CrossMark

Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times

Patrick Bonin [1] · Alain Méot [2] · Aurélia Bugaiska [1]

Abstract

Words that correspond to a potential sensory experience—concrete words—have long been found to possess a processing advantage over abstract words in various lexical tasks. We collected norms of concreteness for a set of 1,659 French words, together with other psycholinguistic norms that were not available for these words—context availability, emotional valence, and arousal—but which are important if we are to achieve a better understanding of the meaning of concreteness effects. We then investigated the relationships of concreteness with these newly collected variables, together with other psycholinguistic variables that were already available for this set of words (e.g., imageability, age of acquisition, and sensory experience ratings). Finally, thanks to the variety of psychological norms available for this set of words, we decided to test further the embodied account of concreteness effects in visual-word recognition, championed by Kousta, Vigliocco, Vinson, Andrews, and Del Campo (Journal of Experimental Psychology: General, 140, 14–34, 2011). Similarly, we investigated the influences of concreteness in three word recognition tasks—lexical decision, progressive demasking, and word naming—using a multiple regression approach, based on the reaction times available in Chronolex (Ferrand, Brysbaert, Keuleers, New, Bonin, Méot, Pallier, Frontiers in Psychology, 2; 306, 2011). The norms can be downloaded as supplementary material provided with this article.

Keywords Concreteness · Imageability · Context availability · Emotional valence · Arousal · Sensory experience ratings · Word recognition

Psycholinguistic norms have been collected for various types of stimuli. To name just a few, norms have been obtained for pictures of objects (e.g., Boukadi, Zouaidi, & Wilson, 2016; Ghasisin, Yadegari, Rahgozar, Nazari, & Rastegarianzade, 2015; Snodgrass & Vanderwart, 1980), pictures of actions (e.g., Masterson & Druks, 1998; Shao & Stiegert, 2016), the faces of celebrities (e.g., Bonin, Perret, Méot, Ferrand, & Mermillod, 2008), multiword expressions (e.g., phrases: Arnon, McCauley, & Christiansen, 2017; or idiomatic

expressions: Bonin, Méot, Boucheix, & Bugaiska, 2018; Bonin, Méot, & Bugaiska, 2013), and, of course, words. Three important aims underpin the collection of norms. First of all, having norms for stimuli permits the methodological or statistical control of the potential influences of confounding variables when investigating a specific factor in a given lexical task. In effect, word characteristics must be very strictly controlled for, in light of the huge number of studies conducted in the past that have revealed a number of variables that affect the processing speed and accuracy of words, as well as their memorability. Second, having psycholinguistic norms helps establish the underlying structure of the norms, and this permits a better understanding of exactly what it is they measure. Last, but not least, the availability of norms makes it possible to investigate the norms' relationships with online (or offline) measures of word processing. Likewise, the use of psycholinguistic norms has played an important role in constraining several models of lexical processing. In the present study, we collected norms of concreteness in French for a set of 1,659 words for which other psycholinguistic norms were already available (e.g., age of acquisition [AoA] and

✉ Patrick Bonin
Patrick.Bonin@u-bourgogne.fr

[1] LEAD-CNRS (UMR 5022), Univ. Bourgogne Franche-Comté, Pôle AAFE—Esplanade Erasme, BP 26513, 21065 Dijon Cedex, France

[2] Université Clermont Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France

imageability; see below). However, in order to address the issues of the structure of concreteness norms and of how concrete words are processed relative to abstract words in visual-word recognition, it was also important to collect psycholinguistic norms that, thus far, have not been available for this set of words: context availability, emotional valence, and arousal. Below we describe the reasons why it is important to take account of such norms when investigating the concreteness variable. Beyond this issue, we think that these additional norms—which are available in the supplementary materials, together with the concreteness norms—will be of interest to researchers who wish to investigate word recognition or memory. We now discuss in more detail the different goals that are pursued when collecting norms for words, focusing on the concreteness variable in particular.

Availability of norms helps the control of variables

In recent years, concreteness norms have been collected for a large number of words (e.g., 40,000 words in Brysbaert, Warriner, & Kuperman, 2014b) and in different languages and cultures. To name just a few recent studies, concreteness norms have been collected in Spanish (Guasch, Ferré, & Fraga, 2016), in Portuguese (Soares, Costa, Machado, Comesaña, & Oliveira, 2018), and in Chinese (Yao, Wu, Zhang, & Wang, 2018). Norms of concreteness correspond to ratings of the degree to which the concepts denoted by words refer to perceptible entities (e.g., objects, persons, places; Brysbaert, Warriner, & Kuperman, 2014). To illustrate, in the English norms of the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982), *table* and *heaven* have values of 6.8 and 2.5 (out of 7), respectively. From these values, it is possible to categorize the former word as concrete and the latter as abstract.

Since concreteness is a well-known variable that affects the processing of words in various lexical-processing tasks, it is important to take this variable into account when investigating the effect of another variable that is the focus of interest. To illustrate this issue, let us consider a recent mnemonic effect referred to as the *animacy effect* in memory (Bonin, Gelin, & Bugaiska, 2014; Nairne, VanArsdall, Pandeirada, Cogdill, & LeBreton, 2013). It corresponds to the observation that words referring to animate concepts are remembered better than words referring to inanimate concepts. To account for this finding, it has been assumed that animate entities have a greater fitness value than do inanimates (Bonin et al., 2014; Nairne et al., 2013). On the basis of an evolutionary–functional account, the idea is that our memory systems were sculpted by a range of pressures faced by our ancestors in the distant past—that is, the adaptive view of memory (Nairne, 2010). Since animates may potentially refer to mates, prey, or predators, they are therefore more important for survival and/or

reproduction than are inanimates. Of course, to provide convincing evidence for this effect and support an evolutionary account of it, researchers have to test for other (and somewhat more trivial) explanations, such as the possibility that animates are easier to imagine or are more concrete than inanimates. In effect, given that concrete words are remembered better than abstract words (Paivio, 1971; Paivio, Yuille, & Smythe, 1966; ter Doest & Semin, 2005), and because concreteness is one of the most important variables accounting for the recall rates of words presented in unrelated lists (Nairne et al., 2013), this variable must be controlled for when investigating animacy effects. As a result, the availability of concreteness norms for words is essential for researchers who seek to investigate memory.

Availability of norms helps the identification of the deep structure of psycholinguistic variables

Concreteness, like imageability, is a semantic variable (but see also Reilly & Kean, 2007, who have shown that word-form properties reliably differentiate between high- and low-imageability/concrete words); *imageability* corresponds to the ease of forming a mental image for words. It has often been assumed that imageability indexes the richness of words, with the result that more imageable words are semantically richer than less imageable words (Yap & Pexman, 2016). In the literature, as was pointed out by Kousta, Vigliocco, Vinson, Andrews, and Del Campo (2011), certain authors have used concreteness and imageability interchangeably (e.g., Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Reilly & Kean, 2007), perhaps because the correlation between the two variables is high (e.g., .78 in Yao et al.'s, 2018, study). However, even though the two variables are closely related, they are not thought to be equivalent. Kousta et al. (2011) found that the distributions of imageability and concreteness ratings were different. In effect, the frequency distribution of concreteness scores is bimodal (see also Della Rosa, Catricalà, Vigliocco, & Cappa, 2010, and Brysbaert, Warriner, & Kuperman, 2014b, for a less extreme distribution), with one mode corresponding to abstract words, and the other mode to concrete words. In contrast, the frequency distribution of imageability scores is unimodal. According to Kousta et al., the difference between the two distributions is due to the fact that the ratings of concreteness index the categorical ontological distinction between concrete and abstract concepts and their corresponding words. Imageability ratings instead capture the differential association of words with sensory, and more specifically visual properties, and the latter sensory property is graded. It is important to note that, at the theoretical level, imageability ratings are taken as a proxy for concreteness within the *dual-code* model of semantic memory (Paivio,

1971, 1986, 1991, 2007). According to this view, there are two distinct cognitive systems for processing the meanings of words. The first system comprises verbal knowledge, and the second system consists of image-based knowledge. All words engage linguistic codes, but concrete words activate image-based codes to a greater extent than do abstract words. By contrast, the context availability account assumes that only one system is used to derive the meanings of abstract and concrete words (Schwanenflugel, 1991; Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Shoben, 1983). The semantics of the words are accessed by linking words with preexisting knowledge stored in long-term memory, referred to as "context." Concrete and abstract words differ in the quantity and quality of their available contexts, with the former words having "richer" contexts than the latter. Given the theoretical implications of imageability and concreteness rating distributions, we will assess whether the patterns for these two types of ratings found in previous studies also apply to the present concreteness ratings and imageability ratings, which were collected in French for the same words as in a previous study (Bonin et al., 2003).

To achieve a better understanding of what aspects of semantics concreteness indexes, one has to analyze precisely how this variable relates to other important psycholinguistic variables. As we said above, among the variables that must be taken into account are imageability and context availability. Kousta et al. (2011) have provided evidence for the hypothesis that sensory-motor information is more important for the representation of concrete than of abstract words and that emotional information is, in contrast, more important for abstract than for concrete words. As a result, we will examine how concreteness ratings collected for the present set of French words relate to other psycholinguistic variables and, in particular, sensory experience ratings (SER), emotional valence ratings, and arousal ratings. Emotional information is generally assessed in terms of valence and arousal. Valence is the extent to which an emotion is positive/pleasant, negative/unpleasant or neutral, whereas arousal is the degree of physiological activation that a word evokes and varies from calm/relaxed to excited (Gilet, Grühn, Studer, & Labouvie-Vief, 2012; Monnier & Syssau, 2014). SER is a new variable that corresponds to the degree to which words evoke sensory and perceptual experiences (Juhasz & Yap, 2013). Indeed, we recently collected SER for the set of French words used in the present study (Bonin, Méot, Ferrand, & Bugaiska, 2015). The view championed by Kousta et al. regarding the influence of concreteness leads us to anticipate that SER should be a critical variable that has to be taken into account for a better understanding of concreteness effects.

To recap, to investigate the relationships between concreteness norms and other psycholinguistic variables, we have collected norms of concreteness for a set of 1,659 French words for which previous norms had been collected (i.e., AoA and

subjective frequency [Ferrand et al., 2008], SER [Bonin et al., 2015], and imageability [Bonin et al., 2003]). We also collected norms of context availability, emotional valence and arousal since these norms were not available for this set of French words and because they are of critical importance for understanding the dimensions that concreteness captures. Moreover, the findings in the literature regarding the relationships between valence, arousal and concreteness remain equivocal (Yao et al., 2018, and below) and more studies are therefore needed to improve our understanding of their relations.

The availability of norms constrains models of lexical processing of psycholinguistic variables

The concreteness effect is the finding that words that refer to concrete entities have a processing advantage in a variety of cognitive tasks over words that are less concrete/more abstract. Likewise it has been found that concrete words are processed more quickly than abstract words in lexical decision (e.g., Binder, Westbury, McKiernan, Possing, & Medler, 2005; de Groot, 1989; Kroll & Merves, 1986; Roxbury, McMahon, & Copland, 2014; Schwanenflugel et al., 1988; Schwanenflugel & Stowe, 1989) and in word naming (e.g., de Groot, 1989; Schwanenflugel & Stowe, 1989). Concrete words are also remembered better than abstract words (Paivio, 1971; Paivio et al., 1966; ter Doest & Semin, 2005). This processing advantage is considered to be a classical effect in the psycholinguistic literature and it is indeed covered in every textbook on cognitive psychology.

As we mentioned above, two classical views of concreteness effects have been influential: the dual-coding view (e.g., Paivio, 2007) and the contextual-availability view (e.g., Schwanenflugel, 1991). However, recent articles (Kousta et al., 2011; Vigliocco et al., 2014; Vigliocco, Kousta, Vinson, Andrews, & Del Campo, 2013) have called for a new approach to concreteness effects. According to Kousta et al. (2011), neither the dual-coding nor the context-availability view exhaustively accounts for concreteness effects, and accordingly, these authors put forward an alternative, *embodied* account of concreteness. Embodied views of cognition assume that cognition is grounded in bodily states, modal simulations, and situated action (e.g., Barsalou, 1999; Barsalou, Simmons, Barbey, & Wilson, 2003; Decety & Grèzes, 2006; Rizzolatti & Craighero, 2004). The core assumption is that the representation and processing of semantic information call on the same neural systems that are involved during perception and action. Using factorial experiments and sophisticated multiple-regression models with lexical-decision times as the dependent variable, Kousta et al. found that, once imageability and context

availability ratings had been taken into account, abstract words were processed faster than concrete words.[1]

It is important to stress that neither the dual-coding view nor the context-availability view predicts a processing advantage of abstract over concrete words. In effect, the dual-coding view predicts that concreteness effects should vanish when imageability values are controlled for. Following the context-availability view, when the number of contexts is controlled for, concrete words should be processed to the same extent as abstract words. According to Kousta et al. (2011), the processing advantage of abstract over concrete words can be explained by the fact that abstract words have greater affective associations. More precisely, two classes of information contribute to the representation of both concrete and abstract concepts: experiential (sensory, motor, and affective) and linguistic (verbal associations arising through patterns of co-occurrence and syntactic information) information. The difference between the two types of words arises as a result of the types and the relative proportions of experiential and linguistic information they bind. Given the findings of Kousta and colleagues, and the strong theoretical implications they have for our understanding of concreteness effects, we thought it important to provide further empirical evidence for Kousta et al.'s view of concreteness in word processing. Beyond simple replication purposes that are now widely acknowledged as being of value in the psychological sciences (Roediger III, 2012; Simon, 2014), we wanted to assess whether the reversed concreteness effect observed in lexical decision would also be observed in other word recognition tasks such as word naming or progressive demasking.

