs of objects. *Frontiers in Human Neuroscience*, 8,
38
39 837.
40
41
- 42 Strain, E., Patterson, K. E., & Seidenberg, M. S. (1995). Semantic effects in single-word
43
44 naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21,
45
46 1140–1154.
47
48
- 49 Strain, E., Patterson, K. E., & Seidenberg, M. S. (2002). Theories of word naming interact
50
51 with spelling-sound consistency. *Journal of Experimental Psychology: Learning*,
52
53 *Memory, and Cognition*, 28, 207–214.
54
55
- 56 Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of
57
58 picture naming. *CRL Newsletter*, 12, 3–33.
59
60

- 1
2
3 Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al. (2005).
4
5 Timed action and object naming. *Cortex*, *41*, 7–26.
6
7 Székely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., (...),
8
9 Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of*
10
11 *Memory and Language*, *51*, 247–250.
12
13 Tsaparina, D., Bonin, P., & Méot, A. (2011). Russian norms for name agreement, image
14
15 agreement for the colorized version of the Snodgrass and Vanderwart pictures and age
16
17 of acquisition, conceptual familiarity, and imageability scores for modal object names.
18
19 *Behavior Research Methods*, *43*, 1085–1099.
20
21
22 Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors
23
24 in picture naming: A trial-by-trial multiple regression analysis from stimulus onset to
25
26 response. *Frontiers in Neuroscience*, *8*, 390.
27
28
29 Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The*
30
31 *Quarterly Journal of Experimental Psychology*, *48A*, 822–848.
32
33
34 Weekes, B. S., Shu, H., Hao, M., Liu, Y., & Tan, L. H. (2007). Predictors of timed picture
35
36 naming in Chinese. *Behavior Research Methods*, *39*, 335–342.
37
38
39 Yoon, C., Feinberg, F., Luo, T., Hedden, T., Gutchess, A. H., Chen, H. Y. M., et al. (2004). A
40
41 cross-culturally standardized set of pictures for younger and older adults: American
42
43 and Chinese norms for name agreement, concept agreement, and familiarity. *Behavior*
44
45 *Research Methods, Instruments, & Computers*, *36*, 639–649.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Authors' notes

The authors wish to thank Patrick Borkowski for his availability and for having provided us with several versions of many drawings in order for us to choose the most suitable for the different concepts present in IMABASE. The authors also wish to thank Manuel Perea and two anonymous reviewers for helpful comments on a previous version of the manuscript.

Peer Review Version

Footnotes

1
2
3
4
5
6
7
8 1. Certain studies have shown that subjective visual complexity ratings (which are often used
9 as a dimension of visual complexiy) are influenced by the familiarity with the objects (e.g.,
10 Forsythe, 2009; Forsythe, Street, & Helmy, 2016). Following Székely and Bates (2000), we
11 think that using objective visual complexity is a way to attempt to have a "less contaminated"
12 measure of visual complexity. However, we acknowledge that, perhaps, it could be useful in
13 the future to collect subjective visual complexity norms on our drawings because these could
14 be (1) directly compared with the objective ones and (2) used in a "complementary manner"
15 to the objective ones (as far as objective word frequency measures are concerned, a similar
16 debate has arisen concerning the importance of subjective word frequency measures in
17 comparison with objective word frequency measures, e.g., Balota, Pilotti, & Cortese, 2001;
18 Brysbaert & Cortese, 2011).

