

# ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words

David S. Schmidtke · Tobias Schröder ·  
Arthur M. Jacobs · Markus Conrad

© Psychonomic Society, Inc. 2014

**Abstract** We present the German adaptation of the Affective Norms for English Words (ANEW; Bradley & Lang in Technical Report No. C-1. Gainesville: University of Florida, Center for Research in Psychophysiology). A total of 1,003 Words—German translations of the ANEW material—were rated on a total of six dimensions: The classic ratings of valence, arousal, and dominance (as in the ANEW corpus) were extended with additional arousal ratings using a slightly different scale (see BAWL: Vö et al. in Behavior Research Methods 41: 531–538, 2009; Vö, Jacobs, & Conrad in Behavior Research Methods 38: 606–609, 2006), along with ratings of imageability and potency. Measures of several objective psycholinguistic variables (different types of word frequency counts, grammatical class, number of letters, number of syllables, and number of orthographic neighbors) for the words were also added, so as to further facilitate the use of this new database in psycholinguistic research. These norms can be downloaded as supplemental materials with this article.

**Electronic supplementary material** The online version of this article (doi:10.3758/s13428-013-0426-y) contains supplementary material, which is available to authorized users.

D. S. Schmidtke (✉) · T. Schröder · A. M. Jacobs · M. Conrad  
Department of General Psychology, Freie Universität Berlin,  
Habelschwerdter Allee 45, 14195 Berlin, Germany  
e-mail: david.schmidtke@fu-berlin.de

T. Schröder  
University of Waterloo, Waterloo, ON, Canada

A. M. Jacobs  
Dahlem Institute for Neuroimaging of Emotion, Freie Universität  
Berlin, Berlin, Germany

M. Conrad  
Universidad de La Laguna, La Laguna, Spain

**Keywords** Emotion · Valence · Arousal · Dominance ·  
Affective dictionary

Abundant evidence illustrates the effects of the affective features of words or text on language processing. In the case of visual word recognition, such effects have been shown for response latencies (Briesemeister, Kuchinke, & Jacobs, 2011b; Huckauf, Heller, & Gouzouli-Mayfrank, 2003; Kousta, Vinson, & Vigliocco, 2009; Vö, Jacobs, & Conrad, 2006), pupil dilation (Kuchinke, Vö, Hofmann, & Jacobs, 2007; Vö et al., 2008), event-related potentials/ERPs (e.g., Conrad, Recio, & Jacobs, 2011; Hofmann, Kuchinke, Tamm, Vö, & Jacobs, 2009; Kayser, Fong, Tenke, & Bruder, 2003; Kissler, Herbert, Peyk, & Junghofer, 2007; Ponz et al., 2013; Schacht & Sommer, 2009; Trauer, Andersen, Kotz, & Müller, 2012; see Citron, 2012, for a review), transcranial magnetic stimulation/TMS (Weigand et al., 2013), or functional magnetic resonance imaging/fMRI (Grimm, Weigand, Kazzner, Jacobs, & Bajbouj, 2012; Hamann & Mao, 2002; Herbert et al., 2009; Kuchinke et al., 2005; Tabert et al., 2001). Regardless of continuously improving methodologies in neuroscience, the validity of such findings hinges on thoroughly balanced stimulus materials with the potential to induce the type of affect or emotion under scrutiny and to control for other potential sources of influence. Only extensive databases including ratings of many words on several emotional dimensions will allow the appropriate selection of stimulus materials capable of providing reliable answers to specific research questions like the existence of differential effects of different emotion dimensions. These are also very useful for theory development, since current computational models of word recognition still do not include emotional dimensions (e.g., Grainger & Jacobs, 1996; Hofmann, Kuchinke, Biemann, Tamm, & Jacobs, 2011; Perry, Ziegler, & Zorzi, 2007).

Thus, meeting a still-growing demand due to an increasing interest in emotion research, various databases containing

emotional evaluations have been provided for different types of stimuli, such as pictures (Lang, Bradley, & Cuthbert, 2008) or sounds (Bradley & Lang, 1999b), but also words in English (Bradley & Lang, 1999a; Heise, 2010; Whissell, Fournier, Pelland, Weir, & Makarec, 1986), German (Schröder, 2011; Vö et al., 2009; Vö et al., 2006), Spanish (Redondo, Fraga, Comesaña, & Perea, 2005; Redondo, Fraga, Padrón, & Comesaña, 2007), Portuguese (Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), French (Silva, Montant, Ponz, & Ziegler, 2012), or Finnish (Eilola & Havelka, 2010).