The present study

To summarize, the present study had three main aims. The first was to provide concreteness norms in addition to context availability, valence and arousal ratings for a large set of French words. These norms will be of interest to researchers investigating memory or language when designing factorial experiments or when conducting multiple regression analyses. The second aim was to analyze the relationships between concreteness and other important psycholinguistic variables in order to get a better understanding of the underlying structure of this variable. The third and final aim was to investigate the influences of concreteness in lexical decision, word naming, and progressive demasking by using response times (RTs) that were available from a previous French study (Ferrand et al., 2011), while including several lexical (e.g., word frequency

and AoA) and semantic (e.g., imageability, arousal, valence, and context availability) variables, in order to provide a rigorous analysis of how concreteness influenced the tasks.

Method

Participants

A total of 482 adults, native speakers of French (368 females, 114 males; mean age 30.18 years) with normal or corrected-to-normal vision, took part. Most of them were undergraduates in psychology, but others were undergraduate students at a speech-therapist school in Besançon (France) or adult acquaintances of the latter. All the participants were volunteers and received course credit for their participation.

Stimuli

We used the 1,493 words for which AoA and subjective frequency ratings were available in French (Ferrand et al., 2008), as well as imageability ratings (Bonin et al., 2003). In addition, we included the modal names (i.e., the corresponding French nouns) of 166 pictures taken from the Snodgrass and Vanderwart (1980) norms, since these are very often used in memory or psycholinguistic experiments. The words had a mean length of 4.96 and ranged from two to 13 letters in length. Most of the words were nouns (1,117), but there were also verbs (157), adjectives (187), adverbs (28), and pronouns (27). The rest of the words consisted of prepositions, conjunctions, articles or onomatopoeia.

Procedure

The 1,659 stimuli were randomly divided into four lists: Three lists contained 415 words and one list contained 414 words. Two words were accidentally omitted from the third (*triste* meaning sad) and fourth (*grain* meaning seed) lists used to collect arousal norms. A paper-and-pencil procedure was used and two to three participants rated the words simultaneously in a quiet room. There were different groups of participants for each questionnaire, corresponding to the different norms. The order of the words was randomized across the four lists used for each type of rating (and for a given type of rating and a given list, the words were presented in the same order to the participants). With the exception of a 25-participant group who completed one of the concreteness questionnaires, there were between 29 and 33 participants in each group.

After having provided socio-demographic information (e.g., age), the participants were informed that they would be presented with a list of words and that, depending on the group, they would have to rate concreteness, contextual availability, valence or arousal. Likert scales ranging from 1 to 5

---

[1] It is important to stress that more efficient processing of abstract than of concrete words has also been reported in patients with semantic dementia (e.g., Bonner et al., 2009; Papagno, Capasso, Zerboni, & Miceli, 2007).

were used to collect the different norms.[2] The instructions for the different norms were either adapted on the basis of original instructions taken from previous published studies (e.g., arousal [Monnier & Syssau, 2014], contextual availability [Altarriba, Bauer, & Benvenuto 1999]), or were taken from a previous French normative study (e.g., concreteness and valence; Bonin et al., 2003).

The instructions were provided both orally by the experimenter and in written form on the first page of the questionnaires (see the supplementary material for the exact wording). For each rating task, the instructions provided examples of ratings of several words that were not included in the list. Although no time limit was imposed, the participants were told to provide spontaneous responses (and thus not take too much time rating each word). They were also instructed that there were no right or wrong responses and that the experimenter was only interested in their personal answers. Finally, the participants were informed that they could make short breaks. An entire rating session lasted about 1 h.

Results and discussion

Five participants who gave scores for fewer than two thirds of the words included in a list were excluded from the arousal rating task. The valence ratings provided by one participant were not taken into account because of a mean correlation of .003 between his/her scores and the ratings provided by the other participants.

Reliability of the norms

Overall, the reliability scores were high for all the collected variables. In effect, as can be seen in Table 1, the correlations between the scores obtained from the means of the even and odd participants were between .80 and .96 for the entire set of words. Within-list reliability scores were also computed, because they provide information about the homogeneity of the ratings across the different lists. As Table 1 shows, the lists were homogeneous, and the patterns that were found for the entire set of words were also found across the different lists. To further gauge the reliability of our ratings concerning concreteness and valence, we correlated them with the ratings taken from a previous French study (Bonin et al., 2003). There were 426 words in common. The correlations were above .75 (the correlation coefficients for concreteness and valence were .85 and .83, respectively, all ps < .001). As far as valence and arousal ratings were concerned, we correlated

our ratings with those taken from two previous French studies (Gobin, Camblats, Faurous, & Mathey, 2017; Monnier & Syssau, 2014). The correlations for arousal were .44 with Gobin et al. (2017) [165 words in common], and .32 with Monnier and Syssau (2014) [445 words in common], all ps < .001.[3] For valence, the correlations were .96 and .86, respectively, all ps < .001. Interestingly, and perhaps in line with the lower correlations across databases that were found for arousal, we replicated the finding that interrater reliability was greater for the valence than for the arousal values reported in other studies (e.g., Gobin et al., 2017; Monnier & Syssau, 2014; Montefinese, Ambrosini, Fairfield, & Mammarella, 2014; Yao et al., 2018). Taken overall, the different reliability analyses indicate that our collected norms of concreteness, context availability, valence, and arousal are highly reliable.

Descriptive statistics

In the supplementary materials, the words are provided in alphabetical order with their English translations. The database also provides the mean ratings and their standard deviations for each word on concreteness, context availability, valence, and arousal. Table 2 shows the means, standard deviations, minima and maxima, medians, Q1 and Q3, and skews for the different collected variables. Descriptive statistics concerning the other norms used to investigate the relationships between the collected norms and other word characteristics (e.g., word frequency, SER) are provided in Table A1 of the supplementary materials.

Table 2 and Fig. 1 show that concreteness scores were distributed relatively homogeneously across the entire scale, with a marked main mode situated at the top-right of the scale. The distributions of context availability, valence, and arousal ratings were all unimodal, with relatively marked negative and positive skews for the latter two variables (Fig. 1). By contrast, there were two modes for the distribution of concreteness ratings (dip = .021, p < .001, bimodality coefficients > .555). The location of the main mode at the top right of the scale suggests an overrepresentation of very concrete words. Thus, in accordance with what has been reported by other researchers (Brysbaert, Warriner, & Kuperman, 2014b; Della Rosa et al., 2010; Kousta et al., 2011), we found a bimodal distribution of concreteness ratings. By contrast, the imageability ratings distribution obtained by Bonin et al. (2003) was unimodal (dip = .009, p > .10, bimodality coefficients < .555).

---

[2] The number of points that should be used in Likert scales is a matter of debate. It seems that there is an apparent consensus to use a limited number of points—that is to say, between 5 and 7 points (e.g., Krosnick & Presser, 2010).

---

[3] It is important to note that the arousal ratings from Monnier and Syssau (2014) were also less correlated with those taken from Gobin et al. (2017) [96 words in common, r = .56, p < .001] than was the case for the valence ratings (r = .83, p < .001). Thus, these findings suggest that there is greater variability in the ways that participants rate words for arousal than for valence.

Table 1 Reliability analyses

	r(even , odd)		Means
	All Words	Within Lists Min & Max	Within Lists Min & Max
Concreteness	.95	.94–.96	3.41–3.46
Context availability	.80	.77–.87	2.97–3.23
Valence	.92	.92–.94	3.06–3.25
Arousal	.84	.85–.89	2.52–2.75

r(even , odd) = correlations from the means of the even and odd participants

Relationships between concreteness, context availability, valence, arousal, and other word characteristics

Figure 2 shows the lowess adjustments obtained between one of the present norms and the remaining ones. The Spearman correlation coefficients are provided in Table 3.

The most important observations were the following. First of all, two pairs of the variables were highly correlated (see Table 3): (1) Concreteness and context availability were highly positively correlated, and (2) valence and arousal were highly negatively correlated. First of all, the high positive correlation of concreteness with context availability was due to a nearly linear relationship. This result is in line with other studies (Altarriba et al., 1999; Guasch et al., 2016; Schwanenflugel et al., 1988; Yao et al., 2018). Thus, more concrete words are associated with more contexts than are abstract words. This observation is precisely what the context-availability view of concreteness predicts. In effect, according to this view, the processing advantage of concrete words is due to differences in the quantity and in the quality of available contexts, with concrete words benefiting from "richer" contexts than abstract words. Second, the correlation between concreteness and valence was also positive, but clearly lower than the correlation between concreteness and context availability. Moreover, it resulted from a nonlinear relation, with more negatively and positively valenced words being less concrete than intermediate ones. In contrast to this finding (of a positive correlation), both Hinojosa et al. (2016) and Guasch et al. (2016) found a negative correlation between valence and concreteness in Spanish. Finally, the

correlation between concreteness and arousal was negative but low, with a linear relation being observed between the two. This type of negative correlation was also found by both Guasch et al. (2016) and Soares et al. (2018). However, Hinojosa et al. (2016) reported a positive relationship, whereas Montefinese, Ambrosini, Fairfield, and Mammarella (2014) found nonlinear relationships of concreteness with arousal ratings, with the result that very abstract or very concrete words were rated as "calm," whereas words of middle concreteness were rated as "exciting."

Concerning the relationship between valence and arousal, previous studies have often described this relationship by means of a U-shaped distribution (e.g., Citron, Weekes, & Ferstl, 2014; Ferré, Guasch, Moldovan, & Sánchez-Casas, 2012; Guasch et al., 2016; Montefinese et al., 2014; Warriner, Kuperman, & Brysbaert, 2013). More precisely, it has been found that (very) positive and (very) negative words are rated as being the most arousing stimuli, whereas items with low positive and negative valence ratings are rated as being the least arousing. However, we did not find this relationship in the present norms (see also Gilet et al., 2012, for a failure to find a U-shaped distribution between valence and arousal in French adjectives, especially among older adults). Instead, as was found by Gilet et al., the correlation between valence and arousal was negative, with a near-linear relationship that was, however, stronger at the left of the scale. It is important to note that this U-shaped relationship was not found when we investigated it using a quadratic model as, for example, Warriner et al. did or, alternatively, when we used simple linear models run separately for the most positive words and the most negative words.

Figure 3 shows the relationships between the present norms and a selection of other subjective norms (e.g., SER, imageability) or objective norms (e.g., subtitle frequency [in log], number of letters).

Spearman correlations are provided in Table 4. As far as concreteness is concerned, Fig. 3 shows that the relationship with imageability was positive and mostly linear. The concrete words were easier to imagine than the abstract words. These findings have been found in other languages (e.g., Guasch et al., 2016; Paivio, 1986, 1991; Yao et al., 2018), and they are in line with the two traditional accounts of concreteness effects: the dual-code view (Paivio, 1971, 1986, 1991, 2007)

Table 2 Descriptive statistics for the norms of concreteness, contextual availability, valence, and arousal