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33 2. We compared, a posteriori, our semantic categories with the semantic categories available
34 for French in a previous study (Bueno & Megherni, 2009). Using a stringent criterium, i.e.,
35 use of exactly the same category name, for the 265 items common to Bueno and Megherni
36 and our own study, the overlap is in the magnitude of 67%. Using a more lenient criterium,
37 that is to say a subcategory is taken as being similar to the main category (e.g., the
38 subcategory "birds" ~ the category "animals"), the overlap amounts to 81%. In addition, it is
39 worth noting that most items belong to several semantic categories. For instance, a hen can be
40 categorized as an animal, or as a bird; it can even be categorized more precisely as a farmyard
41 animal. The different categories provided in the Supplementary Material for our database will
42 make it possible for researchers to select more easily the items they need in order to build
43 their experiments.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 3. Two and six words, from even and odd participants, respectively, were excluded from the
4
5 computations because of discrepancies between the modal names obtained from even and odd
6
7 participants.
8
9

10
11 4. $H = \sum_{i=1}^k p_i \log_2(1/p_i)$ where k refers to the number of different names given for each picture
12
13 and p_i is the proportion of participants giving each name.
14
15

16
17 5. It is worth noting that Snodgrass and Vanderwart (1980) did not exclude “DK items” and
18
19 “idiosyncratic responses”. The latter procedure leads to an overestimation of name agreement
20
21 percentages.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure Captions

Figure 1A. Distributions for words common to IMABASE and Rossion and Pourtois ([2014]; RP) databases (see Table 1A for descriptive statistics linked to this figure in the Supplementary Material)

Figure 1B. Distributions for words common to IMABASE and MultiPic databases (see Table 1B for descriptive statistics linked to this figure in the Supplementary Material)

Figure 2. Distributions of the new norms (y-axis = word percentages)

Figure 3. Partial effects of the significant predictors. The effects of subjective variables are indicated by black symbol curves (%NA, familiarity, image agreement, AoA and imageability), whereas the effects of objective variables (visual complexity, length and lexical frequency) are shown with grey curves. In order to obtain clearer figures, independent variables were z-transformed and values below -2 and above 2 were eliminated. These represented less than 5% of trimmed values except in the case of name agreement [7.9%].

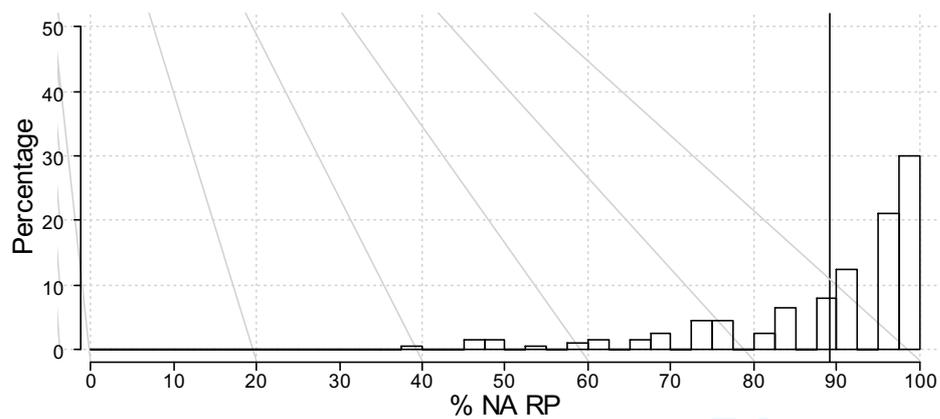
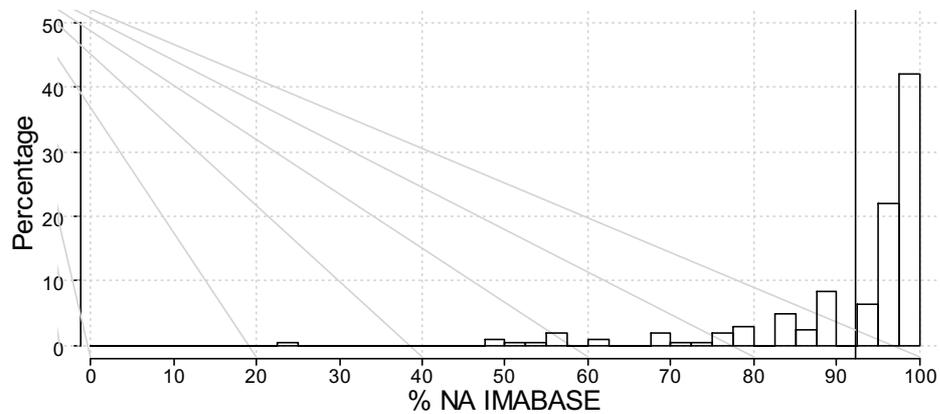