Among these, the Affective Norms for English Words (ANEW; Bradley & Lang, 1999a) represent the best-known affective dictionary, providing normative evaluative ratings for 1,034 words. The adaptation of this database to other languages enables the comparison of empirical findings on emotion processing across different language contexts. Accordingly, the ANEW has already been adapted for Spanish (Redondo et al., 2007) and European Portuguese (Soares et al., 2012). Here, we present an adaptation for German—the native language of three European countries, where a quickly growing amount of research on emotion processing has been published in recent years.

Note that the comparability between available databases providing emotion ratings for German words (Briesemeister, Kuchinke, & Jacobs, 2011a; Kanske & Kotz, 2010; Schröder, 2011; Vö et al., 2009; Vö et al., 2006) and the ANEW corpus is limited. First, the overlap in the respective word materials only amounts to 40 percent of the ANEW—when compared to the recent version of the most extensive German database, the Berlin Affective Wordlist (BAWL; Vö et al., 2009). Second, these German databases do not provide ratings on all dimensions listed in the ANEW, or they do so for—at least potentially—slightly different dimensions: Unlike the ANEW, the BAWL database (Vö et al., 2009; Vö et al., 2006) does not contain ratings of dominance, and the arousal dimension was operationalized in a different way. All of these factors limit potential comparisons of the emotional connotations of specific words or their respective effects on emotion processing in order to contrast German with other languages/cultures (see, e.g., Conrad et al., 2011; Ponz et al., 2013).

In contrast to discrete conceptualizations of emotion (e.g., Panksepp, 1998), for which affective evaluations are available for English (Stevenson, Mikels, & James, 2007) and German (Briesemeister et al., 2011a), a dimensional approach underlies the ANEW corpus and its adaptations in different languages. This position, prominently advocated by Wundt (1896) and Russell (1980), is based on the seminal studies of Osgood, Suci, and Tannenbaum (1957), who applied factor analyses to semantic differential judgments of words. They identified three dimensions, named “evaluation,” “potency,” and “activity,” which account for a major portion of variance. Applied specifically to emotional states, Mehrabian and Russell (1974) later termed these dimensions “pleasure,”

“arousal,” and “dominance”—in order of importance, as defined by the amount of variance explained. ANEW characterizes these three dimensions as “pleasure/valence” (ranging from *pleasant* to *unpleasant*), “arousal” (ranging from *calm* to *excited*), and “dominance/control” (ranging from *dominant* to *dominated*) (Bradley & Lang, 1999a).

Besides these variables, imageability is a further dimension important for understanding the relations between emotion and language. Recent studies have demonstrated a significant influence of words’ imageability on language processing (Altarriba & Bauer, 2004; Altarriba, Bauer, & Benvenuto, 1999; Huang, Lee, & Federmeier, 2010; Kanske & Kotz, 2007; Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). This variable should also provide a close match to the concreteness or abstractness of words—often proposed to determine the processing of emotion content (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Accordingly, unlike the ANEW, more recent normative databases, such as the Berlin Affective Word List (BAWL; Vö et al., 2009; Vö et al., 2006), provide imageability ratings together with valence and arousal ratings.

Therefore, our adaptation of the original ANEW corpus to German serves a twofold purpose: First, we wanted to improve the match between normative emotion databases in German and other languages in order to enable future cross-linguistic research on emotion. Second, we also aimed at further developing the general understanding of different emotion dimensions proposed so far across different databases—investigating the mutual relations between these dimensions, after combining them within one, single database.

Our German ANEW database was collected by means of the following steps:

- 1 *ANEWing the BAWL*: The ratings of valence, arousal, and imageability for the German translations of ANEW words contained in the Berlin Affective Word List (Vö et al., 2009) served as a basis to build upon.
- 2 We also added ratings of valence, arousal, and imageability for German translations of the ANEW words not contained in the BAWL (Vö et al., 2009).
- 3 Finally, we collected dominance ratings for all German translations.
- 4 *Contrasting alternative but closely related dimensions*: We re-collected arousal ratings for all of the German words, this time using scales and instructions closely matching the original ANEW procedure. The arousal ratings for the BAWL corpus were based on a slightly different scale and instructions, to emphasize the bipolar character of this dimension (see Vö et al., 2009; Vö et al., 2006).
- 5 We also collected ratings for all words on the dimension of “potency” (Mehrabian & Russell, 1974; Schröder, 2011). Though this dimension is closely related to dominance,

potency mainly differs in its independence from the raters' perspective. A more detailed account of these variables will be presented in the discussion of our methods for creating the database.