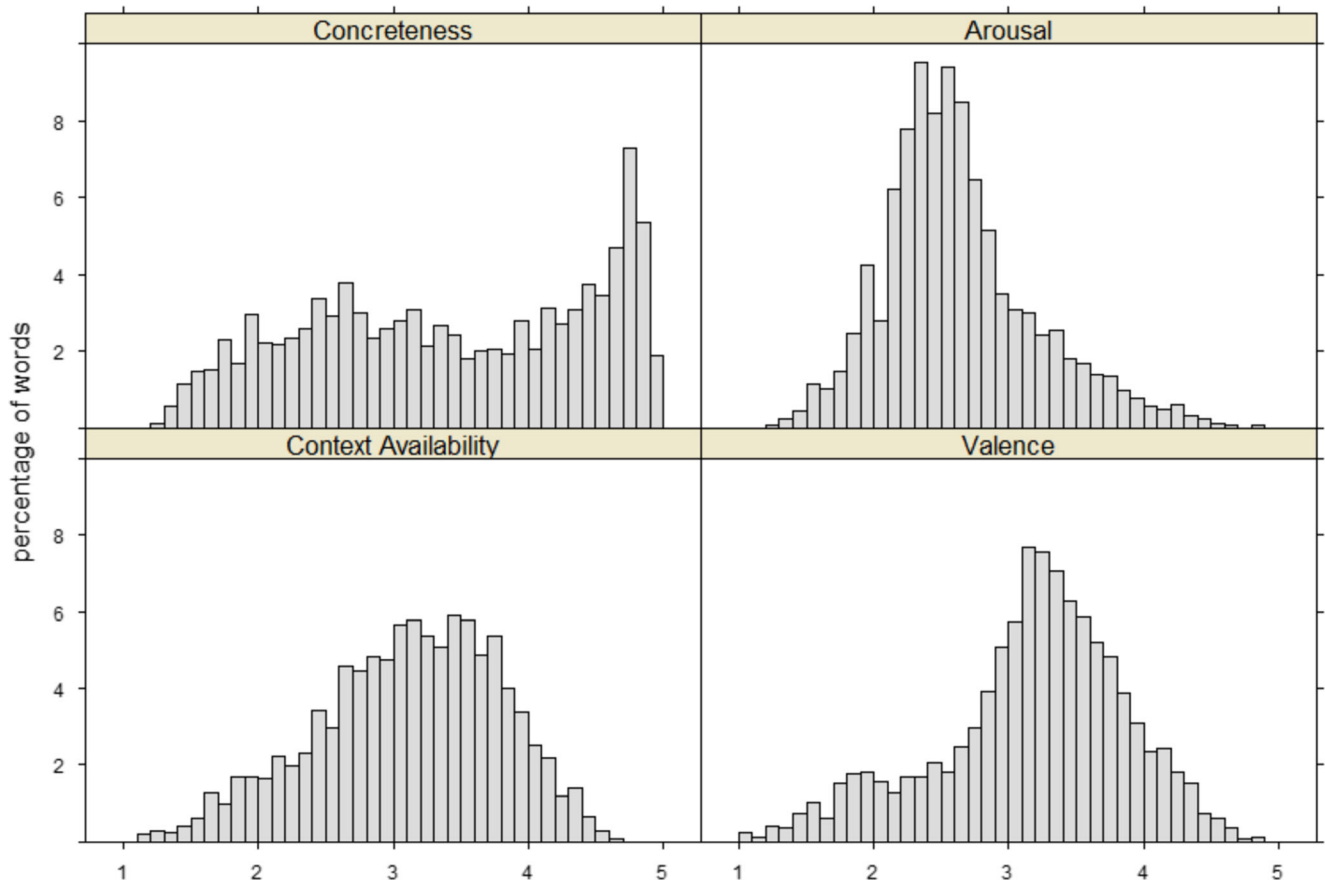|	N	Min	Max	Mean	SD	Q1	Median	Q3	Skew
Concreteness	1,659	1.23	5.00	3.43	1.05	2.53	3.43	4.47	− .16
Contextual availability	1,659	1.13	4.63	3.13	0.68	2.70	3.20	3.63	− .37
Valence	1,659	1.00	4.90	3.18	0.70	2.83	3.27	3.66	− .58
Arousal	1,657	1.21	4.90	2.63	0.57	2.28	2.55	2.90	.75

Fig. 1 Distributions of ratings of concreteness, arousal, context availability, and valence

and the context-availability view (e.g., Schwanenflugel et al., 1988). By contrast, negative correlations were observed

between concreteness and subtitle frequency, subjective frequency, and, to lesser extents, AoA and number of
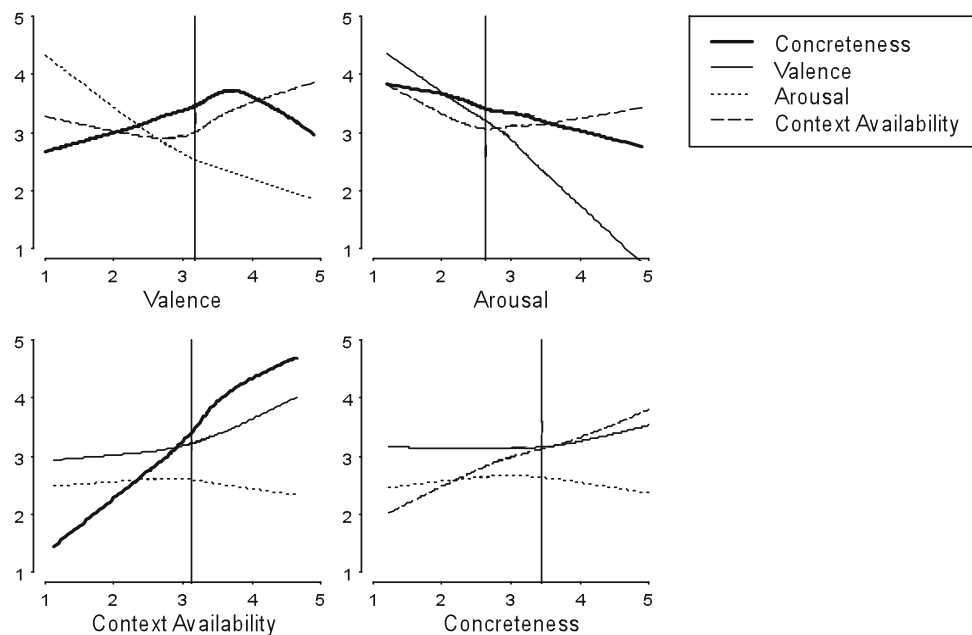

Fig. 2 Relationships between the collected norms (concreteness, valence, arousal, and context availability). Lines represent lowess regressions

between the represented norms and the variable on the abscissa. Vertical line = mean of the norm on the abscissa

Table 3 Spearman correlations between concreteness, context availability, valence, and arousal

	Context Availability	Valence	Arousal
Concreteness	.64	.20	−.15
Context availability		.32	−.09
Valence			−.63

All ps < .001

homographs. For objective word frequency, the decreasing relationship was observed, however, only at the right of the scale, whereas concreteness was relatively stable for low frequency values, with the decrease setting in only close to the frequency mean. A negative correlation between concreteness and subjective frequency was also reported by Soares et al. (2018) in Portuguese. A number of studies have found that concrete words are acquired earlier than abstract words (e.g., Barca, Burani, & Arduino, 2002; Bird, Franklin, & Howard, 2001; Cameirão & Vicente, 2010; Kuperman, Stadthagen-Gonzales, & Brysbaert, 2012), and we also found this relationship in our data.

Concreteness scores increased continuously with SER. However, the curve was nonlinear, with the growth rate diminishing with higher SER ratings. This observation is in line with Kousta et al.'s (2011) view of concreteness, which assumes that concrete words are richer than abstract words in terms of sensory information. However, the correlation between concreteness and imageability was stronger than the correlation between concreteness and SER. In Bonin et al. (2015), we suggested that imageability more readily captures the visual information corresponding to the semantic aspects of the words, whereas SER captures this dimension along with other sensory information (e.g., gustatory, olfactory, and auditory information). Thus, the correlations found here suggest that concreteness seems to be related more to the visual–semantic aspects of words than to other types of sensory information.

Given their high correlation (see Table A2 in the supplementary materials), number of letters and orthographic unicity point exhibited the same relationship with concreteness. As was found for the English language, nouns denoting abstract concepts were longer than nouns denoting concrete ones (Reilly & Kean, 2007). Interestingly, the same positive
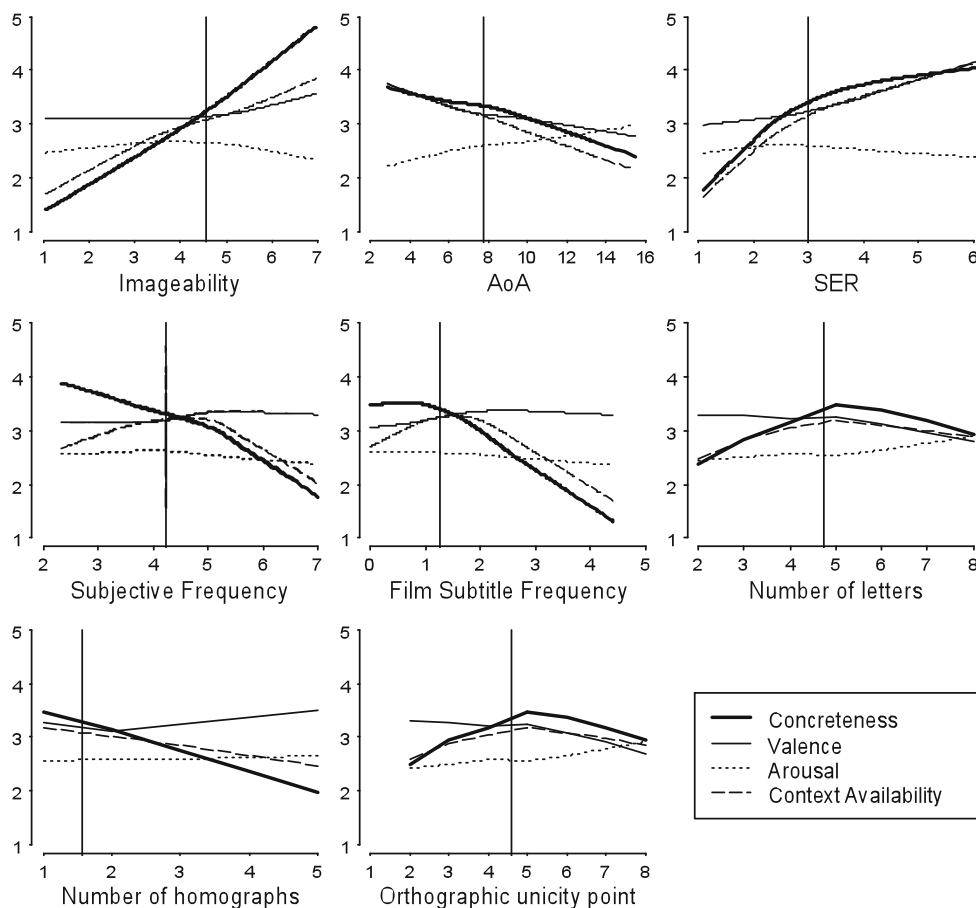

Fig. 3 Relationships between the present norms and other psycholinguistic norms. Lines represent lowess regressions between the represented norms and the variable on the abscissa. Vertical line = mean of the norm on the abscissa

Table 4 Spearman correlations between the collected norms and other psycholinguistic variables

	Image-ability	AoA	SER	Subjective Frequency	Log Subtitle Frequency	Number of Letters	Number of Homographs	Orthographic Unicity Point
Concreteness	.89	− .18	.43	− .33	− .30	.16	− .20	.14
Context availability	.74	− .36	.68	.01 ns	.06[*]	.09	− .14	.07[**]
Valence	.22	− .24	.24	.12	.15	− .08[**]	− .08[**]	− .09
Arousal	− .15	.23	− .02 ns	− .09	− .05[*]	.10	.04 ns	.10

Spearman correlations were computed over the 1,478 words for which all norms were available. All $ps < .001$, except for the following symbols: ns = not significant; [*] $p < .05$; [**] $p < .01$

relationship was observed in Bonin et al.'s (2003) study. However, in the present study, a nonlinear relation was also present, with both short and long nouns being less concrete than intermediate ones (see Fig. 3).

It is important to note that certain relationships with other variables are not reported here (the correlations can be found in Table A3 of the supplementary materials). The relationships between context availability and other psycholinguistic norms generally mimic those between concreteness and the same norms, whereas those between arousal and the other variables were weak. There were also weak relationships between valence and the other norms, the most salient being with imageability (positive, as found in Citron et al., 2014), with SER (positive), and with AoA (negative). Kousta et al. (2011, Fig. 7) found that, in the case of abstract words, AoA and valence were related by an inverted U-shaped function. Using exactly the same procedure employed by those authors to analyze the data, we also found this relationship when we used a quadratic model restricted to words with concreteness values below the mean, $F(2, 711) = 38.425, p < .001, R^2 = .098$; quadratic alone: $F(1, 711) = 13.29, p < .001, \Delta R^2 = .017$.

Multiple linear regressions were performed in order to examine the relationships between concreteness and other psycholinguistic norms. In a first analysis, the independent variables that were included were context availability, valence, arousal, SER, AoA, subtitle frequency, number of letters, and number of homographs. A second analysis included imageability in the regression equation. The reason why we did not include imageability directly in the first analysis is because there are strong relationships between this variable and both concreteness and other psycholinguistic variables, thus leading to results that differ considerably when imageability is not included in the equation.

On the basis of the previous analyses reported above, we excluded certain variables from the analyses. This was because of (1) very weak relationships with the dependent variable or because (2) the relationships with the dependent variables were comparable with those with the independent variables that were included (though less pronounced). All linear terms were included. Nonlinearities were modeled using restricted cubic splines for some independent variables. To obtain smoothings that did not have functional forms that were too specific to only limited parts of the independent variable scales, the maximum number of knots was set to six. In addition, given that there were only five unique values for the number of homographs and seven for the number of letters, no splines were used for the former variable, and the number of knots was limited to four for the latter variable. A forward approach was used. At each step, nonlinear terms were included for the independent variable whose inclusion led to the greatest increase in variance compared with the model that did not include such terms. At each step, the number of knots was fixed at the value at which the addition of one or two more knots did not cause R^2 to rise above .001. If nonlinear terms were, in general, not significant irrespective of the number of knots, the next independent variable was considered. The result of this procedure is summarized in Table A4 of the supplementary materials.