Figure 1A. Distributions for words common to IMABASE and Rossion and Pourtois ([2014]; RP) databases (see Table 1A for descriptive statistics linked to this figure in the Supplemental Material)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

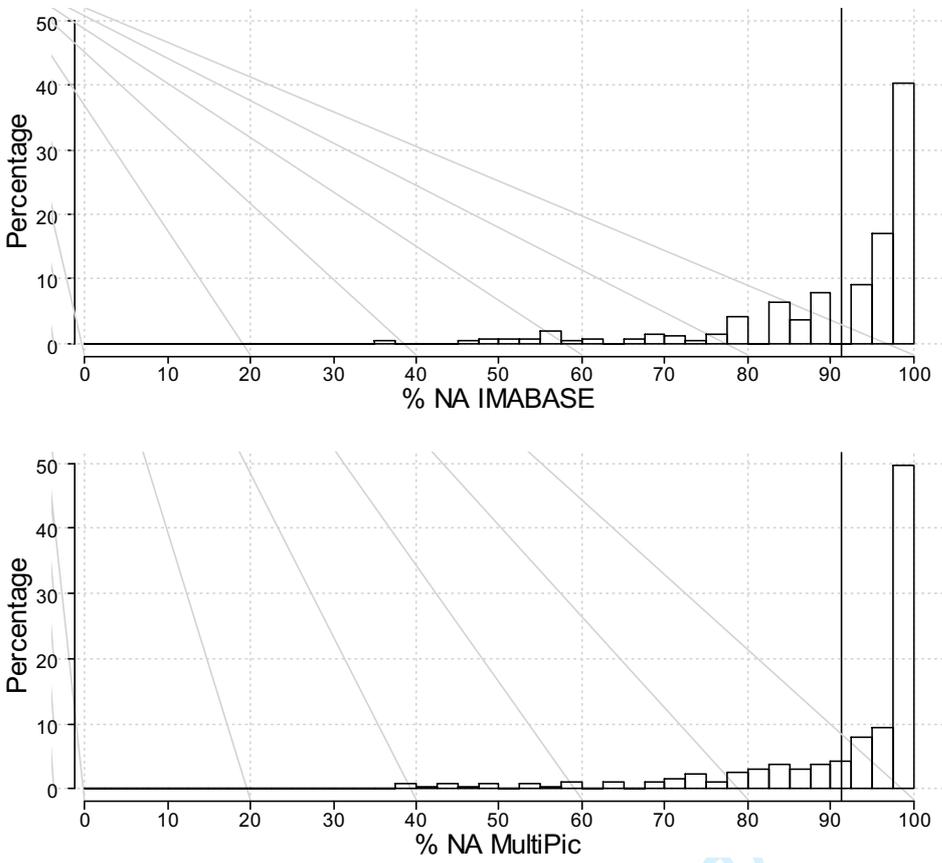


Figure 1B. Distributions for words common to IMABASE and MultiPic databases (see Table 1B for descriptive statistics linked to this figure in the Supplemental Material)