- 6 *Shaping up the database for comfortable use in psycholinguistic research:* In addition to these evaluative variables, language statistical measures and other variables of psycholinguistic relevance were added as objective indices for all of the words, in order to make the database most beneficial for further scientific research. These are word frequencies, taken from two recent extensive corpora representing different contexts of language use: the print-based corpus of the Leipzig Wortschatz Projekt (*Wortschatz Universität Leipzig*, 2013), including over 50 million words, and the SUBTLEX corpus (Brysbaert et al., 2011), with more than 25 million words taken from movie subtitles—together with grammatical class, number of letters, number of syllables, and number of orthographic neighbors.

## Method

### Study I: ANEWing the BAWL

#### 1. + 2. Expanding the BAWL ratings

The translation of the 1,034 words provided by the ANEW (Bradley & Lang, 1999a) resulted in a total of 1,003 German words, because some of the English words have the same German translation. Ratings for valence, arousal, and imageability for 400 of the words could be retrieved from the BAWL (Vö et al., 2009). Additional ratings for the remaining 603 German words were collected separately for each dimension.

*Procedure* Ratings for valence, arousal, and imageability were collected according to the procedure reported in the BAWL (Vö et al., 2009). Note that this procedure involves the use of the Self-Assessment Manikin (SAM; see Fig. 1) only for the arousal dimension, but not for valence.

The SAM, a nonverbal pictorial measure derived from the semantic differential scale developed by Mehrabian and Russell (1974) and adapted by Lang (1980), has often been used to assess the three dimensions of valence, arousal, and dominance. Each of these scales is represented by five figures. To facilitate scale comprehension, SAM pictures are normally accompanied by verbal anchors at the extreme ends of the scale. But note that the capacity of the SAM to adequately represent these constructs or dimensions of emotional content is a matter of debate (see Vö et al., 2009, and Vö et al., 2006, for discussions). This is why, when designing the data

collection for the BAWL, we had opted for a slightly different procedure—presented below—which we continued to use now to complete the ratings on valence and arousal for the 1,003 German words.

*Valence.* Following the procedure for obtaining BAWL ratings, participants were presented with the verbal anchors *positiv* (“positive”) and *negativ* (“negative”), defining the ends of a bipolar scale ranging from −3 to +3, with “neutral” (0) in the center.

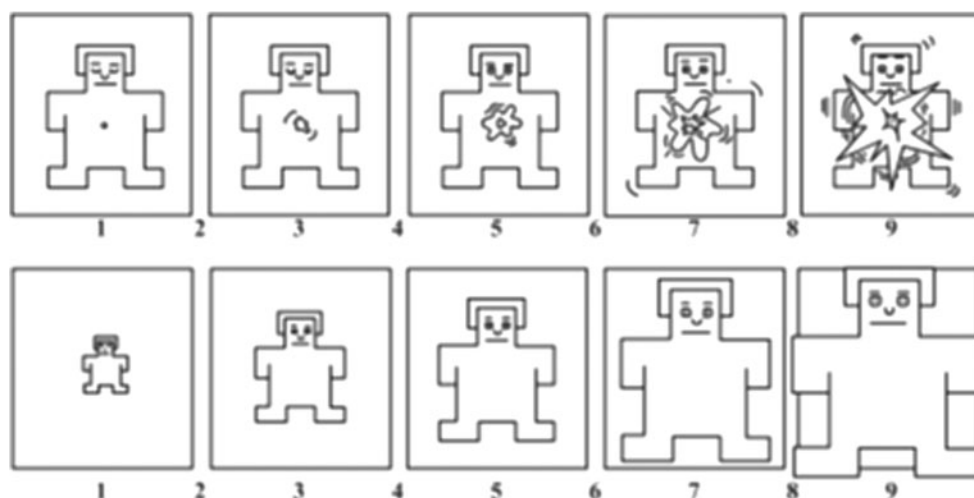
*Arousal (BAWL).* Unlike valence, the concept of *arousal* is difficult to represent in a purely verbal way. For this reason, we opted to use SAMs for data recollection in both the original BAWL study and the present study (see Fig. 1). But, unlike Bradley and Lang (1999a), we had participants rate arousal on a 5-point scale (rather than a 9-point scale) and only used the SAMs, but not the intervals between them, for the rating.