The model that took the splines (without imageability) into account explained 4.49% more of the variance than did the model that included only linear terms, with valence being the first independent variable entered, $F(5, 1461) = 22.93, p < .001$ [nonlinear, $F(4, 1461) = 28.39, p < .001$]. The relationship between concreteness and valence was roughly an inverted U, with the lowest concreteness scores being located at the extremes of the valence scale (Fig. 4), which is to say that more abstract words are more valenced, whereas neutral words tend to be more concrete. This observation is in line with other findings (in Spanish, Guasch et al., 2016; in English, Vigliocco et al., 2014; in Chinese, Yao et al., 2018), and it also fits with the embodied view of concreteness processing championed by Kousta et al. (2011), which assumes that abstract words are more closely associated with emotional information. Interestingly, and in accordance with behavioral

Fig. 4 Partial effects in the regression model without imageability—all ▶ predictors. Concreteness z scores are shown on the y-axis; standardized scores are used for all independent variables, on the abscissa. All independent variables except arousal were significant at $p < .05$
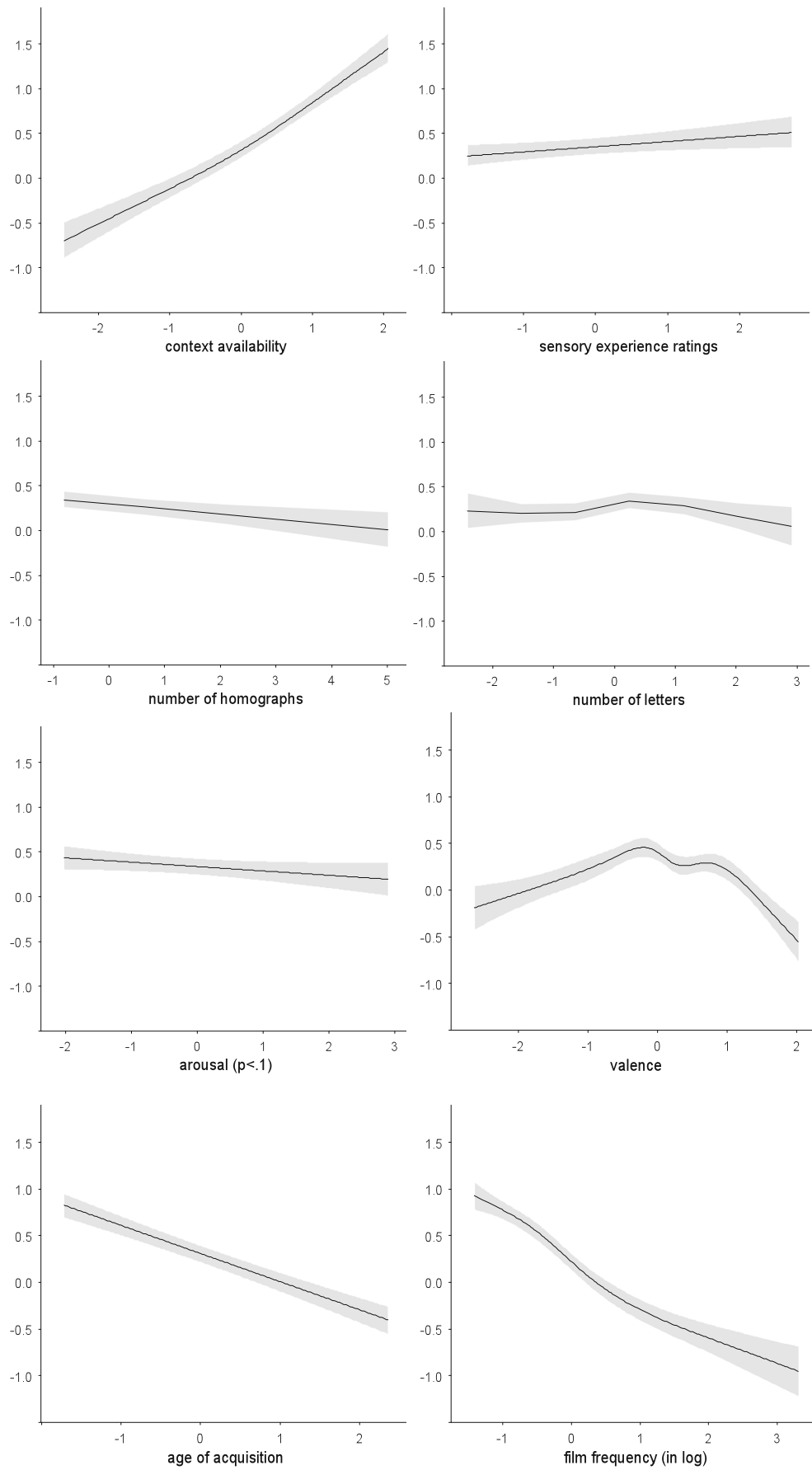
data, an fMRI study by Vigliocco et al. (2014) showed that emotionality ratings predicted modulation of the blood-oxygenation-level-dependent signal in the rostral anterior cingulate cortex, which is a brain region associated with the processing of emotional information.

Context availability, $F(2, 1461) = 160.57$, $p < .001$, had a positive effect on concreteness, with a small nonlinearity proportion, $F(1, 1461) = 5.41$, $p < .05$. The opposite relation was observed for both AoA, $F(1, 1461) = 130.76$, $p < .001$, and word frequency, $F(3, 1461) = 113.24$, $p < .001$ [nonlinear: $F(2, 1461) = 8.22$, $p < .001$], with the latter variable exhibiting relatively pronounced nonlinearity. The effects of the other variables were weaker: number of letters, $F(3, 1461) = 4.2$, $p < .01$ [nonlinear, $F(2, 1461) = 5.6$, $p < .01$]; number of homographs, $F(1, 1461) = 10.02$, $p < .01$; SER, $F(1, 1461) = 5.25$, $p < .05$. Arousal was marginally significant, $F(1, 1461) = 3.34$, $p = .0678$. The partial effects (see Fig. 4) were broadly consistent with univariate relationships, with the following exception: (1) word frequency, for which the flat portion of the relation at the beginning of the scale ceased to be observed when other independent variables were controlled for, and (2) arousal and SER, whose effects practically disappeared when other independent variables were controlled for.

The procedure described above was also used with imageability included. The order of entry of nonlinear terms is summarized in Table A5 of the supplementary materials. The inclusion of the imageability variable led to a range of results. First of all, we found an increase in R^2 of .26 in comparison with the previous regression analysis, and the explained variance was mostly due to imageability, $F(3, 1460) = 781.14$ $p < .001$ [nonlinear, $F(2, 1460) = 44.37$, $p < .001$]. Second, the effects that were weak in the first step of the analysis were no longer reliable: number of letters and SER, $F(1, 1460) < 1$; number of homographs, $F(1, 1460) = 1.62$, $p > .1$. Third, the following effects were still significant: valence, $F(5, 1460) = 4.82$, $p < .001$ [nonlinear, $F(4, 1460) = 5.42$, $p < .001$]; word frequency, $F(2, 1460) = 29.61$, $p < .001$ [nonlinear, $F(1, 1460) = 36.85$, $p < .001$]; context availability, $F(1, 1460) = 6.33$, $p < .05$; AoA, $F(1, 1460) = 15.35$, $p < .001$; and arousal, $F(3, 1460) = 4.88$, $p < .01$ [nonlinear, $F(2, 1460) = 7.26$, $p < .001$]. However, with the exception of arousal, which was only marginally significant in the analysis excluding imageability, the reliable effects were less strong.

Fourth, as can be seen in Fig. 5, the effects of context availability, AoA, and arousal were reversed as compared to the previous analysis. Fifth, the decrease in concreteness ratings for the higher objective frequency values disappeared. Finally, and interestingly, the effect of valence had the same form—that is to say, roughly an inverted U.

The analyses reported earlier suggest that both low- and high-valenced words are less concrete. Moreover, this property was observed whether or not imageability was included in

the regression equation used to predict concreteness scores (cf. Figs. 4 and 5).

Yao et al.'s (2018) study focused on affective variables (valence and arousal) and their relationships with other important psycholinguistic (semantic) variables such as concreteness, imageability and context availability. In Chinese, they found that concreteness was predicted by valence and its square by arousal (linear function). We share the claim made by Yao et al. that the relationship between arousal and concreteness is still uncertain because of discrepant findings in the literature (e.g., positive correlation: Vigliocco et al., 2014; negative: Guasch et al., 2016; the present study), which could be due to both the greater variability in the way participants rate the words for arousal (see the Reliability section) and differences in the ranges of the variables.

Imageability versus concreteness: Different measures but very similar

Certain researchers have claimed that imageability and concreteness could be used interchangeably (e.g., Connell & Lynott, 2012; Reilly & Kean, 2007). Others have assumed that concreteness is a semantic variable closely related to but different from imageability (Kousta et al., 2011), and that they are therefore not equivalent (e.g., Dellantonio, Mulatti, Pastore, & Job, 2014; Soares et al., 2018). Accordingly, the former view relates to the categorical ontological distinction between concrete and abstract concepts, whereas the latter differentiates between words in terms of their sensory (and more particularly, visual) properties. Given these theoretical divergences, and because we found different distributions of concreteness and imageability in the same way as in Kousta et al.'s (2011) and Montefinese et al.'s (2014) studies, we thought it important to study in more detail the similarities and the differences in the way the two variables are related to other important psycholinguistic variables.

As we will now describe, several analyses do indeed strongly suggest that the two constructs, though not equivalent, are very close and that they therefore detect analogous properties (Dellantonio et al., 2014). First of all, as reported in several studies (Guasch et al., 2016; Paivio, 1991; Schwanenflugel et al., 1988; Yao et al., 2018) and as was found here, the correlation between the two variables is very high, which suggests that the same construct is captured by both variables. One possible reason could be that in order to evaluate the concreteness of a word, people try to create a mental (mainly visual) image of what it refers to. Thus, trying to associate a word with (mainly visual) sensory features could be a first step when assessing concreteness. This assumption is consistent with the dual-code view (Paivio, 1971, 1986), according to which words vary as a function of the richness of their imagery information, with concrete words being richer in terms of imagery than abstract words. Second,
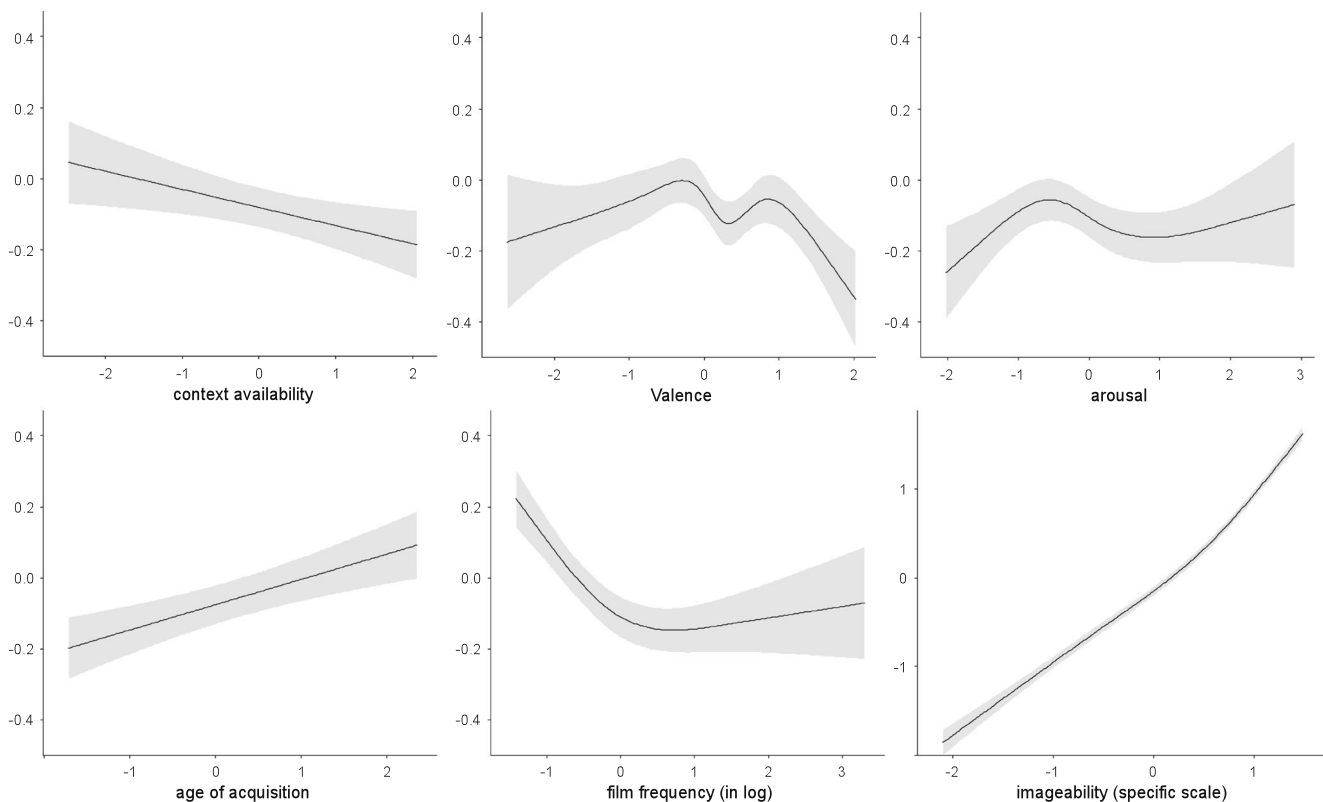
Fig. 5 Partial effects in the regression including imageability— significant effects only. Concreteness *z* scores are shown on the *y*-axis; standardized scores are used for all independent variables, on the abscissa; for imageability, a specific scale is used to make it possible to visualize all the predicted values

imageability is a composite measure (Bonin, Méot, Ferrand, & Roux, 2011; Dellantonio et al., 2014). According to Dellantonio et al., imageability measures not only the degree to which the meaning of words relies on external sensory information, but also the degree to which this is dependent on internal bodily-related sensory experience. Indeed, when the same forward procedure as described above was used to model the relationships between imageability and the other independent variables—with concreteness excluded—the R-square value was equal to .72, with partial effects roughly comparable to those that were found for concreteness when excluding imageability from the model (Fig. 4). Moreover, the R-square between the fitted values obtained for imageability and concreteness was .93. The latter finding strongly suggests that a common dimension underpins the two variables. However, it must be stressed that the explained variance was higher for imageability than for concreteness (.72 vs. .58), which suggests that potential sources of variance beyond this common dimension are more important for concreteness than for imageability.