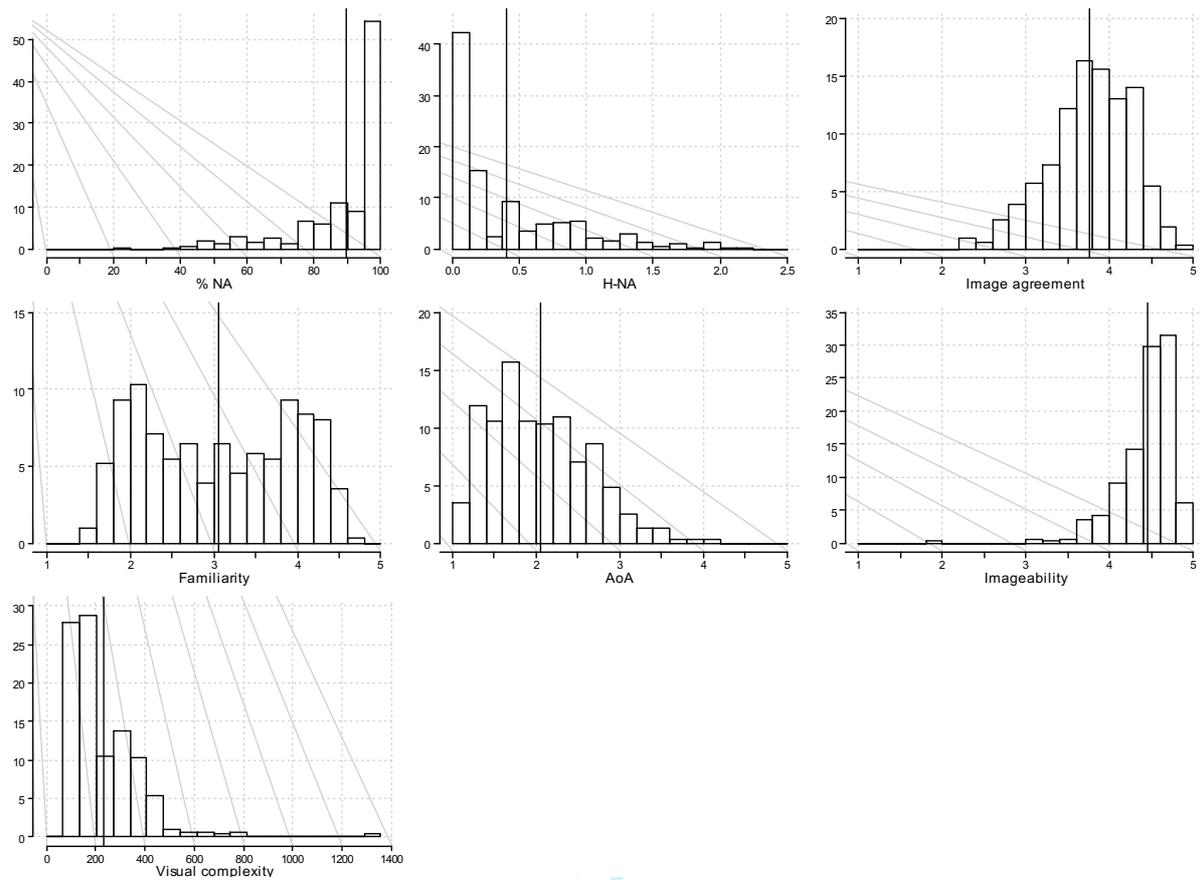


Figure 2. Distributions of the new norms (y-axis = word percentages)