Probably more important than this difference between a 5- versus a 9-point scale might be the following change concerning the verbal anchors: The anchors for the extreme ends of the arousal dimension, as implemented in the ANEW (Bradley & Lang, 1999a), were *calm* versus *excited*. Although the authors' theoretical account of this scale stressed its bipolar character—ranging from a relaxing extreme to an exciting extreme, with a neutral midpoint—when designing the BAWL ratings (Vö et al., 2006), we felt that these anchors might rather lead participants to interpret arousal as a unipolar and uniformly increasing dimension.

Therefore, when collecting German arousal ratings, we applied the anchors *aufregend* (“exciting”) and *beruhigend*, which might best be translated as “calming,” to emphasize an actively relaxing aspect of low arousal, as opposed to the “exciting” opposite end of the scale—understanding arousal as a bipolar concept (see the Appendix).

*Imageability.* Imageability was represented in terms of a monopolar, uniformly increasing 7-point scale according to the procedure in BAWL (Vö et al., 2006).

Ratings were collected separately for each dimension: Each participant was presented with a randomized list of words for which ratings had to be given on only one dimension, to avoid transfer effects. To perform the rating, participants were seated in a quiet room. Each list contained 603 items and was rated in a self-paced procedure. Each word was rated on each dimension by at least 20 participants. The words were presented in white letters, together with the rating scale, on a black background on a computer screen. Participants were given written instructions and five practice trials. The same general procedure was applied for all of the following data collections.



**Fig. 1** Self-Assessment Manikin (SAM): Self-evaluation scales used for the dimensions of arousal and dominance

**Participants** A total of 65 participants took part in the study (36 women, 29 men; the female ratio for any particular rating was no more than 2:1), all of whom were psychology students from the Freie Universität Berlin from the age of 18 to 37 years (mean = 24.9,  $SD = 4.6$ ). All participants (here and for all of the following data assessments) were native speakers of German and had normal or corrected-to-normal vision. Their participation was rewarded with course credit or a small amount of money.

### 3. Collecting ratings on dominance

Because dominance ratings were not part of the BAWL, we replicated the procedure used by Bradley and Lang (1999a) in order to achieve optimal cross-language comparability.

**Procedure** The SAM for dominance ranges from a small-sized (dominated) figure to a large one (in control). Due to repeatedly expressed difficulties of interpretation, more explicit anchors for the dimensions of dominance and potency were used (see the Appendix). Ratings could also be placed in the spaces between the figures, replicating the 9-point scale originally used by Bradley and Lang (1999a).

For all of the data collections reported from now on, a total of 1,003 words had to be rated. The corresponding list of stimuli was randomized and divided into two subsets containing about half of the stimuli. Each participant thus rated a total of 501 or 502 words. Each half was subsequently split, and the resulting parts were alternately assigned to one of two new experimental lists for the next two participants—a procedure that should assure that all words had similar probabilities

to co-occur in a given list for the collection of ratings with any other word from the total list.

**Participants** A group of 40 students (24 women, 16 men) from the Freie Universität Berlin participated in the study; they were 21 to 33 years of age (mean = 25.1,  $SD = 4.2$ ).

**Results and discussion** To obtain an estimate of the general comparability of our data with the original ANEW values and corresponding databases in Spanish and Portuguese, we computed bivariate correlations between the values obtained in the four different languages on the three dimensions of valence, arousal, and dominance. The substantial correlation coefficients between, for instance, German and English evaluations for valence ( $r = .90, p < .001$ ), arousal ( $r = .62, p < .001$ ), and dominance ( $r = .60, p < .001$ ) support a reasonable general quantitative comparability across the languages (see Table 1).

In particular, the very close relationship between the valence ratings in the four languages (all  $r$ s  $> .9$ ) demonstrates that emotion ratings can—in principle—almost perfectly be replicated across different languages/cultures. However, the substantially lower correlations concerning the other two dimensions suggest that these cross-cultural consistencies also

**Table 1** Bivariate correlations of valence, arousal, and dominance between languages

Dimension	E-S	E-P	E-G	S-G	P-G	P-S
Valence	.92	.92	.90	.92	.91	.94
Arousal	.75	.65	.62	.68	.75	.75
Dominance	.72	.73	.60	.77	.68	.78

E = English, S = Spanish, G = German, P = Portuguese



encounter some limitations—apparently affecting different dimensions in specific ways.

In the following discussion, we focus on these apparent discrepancies, considering especially the arousal variable and its relation to the—otherwise cross-culturally remarkably stable—concept of valence.