When concreteness was included in order to predict imageability, a number of noteworthy results were observed. First, R^2 was now equal to .89—that is to say, .16 points more than in the model without concreteness—with the explained variance being mostly due to concreteness, $F(2, 1465) =$

1197.76, $p < .001$ [nonlinear, $F(1, 1465) = 30.54, p < .001$]. Second, variables that had weak effects in the first step of the analysis were no longer significant—that is, numbers of letters and homographs; valence, $F(1, 1465) < 1$; and arousal, $F(1, 1465) = 1.68, p > .1$. Third, the reliable effects—that is, context availability, $F(3, 1465) = 103.13, p < .001$ [nonlinear, $F(2, 1465) = 6.71, p < .001$]; word frequency, $F(2, 1465) = 47.3, p < .001$ [nonlinear, $F(1, 1465) = 29.93, p < .001$]; AoA, $F(1, 1465) = 169.81, p < .001$; and SER, $F(1, 1465) = 4.04, p < .05$—had roughly the same functional form (see Fig. 6) as those obtained in the model that did not include concreteness. One exception was word frequency, for which imageability was stable for low-frequency words and then decreased.

Soares et al. (2018) hypothesized that imageability might be a more valenced construct, whereas concreteness would be a more arousable construct, thus supporting the idea that the two variables are distinct and should not be used interchangeably. However, the present findings do not provide clear support for this proposal. In the multivariate model that included imageability, we found that both valence and arousal were significant predictors of concreteness scores; these variables had, however, no significant effects on imageability when concreteness was controlled for.

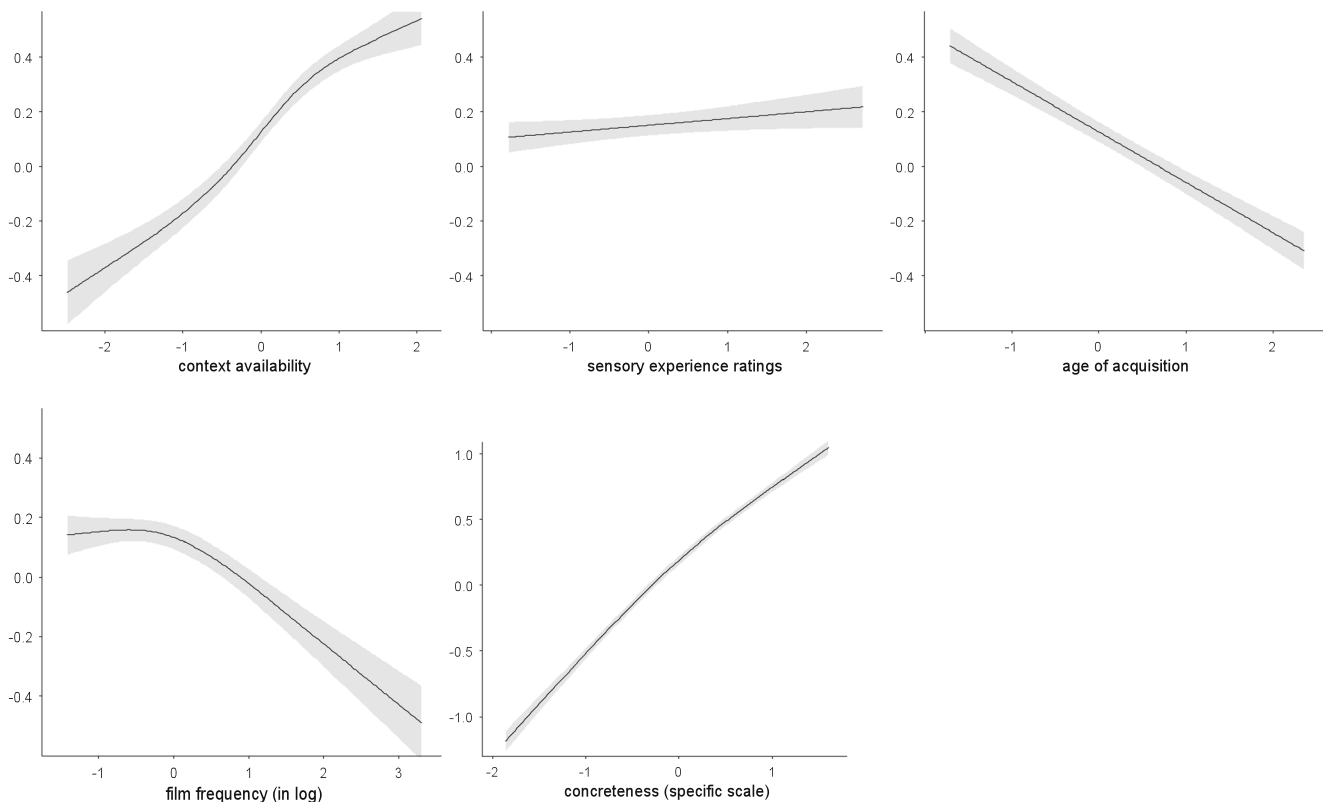To summarize, the same significant effects with very close functional relationships were observed in the

Fig. 6 Partial effects in the regression including concreteness— significant effects only. Imageability z scores are shown on the y-axis; standardized scores are used for all independent variables, on the abscissa. A special scale is used for concreteness to make it possible to visualize all the predicted values

analyses of concreteness and imageability (without considering imageability and concreteness, respectively). Moreover, the fitted values of both variables were very highly correlated. These findings suggest that a strong common dimension underlies both imageability and concreteness. Our analyses are in line with the suggestion that both variables are able to detect analogous properties (Dellantonio et al., 2014). However, this does not mean that the two variables are the same (Dellantonio et al., 2014). Thus, one implication is that researchers should control items for both variables and not for only one of the two. Imageability and concreteness are complex constructs, and it is already clear that future work will be needed to achieve a full understanding of what exactly each of them measures. In particular, when all psycholinguistic variables are taken into account, more variance remains to be explained for concreteness than for imageability.

Influence of concreteness in lexical decision, word naming, and progressive demasking (without inclusion of the emotional variables)

As we reviewed in the introduction, the *classical views* on concreteness effects are the dual-coding (e.g., Paivio,

2007) and the context-availability view (Schwanenflugel et al., 1988) (these accounts were initially evaluated without considering emotional variables). The dual-coding view predicts that concreteness effects should vanish when imageability values are controlled for. In contrast, according to the context-availability view, concreteness effects should vanish when the number of contexts available for words is controlled for. We therefore first analyzed the influence of concreteness in lexical decision, word naming, and progressive demasking without taking emotional variables into account. In lexical decision, Kousta et al. (2011) found that neither the dual-coding nor the context-availability views fully explained concreteness effects because when imageability and context availability ratings were taken into account, abstract words were processed faster in lexical decision than concrete words. According to Kousta et al., the processing advantage of abstract over concrete words would be due to the more numerous affective associations of abstract words. If the "abstractness" effect in lexical-decision times found by Kousta et al. can be explained by differences in emotional information between abstract and concrete words, it should be eliminated once affective associations are entered in the model. We therefore conducted another set of analyses (see below) to assess the influence

Table 5　Correlations between word recognition performance in lexical decision, word naming, and progressive demasking and the various psycholinguistic variables

	Lexical Decision (Errors)	Word Naming	Progressive Demasking	AoA	Number of Letters	Old20	Concreteness	SER	Image-ability	Context Availability	Log Subtitle Frequency
Lexical decision (RT)	.68	.34	.56	.57	.34	.16	.05 ns	−.20	−.12	−.26	−.61
Lexical decision (Errors)		.22	.40	.48	.04 ns	.03 ns	−.02 ns	−.20	−.18	−.31	−.44
Word naming			.22	.23	.23	.19	*.07*	−.11	.00 ns	−.06	−.22
Progressive demasking (RTs)				.31	.38	.12	.01 ns	−.13	**−.09**	−.16	−.34

Except for nonsignificant results and results in italics and in bold, for which $p < .05$ if $|r| > .05$, and $p < .01$ if $|r| > .07$, all tests were significant at $p < .001$. Correlations with independent variables were classified from the highest to the lowest on the basis of the lexical-decision RTs. AoA = age of acquisition; OLD20 = orthographic Levenshtein distance to the 20 nearest neighbors; SER = sensory experience ratings

of concreteness not only in lexical decision, but also in word naming, and progressive demasking when valence and arousal are taken into account.

Ferrand et al. (2011) analyzed reaction times for 1,482 words in lexical decision, word naming, and progressive demasking. Error rates were also analyzed in lexical decision. We included these variables as dependent variables in linear regressions including concreteness, imageability, context availability, SER, AoA, number of letters, orthographic neighborhood density (OLD20, which corresponds to the orthographic Levenshtein distance to the 20 nearest neighbors; Yarkoni, Balota, & Yap, 2008) and subtitle frequency as independent variables. In addition, because they account for a large amount of the variance in word naming, the articulatory features of initial phonemes were included. Words for which not all the scores were available were excluded from the analyses, leaving us with 1,478 words.

As can be seen in Table 5, the correlations of all dependent variables with concreteness were low, and only the correlation with word-naming times was significant, $p < .01$. The observation that the processing advantage for concrete words was not found to be reliably correlated in lexical decision is surprising, given that this effect has been considered to be a classical effect in the psycholinguistic literature. The lack of a significant correlation between lexical-decision times and concreteness was also observed by Guasch et al. (2016), whereas Kousta et al. (2011) reported a significant, but low, correlation on a large set of words. We conjectured that more recent studies that have used larger sets of words differ somewhat from earlier studies in terms of the diversity of words that have been taken into account, with the latter studies also including more limited types of words. It is clear that more in-depth analyses will be required in order to test this assumption. However, this was beyond the scope of the present article.

The procedure that was used to investigate the predictors of concreteness and imageability scores was also used here.

However, word frequency, which is known to have a nonlinear effect, was always entered first in the regression equation. Table 6 summarizes the results.[4]

Lexical decision　The overall R^2 in lexical-decision times was .59.

As is shown in Table 6 and Fig. 7, word frequency, $F(3, 1461) = 162.87$, $p < .001$ [nonlinear $F(2, 1461) = 61.77$, $p < .001$], and imageability, $F(3, 1461) = 13.73$, $p < .001$ [nonlinear $F(2, 1461) = 7.19$, $p < .001$], were the two most important predictors, and both of these were facilitatory effects, with decisions for more frequent/imageable words being faster than those for less frequent/imageable ones.

This was also the case for context availability, $F(1, 1461) = 12.81$, $p < .001$, with decisions for words associated with more contexts being faster than those for words with fewer contexts. By contrast, the effect of number of letters, $F(3, 1461) = 48.33$, $p < .001$ [nonlinear $F(2, 1461) = 3.45$, $p < .05$], was inhibitory (longer words yielded longer RTs than did shorter words), as was also the case for orthographic neighborhood (i.e., OLD20), $F(1, 1461) = 30.74$, $p < .001$, with decision times being longer for words with more neighbors than for words with fewer neighbors (note that the OLD20 scale is inverted; that is to say, words with more neighbors have lower OLD20 values, whereas higher values are associated with words with a more restricted neighborhood). Importantly, concreteness had a positive effect, $F(1, 1461) = 24.23$, $p < .001$, with less concrete (abstract) words yielding shorter RTs than more concrete words. This was also the case with AoA, $F(1, 1461) = 17.53$, $p < .001$: Decision times were faster for early-acquired than for late-acquired

_____

[4] It is worthy of note that excluding imageability from the equations led to the same significant effects in the analyses of naming RTs and lexical decision errors, with small differences in the forms of the partial effects. Concreteness, however, turned out to be unreliable in the RT analyses in the lexical decision and progressive demasking tasks.