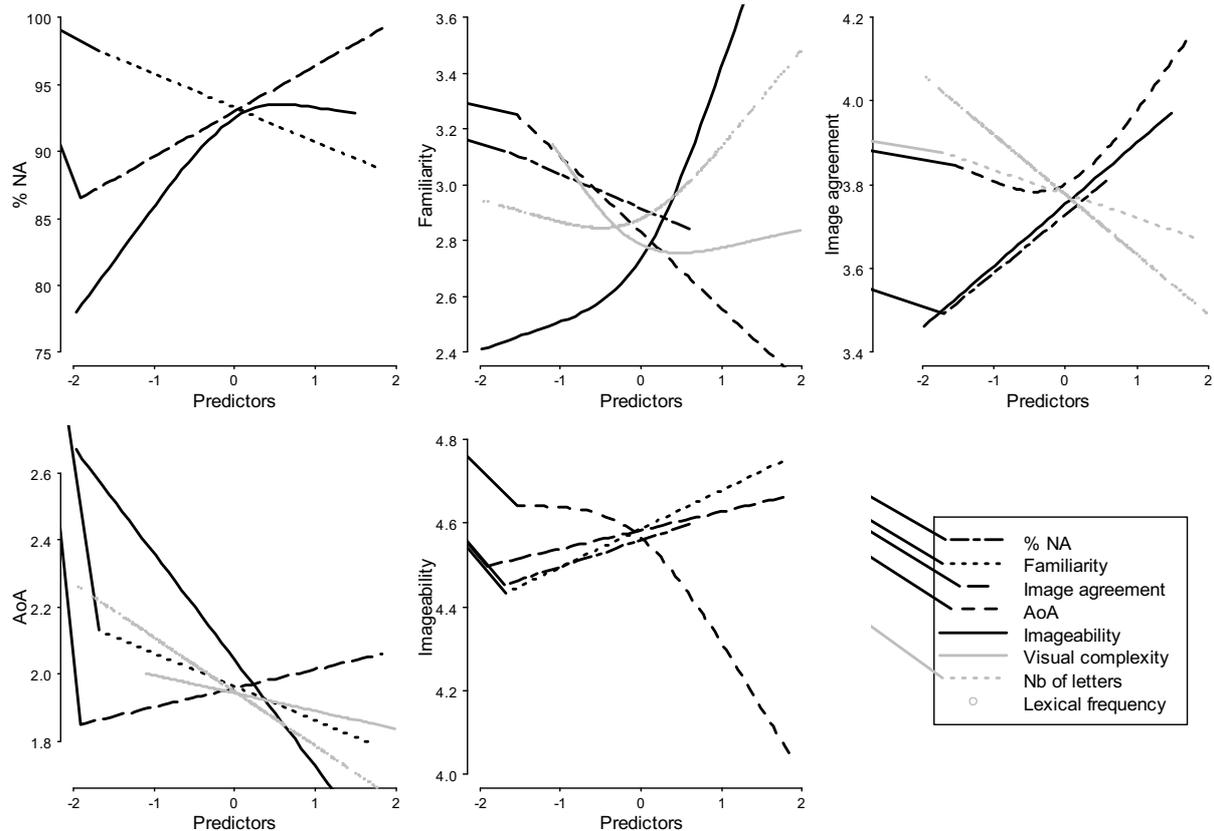


Figure 3. Partial effects of the significant predictors. The effects of subjective variables are indicated by black symbol curves (%NA, familiarity, image agreement, AoA and imageability), whereas the effects of objective variables (visual complexity, length and lexical frequency) are shown with grey curves. In order to obtain clearer figures, independent variables were z-transformed and values below -2 and above 2 were eliminated. These represented less than 5% of trimmed values except in the case of name agreement [7.9%].

Table 1

Semantic categories [French names in brackets] and number of items in each category

Category	Count
Items of furniture [Meubles]	10
School equipment [Fournitures scolaires]	12
Food [Aliments]	16
Aquatic animals [Animaux aquatiques]	12
Farm animal [Animaux de la ferme]	12
<u>Forest and garden animals [Animaux de la forêt et du jardin]</u>	<u>16</u>
<u>Wild animals [Animaux sauvages]</u>	<u>15</u>
<u>Weapons [Armes]</u>	<u>7</u>
Buildings/Parts of buildings [Bâtiments/Parties de bâtiments]	10
<u>Household items [Accessoires de la maison]</u>	<u>19</u>
<u>Birds [Oiseaux]</u>	<u>9</u>
<u>Types of vehicle [Transports/Partie de transports]</u>	<u>13</u>
Make-up [Beauté/Toilette]	9
DIY tools [Bricolage/Jardinage]	15
Parts of the (human) body [Corps humain]	16
Childhood [Enfance]	4
Kitchen utensils [Ustensiles de cuisine]	13
Fruits [Fruits]	12
Articles of clothing [Vêtements]	16
Musical instruments [Instruments de musique]	13
Vegetables [Légumes]	11
Natural elements [Nature/Éléments naturels]	13
Sports/Hobbies [Sports/Loisirs]	21
<u>Other [Autre]</u>	<u>14</u>