Although, following the classic work by Wundt (1896), the dimensions of valence and arousal have initially been introduced as two independent factors constituting a two-dimensional affective space (Bradley & Lang, 1999a; Russell, 1980), empirical reports on these variables consistently evidence a U- or boomerang-shaped distribution in which the two variables are positively correlated within the domain of positive, but negatively correlated within the domain of negative valence (Bradley, Codispoti, Cuthbert, & Lang, 2001).

For the present 1,003 German words, the distribution in the bidimensional affective space determined by the dimensions of valence and arousal (see Fig. 2) approximately fits the typical boomerang shape reported by Bradley and Lang (1999a). Accordingly, both the typical patterns of a positivity offset and a negativity bias could be replicated (Cacioppo, Gardner, & Berntson, 1997), as is revealed by a positive correlation for valence and arousal for words of positive valence (i.e., above 0, the midpoint of the valence scale;  $r = .2, p < .001$ ), and a negative correlation with a considerably steeper slope for words of negative valence ( $r = -.63, p < .001$ ), for the German sample. In general, a comparable pattern involving a positive offset and a negativity bias is, therefore, observable in all four versions of the ANEW corpus in four different languages. But note, also, that the data from the four languages display interesting discrepancies with respect to the relative strengths of the two effects. As we noted

above, the correlation between valence and arousal has a much steeper slope in the negative than in the positive valence domain in our German data, which is comparable to what has been found in the Spanish and Portuguese data, but is opposite the situation for English, where arousal increases especially strongly in positive valence ( $r = .64$  in the positive range,  $r = -.46$  in the negative range; see Table 2). Moreover, both the German and Portuguese data involve a much attenuated positive correlation ( $r < .27$ ) between the two variables concerning the range of positive words, relative to the English and Spanish data ( $r > .45$ ).

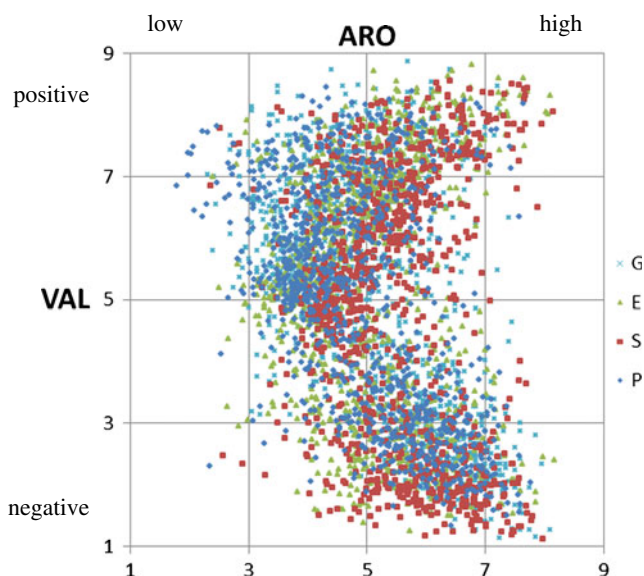
As a consequence, the overall correlations between valence and arousal continuously increased, from the English ( $r = -.05$ ), over the Spanish ( $r = -.15$ , with  $z = -2.69$ ;  $p < .01$  for the comparison to English) and Portuguese ( $r = -.49$ , with  $z = -6.2$ ;  $p < .001$  for the comparison to Spanish) data, and finally to the German data ( $r = -.56$ , with  $z = -1.03$ ;  $p = .3$  for the comparison to Portuguese). This clearly suggests that, at least for the Portuguese and German data, the two dimensions do not seem to be orthogonal to one another, challenging the assumption that the dimensions are independent.

However, the interpretation of these apparent differences concerning the internal relations of the two-dimensional space between the original ANEW and our database for German words faces a serious problem: For all emotion ratings of German words that are available so far, the operationalization of the arousal concept has been slightly modified with regard to the original instructions of Bradley and Lang (1999a)—see the Procedure section above for details.

Probably more important than the switch between a 9-point and a 5-point scale could be the more bidimensional interpretation of the arousal variable that potentially was suggested to the participants in our German sample by the translation of the verbal anchor “calm” as *beruhigend* (which might rather be understood as “calming”).

Certainly, these differences in the use of scales make an interpretation of the present results as evidence for cultural differences in the use of the arousal concept difficult.