Table 6 Nonlinear terms included for the different lexical tasks

	Lexical Decision (RT)	Word Naming	Progressive Demasking	Lexical Decision (Errors)
R^2 init.	.5412	.5491	.257	.3702
Frequency	.0435 (1 ; 4)	.0081 (1 ; 5)	.0452 (1 ; 4)	.0483 (1 ; 6)
Imageability	.0049 (2 ; 4)	.0020 (3 ; 4)	.0043 (3 ; 4)	.0026 (4 ; 4)
SER	.0018 (3 ; 4)			
Number of letters	.0019 (4 ; 4)		.0237 (2 ; 3)	
Old20		.0056 (2 ; 4)		
AoA				.0048 (2 ; 4)
Context availability				.0018 (3 ; 3)
R^2 final	.5932	.5648	.3302	.4277

R^2 init. = R-square with linear terms only. In each row, the first term = R^2 increase when nonlinear terms are included for the independent variable; in parentheses = rank order for entry in the equation and knots number. R^2 final = R-square with all terms (linear and nonlinear) included. AoA = age of acquisition; SER = sensory experience ratings; OLD20 = orthographic Levenshtein distance to the 20 nearest neighbors

words. Finally, the effect of SER was roughly flat, with few nonlinearities, $F(3, 1461) = 2.46$, $p < .10$ [nonlinear, $F(2, 1461) = 3.13$, $p < .05$].

As far as error rates are concerned, the whole set of independent variables explained 42.8% of the variance in lexical decision. Most of the independent variables found to be significant for lexical-decision times were also significant for error rates, with the effects being in the same direction (that is to say, facilitatory vs. inhibitory). This was the case for word frequency, $F(5, 1460) = 54.41$, $p < .001$ [nonlinear, $F(4, 1460) = 22.73$, $p < .001$]; context availability, $F(2, 1460) = 17.12$, $p < .001$ [nonlinear, $F(1, 1460) = 5.4$, $p < .05$]; AoA, $F(3, 1460) = 10.95$, $p < .001$ [nonlinear, $F(2, 1460) = 6.33$, $p < .01$]; imageability, $F(3, 1460) = 7.57$, $p < .001$ [nonlinear, $F(2, 1460) = 3.38$, $p < .05$]; and orthographic neighborhood, $F(1, 1460) = 12.92$, $p < .001$. There were three noticeable exceptions: concreteness, word length, and SER. Even though the influence of concreteness was significant, $F(1, 1460) = 24.95$, $p < .001$, it was in the direction opposite to that found in the RT analysis; that is to say, concrete words yielded more errors than did abstract words. The effect of length was reliable, $F(1, 1460) = 6.37$, $p < .05$, and its influence also ran in the direction opposite that found for RTs; that is to say, there were fewer errors for longer words. Finally, the effect of SER was not significant, $F(1, 1460) = 2.52$, $p > .1$.

Word naming The overall R^2 for word-naming times was $R^2 = .565$. The independent variables other than the initial phoneme characteristics accounted for 11.8% of the variance when these characteristics were controlled for. As is shown in Fig. 8, there were significant positive effects of number of letters, $F(1, 1452) = 41.74$, $p < .001$; orthographic neighborhood

(OLD20), $F(3, 1452) = 10.09$, $p < .001$ [nonlinear, $F(2, 1452) = 9.66$, $p < .001$]; AoA, $F(1, 1452) = 21.76$, $p < .001$; and concreteness, $F(1, 1452) = 7.94$, $p < .01$. Thus, words with fewer letters, acquired early in life, or with more neighbors were named aloud faster than words with more letters, that were late-acquired, or that had fewer neighbors, whereas longer RTs were observed for more concrete words. The effect of word frequency was significant and clearly nonlinear, $F(4, 1452) = 7.18$, $p < .001$ [nonlinear, $F(3, 1452) = 7.39$, $p < .001$], with RTs decreasing in the left part of the scale and increasing in the remaining part of the scale. Imageability had a negative effect, $F(3, 1452) = 2.71$, $p < .05$ [nonlinear, $F(2, 1452) = 3.26$, $p < .05$], with high-imageability words being named faster than low-imageability words. As in lexical decision, inhibitory and facilitatory effects of concreteness and imageability, respectively, were found (Fig. 8). Finally, the effects of context availability, $F(1, 1452) = 0.00$, $p > .10$, and SER, $F(1, 1452) = 0.01$, $p > .10$, were not significant.

Progressive demasking Thirty-three percent of the variance of the RTs was explained in the progressive-demasking task. Significant effects were found for number of letters, with RTs being longer for words having more letters, $F(2, 1464) = 138.7$, $p < .001$ [nonlinear, $F(1, 1464) = 49.32$, $p < .001$]; word frequency, for which there was a nearly linear decrease in RTs down to the mean frequency, followed by a stabilization, $F(3, 1464) = 44.67$, $p < .001$ [nonlinear, $F(2, 1464) = 36.92$, $p < .001$]; orthographic neighborhood, with decreasing RTs being associated with words with fewer neighbors, $F(1, 1464) = 28.45$, $p < .001$; imageability, $F(3, 1464) = 4.88$, $p < .01$ [nonlinear, $F(2, 1464) = 4.67$, $p < .01$]; and concreteness, $F(1, 1464) = 4.55$, $p < .05$. As in the other tasks, the effect of imageability was facilitatory, whereas the reverse was observed for concreteness (see Fig. 9); that is to say, RTs were shorter for higher-imageability and abstract words. The other
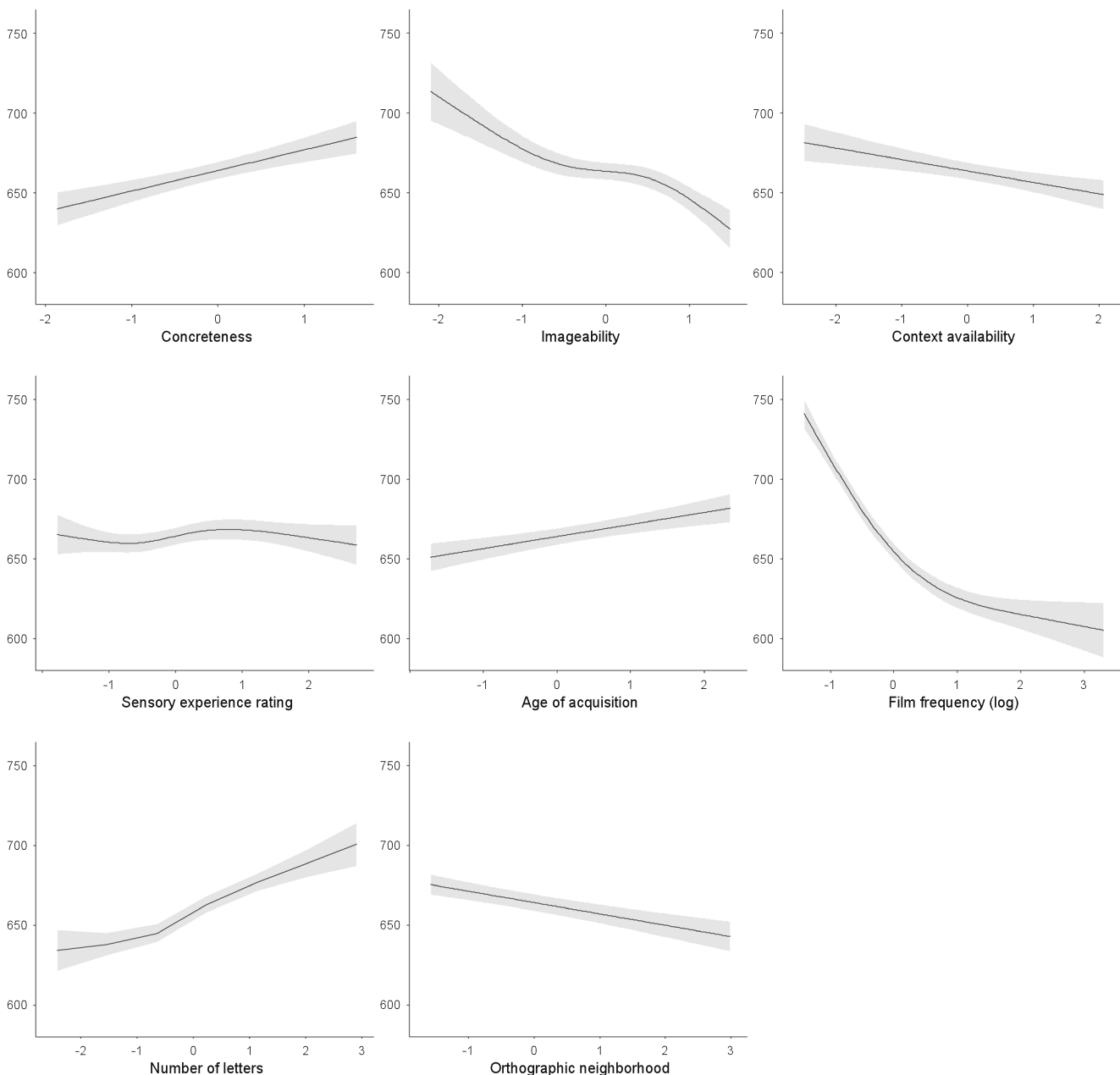
Fig. 7 Significant partial effects in the analysis of lexical-decision reaction times (*y*-axis). For the purposes of comparison, *z* scores are used for the independent variables, shown on the abscissa

independent variables had no significant effect: AoA, $F(1, 1464) = 0.77$, $p > .10$; context availability, $F(1, 1464) = 0.01$, $p > .10$; SER, $F(1, 1464) = 0.38$, $p > .10$.

Table 7 provides a summary of the directions of the effects of the different variables in the three lexical tasks. The directions of the effects were the same in the three tasks in the analyses run with or without valence and arousal included, except for SER, whose effect in lexical decision was not significant in the analyses that included valence and arousal. Thus, only the findings without valence and arousal included are reported in Table 7.

Influence of concreteness in lexical decision, word naming, and progressive demasking when valence and arousal were included

We used the same procedure described above to investigate the effects of valence and arousal. The order of entry of the variables and the numbers of knots are summarized in Table 8. In the analysis of lexical-decision times, the inclusion of valence and arousal in the regression equation resulted in an R^2 increase of only .0057.