Note. Two categories were chosen at an equal level for five words (arc, bougie, cerf, cerf-volant, chapiteau). These words were not included in the count reported in the table (see Supplemental Materials for details).

Table 2

Reliability measures for the norms of age-of-acquisition (AoA), conceptual familiarity, imageability and name agreement (%NA: percentage of name agreement)

Intraclass correlation coefficient				<i>r</i> (even, odd)	
				%NA	
AoA	Conceptual familiarity	Image agreement	Imageability	List 1	List 2
.967	.950	.859	.833	.720	.816

Table 3

Correlations between AoA and imageability scores obtained in the present study and in other studies

	AoA	Imageability
Alario & Ferrand (1999)	.868 (236)	
Bonin, Peereman, et al. (2003)	.929 (32)	
Bonin, Méot, et al. (2003)		.728 (277)
Ferrand et al. (2008)	.923 (99)	
Bonin et al. (2011)		.647 (99)

Note. The numbers of words used in the computations are given in brackets.

Table 4

Results obtained in the picture-choice task

Picture found in	Number and percentage of pictures to be the best match for the object names			Additional information
	IMABASE	RP	MultiPic	
The three databases (N = 178)	138 (77.5%)	10 (5.6%)	28 (15.7%)	1 I = RP > M 1 I = M > RP
IMABASE and MultiPic (N = 85)	66 (77.6%)	–	15 (17.6%)	4 equal (4.7%)
IMABASE and RP (N = 22)	21 (95.5%)	0 (0%)	–	1 equal (4.5%)

Notes. RP = Rossion and Pourtois ; I = IMABASE; M = MultiPic

Table 5

Descriptive statistics of the collected norms

	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	Percentiles			Asymmetry
						25	50	75	
Name agreement (%)	313	23.33	100.00	89.64	14.45	83.33	96.55	100.00	-1.80
Name agreement (<i>H</i>)	313	0.00	2.24	0.40	0.49	0.00	.22	.66	1.35
Familiarity ^a	312	1.51	4.66	3.07	0.90	2.18	3.06	3.94	0.04
Image agreement ^a	313	2.21	4.82	3.77	0.49	3.46	3.82	4.15	0.54
AoA ^a	312	1.09	4.17	2.05	0.59	1.58	1.97	2.46	0.61
Imageability ^a	312	2.00	4.97	4.46	0.34	4.31	4.54	4.69	-2.22
Visual complexity ^c	313	76	1357	231	140	133	178	307	2.57
Nb of letters ^b	313	3.00	16.00	6.80	2.53	5.00	6.00	8.00	1.17
Nb of phonemes ^d	291	2.00	11.00	4.56	1.58	3.00	4.00	6.00	0.69
Subtitle frequency ^d	291	0.00	570.30	29.56	59.66	3.46	9.83	29.61	4.98
Book frequency ^d	291	0.07	788.72	41.87	89.73	4.05	13.85	32.77	4.61

Notes. *N* = number of items; Objective norms were taken from Lexique 3 (New, Pallier, Brysbaert, & Ferrand, 2004).

^aOne word (ski) was accidentally omitted

^bBlanks were not counted for compound words

^cexpressed in 10³ bytes

^dAs compound words are generally not included in the Lexique database, the statistics taken from this database are given for only 291 words out the 313 items used in IMABASE.

Table 6

Bivariate correlations

		Name agreement (H)	Familiarity	Image agreement	AoA	Imageability	Visual complexity	Nb of letters	Subtitle frequency (log)	Book frequency (log)
Name agreement (%)	<i>r</i>	-.942***	.066	.300***	-.288***	.447***	-.017	.088	.029	.057
	95% CI	[-.96, -.92]	[-.06, .18]	[.18, .41]	[-.43, -.11]	[.26, .59]	[-.13, .10]	[-.04, .20]	[-.10, .15]	[-.06, .17]
Name agreement (H)	<i>r</i>		-.052	-.314***	.239***	-.404***	.026	-.080	.026	-.003
	95% CI		[-.17, .07]	[-.42, -.20]	[.08, .37]	[-.53, -.24]	[-.09, .14]	[-.19, .05]	[-.10, .16]	[-.12, .11]
Familiarity	<i>r</i>			-.070	-.588***	.534***	-.176**	-.169**	.457***	.531***
	95% CI			[-.18, .04]	[-.66, -.51]	[.46, .61]	[-.29, -.07]	[-.28, -.06]	[.37, .54]	[.44, .61]
Image agreement	<i>r</i>				.079	.144*	.026	.052	-.257***	-.217***
	95% CI				[-.04, .20]	[.03, .24]	[-.10, .14]	[-.06, .15]	[-.37, -.13]	[-.31, -.10]
AoA	<i>r</i>					-.710***	.039	.233***	-.533***	-.539***
	95% CI					[-.77, -.64]	[-.07, .14]	[.10, .37]	[-.62, -.44]	[-.61, -.46]
Imageability	<i>r</i>						-.108	-.048	.276***	.273***
	95% CI						[-.21, .14]	[-.21, .004]	[.18, .38]	[.18, .37]
Visual complexity	<i>r</i>							.05	-.206***	-.246***
	95% CI							[-.06, .16]	[-.32, -.09]	[-.35, -.14]
Nb of letters	<i>r</i>								-.399**	-.446**
	95% CI								[-.49, -.31]	[-.53, -.36]
Subtitle frequency (log)	<i>r</i>									.901***
	95% CI									[.88, .92]

Notes. First line = Pearson correlation; ***: $p < .001$, **: $p < .01$, *: $p < .05$

Second line: 1000 bootstrapped 95% interval for the Pearson correlation

Computations were calculated for the 290 words/pictures for which all characteristics were available both in IMABASE and Lexique 3 (New, Pallier, Brysbaert, & Ferrand, 2004). Frequencies and visual complexity were log-transformed.

Table 7

Multiple regressions including one norm as DV and the others as IVs

		Dependent variables				
		Name agreement (%NA)	Familiarity	Image agreement	AoA	Imageability
Forward selection of splines	R2lin	.2993	.4395	.1950	.6620	.6308
	1st NL term	Imty (.0195**)	Imty (.0335***)	AoA (.0170*)		AoA (.0338***)
	2nd NL term		VC (.0124**)			
	3rd NL term		FF (.0119*)			
	R2 final	.3188	.4973	.2121	.6620	.6650
Independent variables (IV)	%NA		-.14**	.29***		.18***
	Familiarity	-.18**			-.17***	.26***
	Image agreement	.24***			.09*	.13***
	AoA		-.30***	(+) 6.15**		(-) 84.74***
	Imageability	(+) 19.71***	(+) 20.76***	.30***	-.53***	
	Visual complexity		(-) 4.28*		-.09*	
	Nb of letters			-.12*		
Subtitle frequency		(+) 9.37***	-.29***	-.27***		

Notes. First part: R2lin = R2 with no nonlinearities; R2 final = R2 with nonlinearities included; 1st (2nd, 3rd) NL = first (second, third) IV for which cubic splines were included with the IV name and, in brackets, the percentage of added variance and its significance. Second part: beta coefficient or *F* value for IV including nonlinearities (for these variables, a (-) or a (+) is added to indicate the direction of change).