Although it was weak, a significant nonlinear effect of valence was found, $F(2, 1460) = 5.66$, $p < .01$ [nonlinear,
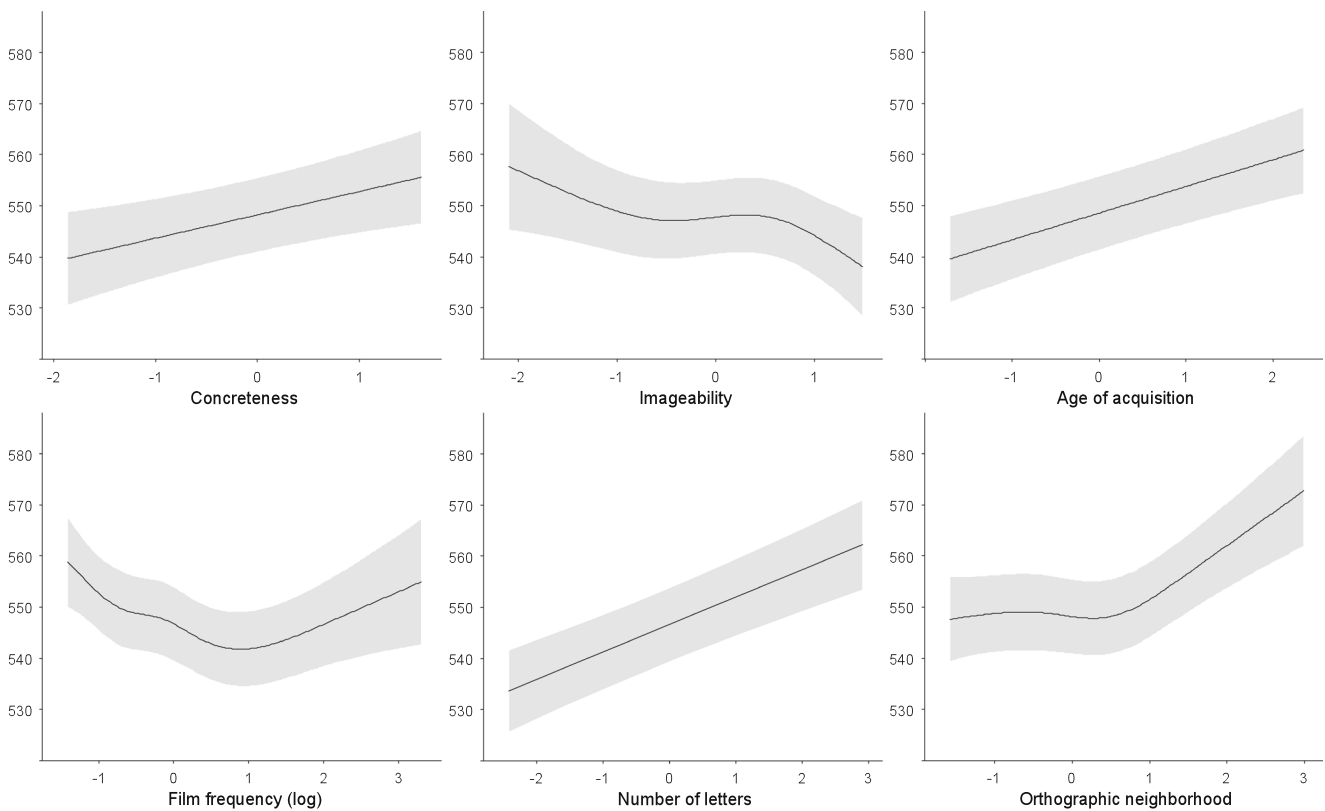
Fig. 8 Significant partial effects in the analysis of RTs (*y*-axis) in the word-naming task. For the purposes of comparison, *z* scores are used for the independent variables, shown on the abscissa
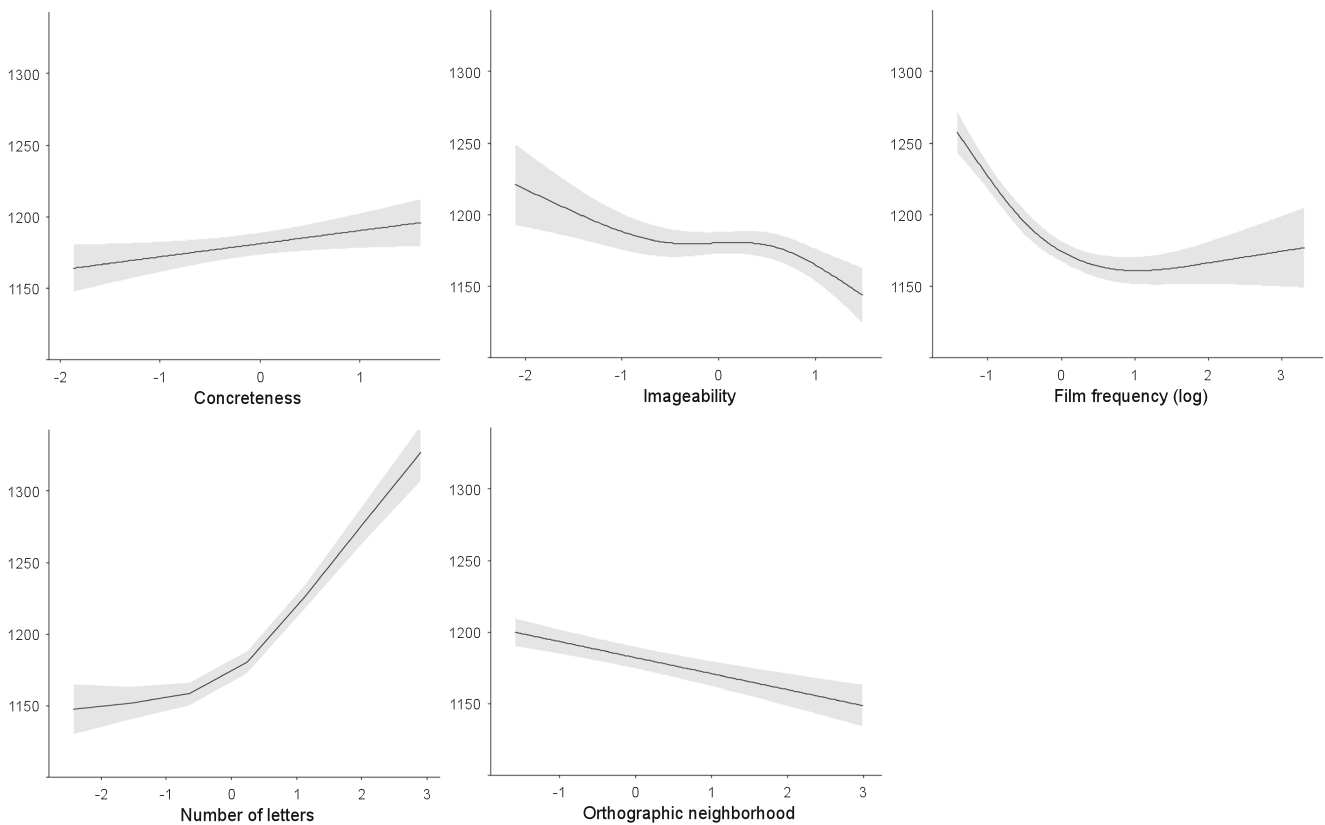

Fig. 9 Significant partial effects in the analysis of RTs (*y*-axis) in the progressive-demasking task. For the purposes of comparison, *z* scores are used for the independent variables, shown on the abscissa

Table 7 Summary of the directions of effects as a function of the different variables in the three lexical tasks (lexical decision, word naming, and progressive demasking)

	Lexical Decision	Word Naming	Progressive Demasking
Concreteness	I	I	I
Imageability	F	F	F
Context availability	F	ns	ns
SER	Flat	ns	ns
AoA	I	I	ns
Subtitle frequency	F	F	F
OLD20	I	F	I
Number of letters	I	I	I

F = facilitatory effect, I = inhibitory effect, flat = flat effect; ns = nonsignificant effect. SER = sensory experience ratings; AoA = age of acquisition; OLD20 = orthographic Levenshtein distance to the 20 nearest neighbors. For all variables except OLD20, which has an inverted scale, higher values on a variable indicate that the corresponding attribute is stronger; for example, higher scores on the concreteness variable mean that the words are more concrete. As a result, the rates of changes on the dependent variables are negative for facilitatory effects and positive for inhibitory effects

$F(1, 1460) = 10.84$, $p < .001$] (Fig. 10). The effect of arousal was marginally significant, $F(1, 1460) = 3.3$, $p < .1$. Except for SER, for which both the linear and nonlinear terms were not reliable, $F(1, 1460) = 0.11$, $p > .10$, the effects of the remaining independent variables were very close to those found in the analysis that did not include valence and arousal. In particular, the inhibitory and facilitatory effects of concreteness and imageability, respectively, were again reliable: concreteness, $F(1, 1460) = 26.81$, $p < .001$; imageability, $F(3, 1460) = 13.52$, $p < .001$ [nonlinear, $F(1, 1460) = 8.92$, $p < .001$]. In the error analysis for lexical decision, with an R^2 equal to .4311, the percentage of explained variance was also increased slightly by the inclusion of emotional variables ($\Delta R^2 = .0034$). For nonemotional variables, we found roughly the same pattern of effects as was reported for the regression model that did not include valence and arousal. Finally, valence had a significant but weak nonlinear

effect, $F(2, 1457) = 3.37$, $p < .05$ [nonlinear, $F(1, 1457) = 5.12$, $p < .05$], whereas the effect of arousal was not significant, $F(1, 1449) = 0.09$, $p > .10$.

In word naming, the R^2 increase was also small (.0014), and neither valence nor arousal reached significance: valence, $F(1, 1449) = 1.67$, $p > .10$; arousal, $F(1, 1449) = 0.62$, $p > .10$. The pattern of findings for the other variables was also very similar to that obtained in the analysis without emotional variables. The same result was obtained for progressive demasking ($\Delta R^2 = .0025$).

In summary, a positive effect of concreteness was observed on lexical-decision times (Table 7). This effect was not significant when imageability was excluded from the model (see note 4). It was also observed in both word naming and progressive demasking. Overall, these findings accord with Kousta et al.'s (2011) lexical-decision findings (their Exps. 1 and 3). Also in accordance with

Table 8 Nonlinear terms included for the different tasks when emotional variables are included

	Lexical Decision (RT)	Word Naming	Progressive Demasking	Lexical Decision (Errors)
R^2 init.	.5442	.5496	.2594	.3709
Frequency	.0449 (1 ; 4)	.0089 (1 ; 6)	.0459 (1 ; 4)	.049 (1 ; 6)
Imageability	.0048 (2 ; 4)	.002 (3 ; 4)	.0039 (3 ; 4)	.0026 (5 ; 4)
Valence	.0028 (3 ; 3)			.0018 (4 ; 3)
SER				
Number of letters	.0021 (4 ; 4)		.0235 (2 ; 3)	
Old20		.0057 (2 ; 4)		
AoA				.0048 (2 ; 4)
Context availability				.002 (3 ; 3)
R^2 final	.5989	.5662	.3327	.4311

R^2 init. = R-square with linear terms only. In each row, first term = R^2 increase when the independent variables are included; in parentheses = rank order for entry in the equation and knots number. R^2 final = R-square with all terms (linear and nonlinear) included. AoA = age of acquisition; SER = sensory experience ratings; OLD20 = orthographic Levenshtein distance to the 20 nearest neighbors
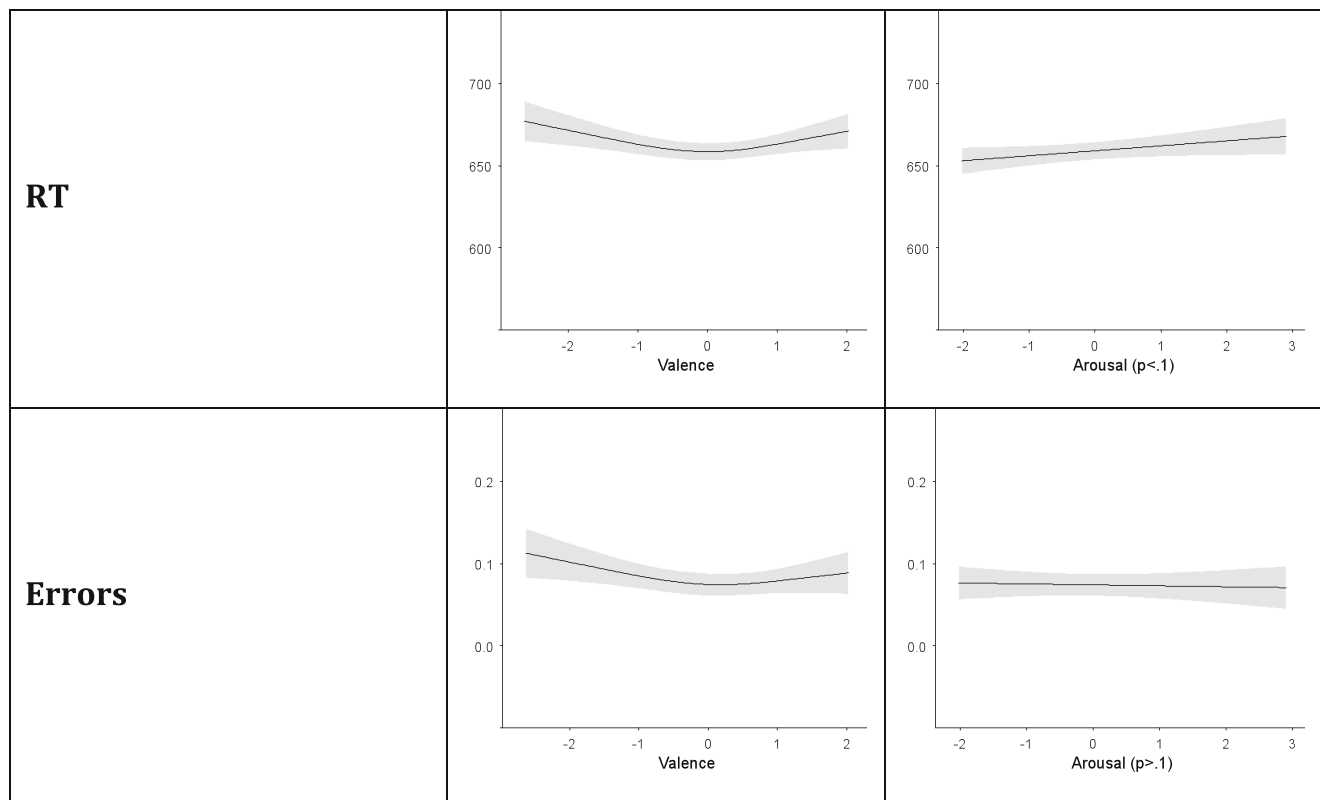
Fig. 10 Effects of emotional variables in the lexical-decision task (*y*-axis). For comparison purposes, *z* scores are used for the independent variables, shown on the abscissa

those authors, the concreteness effect on lexical-decision errors was not facilitatory in the present study (and was observed whether or not imageability was included in the regression model). In their Experiment 2, Kousta et al. did not find a reliable effect of concreteness for neutral words—that is, words with ratings located in the center of the valence scale. To compare this finding with our present data, the interaction terms between concreteness and valence were added in the regression models of RTs and error rates. Valence was included only with a linear term and (as in the lexical-decision task) with three knots splines. In none of our analyses did we find significant interactions between concreteness and valence, which suggests that the effect of concreteness is homogeneous across the valence scale. It is important to note that when we followed Kousta et al.'s analyses that were limited to the words with valence ratings located in the center of the valence scale (= neutral words)—that is, at the center of the scale ± 9.375% of the scale amplitude, thus corresponding to 613 words in the present study—the results suggested the same property (prior to the analyses, the distributions of imageability and concreteness scores were plotted in order to check that the neutral words did indeed span the entire range of the concreteness and imageability scales). For neutral words, both imageability and concreteness were significant, with facilitatory and inhibitory effects being observed for both RTs and errors, respectively. In addition, concreteness also had an inhibitory effect on word naming, whereas the effect of imageability was not reliable. Finally, there were no reliable effects of these two variables in the analysis of progressive-demasking RTs. Valence and arousal had significant effects in none of the lexical tasks.

Finally, one aspect worth noting concerns the influence of SER. Contrary to the claim made by Connell and Lynott (2012) about the strength of perceptual experience (which can be indexed by SER), we did not find that this variable made a reliable contribution in lexical decision, word naming, or progressive demasking when either concreteness or imageability was taken into account. (In effect, Connell & Lynott, 2012, found that the strength of perceptual experience was more important than concreteness and imageability in accounting for variance in both lexical decision and naming.) In a previous study, Bonin et al. (2015) found that SER was a reliable predictor of French lexical-decision times (obtained from Ferrand et al., 2011; Ferrand et al., 2010). However, in the present study, this variable was reliable for neither word naming nor progressive demasking. It must be stressed, however, that since concreteness norms were not available at the time the Bonin et al. (2015) study was conducted, they could not be taken into account in their analyses.

From a general standpoint, concreteness, imageability, valence, arousal, and SER are all semantic variables whose (specific and combined) influences still need to be explored more thoroughly in different lexical-processing tasks (see Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012, for a study of this kind). Critically, thanks to the collection of psycholinguistic norms on words, study of this type is now possible.

Conclusions

The present study has provided a database with concreteness ratings for 1,659 French words, together with ratings for context availability, valence, and arousal that supplement the norms that were already available in French for this set of words (e.g., imageability, SER).[5] The norms are available in an Excel file as supplementary materials, and they will be very useful to researchers when selecting materials for designing experiments in French on language and memory. We have focused on the relationships between concreteness and other psycholinguistic variables, and taken as a whole, several aspects of the present findings are in agreement with previous findings reported in other languages. Importantly, our findings generally support the embodied view of concreteness championed by Kousta et al. (2011) and Vigliocco et al. (2014). Indeed, the two types of words vary as a function of their reliance on perceptual information, internal information, and linguistic information, with abstract words having more affective/internal and linguistic associations than concrete words, whereas concrete words have more sensory/perceptual associations than abstract words. Finally, the analyses of the performance in lexical decision replicated a key finding of Kousta et al.—namely, that when both imageability and context availability scores are taken into account, abstract words are processed more quickly than concrete words. A growing body of research now provides evidence that the processing of concrete versus abstract words cannot be exhaustively accounted for by the two traditional accounts of concreteness effects, namely the dual-code view (e.g., Paivio, 1971, 1986) and the context-availability view (e.g., Schwanenflugel et al., 1988). These two accounts have long assumed that the representations of concrete words are richer (in terms of imagery features or contexts) than those of abstract words. However, Kousta et al.'s findings, the present findings, and other findings (e.g., Pexman, Hargreaves, Edwards, Henry, &

Goodyear, 2007) are more in line with an embodied account of concreteness effects in lexical processing.

---

[5] As was suggested by an anonymous reviewer, future studies could use the technique described in Brysbaert, Stevens, De Deyne, Voorspoels, and Storms's (2014a) study in order to collect norms for many more words in French (e.g., concreteness norms for 30,000 Dutch words in Brysbaert, Stevens, et al.'s, 2014a, study).

References

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers*, *31*, 578–602. https://doi.org/10.3758/BF03200738

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, *92*, 265–280.

Barca, L., Burani, C., & Arduino, L.S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments & Computers*, *34*, 424–434. https://doi.org/10.3758/BF03195471

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–609, disc. 609–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*, 84–91. https://doi.org/10.1016/S1364-6613(02)00029-3

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*, 905–917.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73–79. https://doi.org/10.3758/BF03195349

Bonin, P., Gelin, M., & Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Memory & Cognition*, *42*, 370–382. https://doi.org/10.3758/s13421-013-0368-8

Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M.-C. (2003). Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots [Concreteness, imageability, subjective frequency, and emotional valence norms for 866 words]. *L'Année Psychologique*, *104*, 655–694. https://doi.org/10.3406/psy.2003.29658

Bonin, P., Méot, A., Boucheix, J. M., & Bugaiska, A. (2018). Psycholinguistic norms for 320 fixed expressions (idioms and proverbs) in French. *Quarterly Journal of Experimental Psychology*. Advance online publication. https://doi.org/10.1080/17470218.2017.1310269

Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods*, *45*, 1259–1271. https://doi.org/10.3758/s13428-013-0331-4

Bonin, P., Méot, A., Ferrand, L., & Bugaiska, A. (2015). Sensory experience ratings (SERs) for 1,659 French words: Relationships with

other psycholinguistic variables and visual word recognition. *Behavior Research Methods*, *47*, 813–825. https://doi.org/10.3758/s13428-014-0503-x

Bonin, P., Méot, A., Ferrand, L., & Roux, S. (2011). L'imageabilité : Normes et relations avec d'autres variables psycholinguistiques [Imageability: Norms and relationships with other psycholinguistic variables]. *L'Année Psychologique*, *111*, 327–357. https://doi.org/10.4074/S0003503311002041

Bonin, P., Perret, C., Méot, A., Ferrand, L., & Mermillod, M. (2008). Psycholinguistic norms and face naming times for photographs of celebrities in French. *Behavior Research Methods*, *40*, 137–146. https://doi.org/10.3758/BRM.40.1.137

Bonner, M. F., Vesely, L., Price, C., Anderson, C., Richmond, L., Farag, C., … Grossman, M. (2009). Reversal of the concreteness effect in semantic dementia. *Cognitive Neuropsychology*, *26*, 568–579.

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

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014a). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, *150*, 80–84.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014b). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904–911. https://doi.org/10.3758/s13428-013-0403-5

Cameirão, M. L., & Vicente, S. G. (2010). Age-of-acquisition norms for a set of 1,749 Portuguese words. *Behavior Research Methods*, *42*, 474–480. https://doi.org/10.3758/BRM.42.2.474

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2014). How are affective word ratings related to lexico-semantic properties? Evidence from the Sussex Affective Word List. *Applied Psycholinguistics*, *35*, 313–331. https://doi.org/10.1017/S0142716412000409

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*, 452–465. https://doi.org/10.1016/j.cognition.2012.07.010

de Groot, A. M. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 824–845. https://doi.org/10.1037/0278-7393.15.5.824

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14. https://doi.org/10.1016/j.brainres.2005.12.115

Della Rosa, P. A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract–concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods*, *42*, 1042–1048. https://doi.org/10.3758/BRM.42.4.1042

Dellantonio, S., Mulatti, C., Pastore, L., & Job, R. (2014). Measuring inconsistencies can lead you forward: Imageability and the x-ception theory. *Frontiers in Psychology*, *5*, 708.

Ferrand, L., Bonin, P., Méot, A., Augustinova, M., New, B., Pallier, C., & Brysbaert, M. (2008). Age-of-acquisition and subjective frequency estimates for all generally known monosyllabic French words and their relation with other psycholinguistic variables. *Behavior Research Methods*, *40*, 1049–1054. https://doi.org/10.3758/BRM.40.4.1049

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., … Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology*, *2*, 306. https://doi.org/10.3389/fpsyg.2011.00306

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., … Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*, 488–496. https://doi.org/10.3758/BRM.42.2.488

Ferré, P., Guasch, M., Moldovan, C., & Sánchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods*, *44*, 395–403. https://doi.org/10.3758/s13428-011-0165-x

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage*, *32*, 1413–1421.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*, 375–399.

Ghasisin, L., Yadegari, F., Rahgozar, M., Nazari, A., & Rastegarianzade, N. (2015). A new set of 272 pictures for psycholinguistic studies: Persian norms for name agreement, image agreement, conceptual familiarity, visual complexity, and age of acquisition. *Behavior Research Methods*, *47*, 1148–1158. https://doi.org/10.3758/s13428-014-0537-0

Gilet, A.-L., Grühn, D., Studer, J., & Labouvie-Vief, G. (2012). Valence, arousal, and imagery ratings for 835 French attributes by young, middle-aged, and older adults: The French Emotional Evaluation List (FEEL). *Revue Européenne de Psychologie Appliquée*, *62*, 173–181. https://doi.org/10.1016/j.erap.2012.03.003

Gobin, P., Camblats, A.-M., Faurous, W., & Mathey, S. (2017). Une base de l'émotionalité (valence, arousal, catégories) de 1286 mots français selon l'âge (EMA) [A base of emotionality (valence, arousal, category) of 1,286 French words according to age (EMA)]. *Revue Européenne de Psychologie Appliquée*, *67*, 25–42.

Guasch, M., Ferré, P., & Fraga, I. (2016). Spanish norms for affective and lexico-semantic variables for 1,400 words. *Behavior Research Methods*, *48*, 1358–1369. https://doi.org/10.3758/s13428-015-0684-y

Hinojosa, J. A., Rincón-Pérez, I., Romero-Ferreiro, M. V., Martínez-García, N., Villalba-García, C., Montoro, P. R., & Pozo, M. A. (2016). The Madrid Affective Database for Spanish (MADS): Ratings of dominance, familiarity, subjective age of acquisition and sensory experience. *PLoS ONE*, *11*, e0155866. https://doi.org/10.1371/journal.pone.0155866

Juhasz, B. J., & Yap, M. J. (2013). Sensory experience ratings for over 5,000 mono- and disyllabic words. *Behavior Research Methods*, *45*, 160–168. https://doi.org/10.3758/s13428-012-0242-9

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, *140*, 14–34. https://doi.org/10.1037/a0021446

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 92–107. https://doi.org/10.1037/0278-7393.12.1.92

Krosnick, J. A., & Presser, S. (2010). Questionnaire design. In J. D. Wright & P. V. Marsden (Eds.), Handbook of survey research (2nd ed.). West Yorkshire: Emerald Group.

Kuperman, V., Stadthagen-Gonzales, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*, 978–990. https://doi.org/10.3758/s13428-012-0210-4

Masterson, J., & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word frequency, familiarity and age-of-acquisition. *Journal of Neurolinguistics*, *11*, 331–354.

Monnier, C., & Syssau, A. (2014). Affective norms for French words (FAN). *Behavior Research Methods*, *46*, 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW)

for Italian. *Behavior Research Methods*, *46*, 887–903. https://doi. org/10.3758/s13428-013-0405-3

Nairne, J. S. (2010). Adaptive memory: Evolutionary constraints on remembering. In B. H. Ross (Ed.), The psychology of learning and motivation (Vol. 53), pp. 1–32. Burlington: Academic Press.

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N. S., Cogdill, M., & LeBreton, J. M. (2013). Adaptive memory: The mnemonic value of animacy. *Psychological Science*, *24*, 2099–2105.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart & Winston.

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287. https://doi.org/10. 1037/h0084295

Paivio, A. (2007). Mind and its evolution: A dual coding theoretical approach. Mahwah: Erlbaum.

Paivio, A., Yuille, J. C., & Smythe, P. C. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology*, *20*, 362–377.

Papagno, C., Capasso, R., Zerboni, H., & Miceli, G. (2007). A reverse concreteness effect in a subject with semantic dementia. *Brain and Language*, *103*, 90–91.

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, *19*, 1407– 1419. https://doi.org/10.1162/jocn.2007.19.8.1407

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high and low imageability nouns: Analyses and theoretical implications. *Cognitive Science*, *31*, 1–12.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192. https://doi.org/10.1146/ annurev.neuro.27.070203.144230

Roediger, H. L., III (2012). Psychology's woes and a partial cure: The value of replication. *Observer*. Retrieved March 16, 2012, from http://tinyurl.com/d4lfnwu

Roxbury, T., McMahon, K., & Copland, D. A. (2014). An fMRI study of concreteness effects in spoken word recognition. *Behavioral and Brain Functions*, *10*, 34. https://doi.org/10.1186/1744-9081-10-34

Schwanenflugel, P. J. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), The psychology of word meanings (pp. 223–250). Hillsdale: Erlbaum.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*, 499–520. https://doi. org/10.1016/0749-596X(88)90022-8

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *9*, 82–102. https://doi.org/10.1037/0278-7393.9.1.82

Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, *24*, 114–126.

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

Simon, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, *9*, 76–80.

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*, 174–215. https://doi.org/10. 1037/0278-7393.6.2.174

Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2018). The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*, *49*, 1065–1081. https://doi.org/10. 3758/s13428-016-0767-4

ter Doest, L., & Semin, G. R. (2005). Retrieval contexts and the concreteness effect: Dissociations in memory for concrete and abstract words. *European Journal of Cognitive Psychology*, *17*, 859–881.

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, *24*, 1767– 1777. https://doi.org/10.1093/cercor/bht025

Vigliocco, G., Kousta, S. T., Vinson, D. P., Andrews, M., & Del Campo, E. (2013). The representation of abstract words: What matters? Reply to Paivio's (2013) comment on Kousta et al. (2011). *Journal of Experimental Psychology: General*, *142*, 288–291.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*, 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Yao, Z., Wu, J., Zhang, Y., & Wang, Z. (2018). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1, 100 Chinese words. *Behavior Research Methods*, *49*, 1374–1385. 10.3758/s13428-016-0793-2

Yap, M. J., & Pexman, P. M. (2016). Semantic richness effects in syntactic classification: The role of feedback. *Frontiers in Psychology*, *7*, 1394. https://doi.org/10.3389/fpsyg.2016.01394

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 72.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*, 971–979. https://doi.org/10.3758/PBR.15.5.971