To overcome this problem, and to test whether the apparent differences in arousal ratings across languages could have arisen from changes in the scales and instructions, we decided to re-collect ratings for our 1,003 words on the arousal dimension, this time perfectly meeting the



**Fig. 2** Correlations of valence and arousal for German ( $r = -.56$ ), English ( $r = -.05$ ), Spanish ( $r = -.15$ ), and Portuguese ( $r = -.49$ ) words

**Table 2** Bivariate correlations of valence (VAL) and arousal (ARO), by languages, for the whole, positive, and negative ranges of words

Data Set Used	E	S	G (ARO BAWL)	G (ARO ANEW)	P
Whole data set	-.05	-.15	-.56	-.58	-.49
VAL neg (<0)	-.46	-.57	-.63	-.66	-.58
VAL pos ( $\geq 0$ )	.64	.45	.20	.21	.27

E = English, S = Spanish, G = German, P = Portuguese

operationalization and instructions previously used for the ANEW (see Bradley & Lang, 1999a); hereafter, the new set of ratings will be termed ARO (ANEW), as opposed to ARO (BAWL).

Study II: Contrasting alternative but closely related dimensions

#### 4. Arousal (ANEW)

To ensure comparability across languages, ratings on the dimension of arousal were collected by perfectly matching the procedures applied in the ANEW corpus.

**Procedure** We used the original 9-point scale with five Self-Assessment Manikins, including spaces between them as optional rating positions, ranging from a relaxed (calm) to an open-eyed, literally exploding (excited) figure, with additional verbal markers being given only during the written instructions. In contrast to the version of the BAWL (Vö et al., 2009), the left verbal anchor of the scale was changed to *ruhig* (“calm”) instead of *beruhigend* (“calming”).

**Participants** A group of 40 students (22 women, 18 men) from the Freie Universität Berlin participated in the study. Their ages ranged from 18 to 34 (mean = 23.8,  $SD = 3.6$ ).

**Results** The resulting data did not differ considerably from those obtained using the BAWL scale and instructions (Vö et al., 2009; Vö et al., 2006) with a correlation coefficient of  $r = .88$  between the two. Again, for the new data, the correlation between valence and arousal was negative across the whole sample ( $r = -.58$ ), moderately positive for the positive range ( $r = .21$ ), and characterized by an especially steep slope for the negative range ( $r = -.66$ ). We thus conclude that the previously mentioned differences concerning valence–arousal correlations in the ANEW data across languages cannot be attributed to the different scale used in the BAWL, and we will again refer to these apparent cross-cultural differences in arousal ratings during the General Discussion.

#### 5. Potency

We would like to point out that the relation between past operationalizations of dominance versus potency might be more complex than has been previously considered. While both dimensions are rated using an identical pictorial SAM scale (Bradley & Lang, 1994), the important difference between them resides in the perspective that the participant has to adopt toward the rated concept. In the

case of dominance, the participant is asked to establish a relation toward the rated object and then to decide whether or not he or she can dominate the object—the central question being, “How dominant do you feel in relation to the word?” (Bradley & Lang, 1994). In the case of potency, the concepts are rated independently of their relation to the participant, who has to evaluate what potency the object might have, as such (e.g., Heise, 2010; Schröder, 2011).

**Procedure** Participants were asked in the instructions to rate the word according to its perceived potency, independently from his or her own person (cf. Heise, 2010; Schröder, 2011). The Self-Assessment Manikin scale ranged from a small-sized figure to a large one using a 9-point scale, which was formally comparable to the one used for the dominance ratings.

**Participants** A group of 40 students (23 women, 17 men) from the Freie Universität Berlin participated in the study; their ages ranged from 19 to 37 (mean = 24.5 years,  $SD = 4.3$ ).

**Results** The correlation coefficient between the dimensions of dominance and potency in our ratings was  $r = -.35$ . This relatively weak correlation clearly suggests that the scales are not simply reversed, but rather that different specific aspects seem to determine the respective ratings on each dimension. Additionally, the correlations with the dimensions of valence ( $r = .65$  for dominance,  $r = .25$  for potency,  $z = -47.96$ ;  $p < .001$ ) and arousal ( $r = -.47$  for dominance,  $r = .64$  for potency,  $z = 66.02$ ;  $p < .001$ ) differ considerably across both variables (see Table 3a), suggesting that distinct information is captured by the two measures.

#### 6. Language statistical measures

Two different measures of word frequency were joined from the print-based corpus of the Leipzig Wortschatz Projekt (Wortschatz Universität Leipzig, 2013), including over 50 million words, and the SUBTLEX corpus (Brysbaert et al., 2011), with more than 25 million words taken from movie subtitles. Finally, grammatical class, number of letters, number of syllables, and number of orthographic neighbors were generated.

#### Structure of the database

The database contains a word number (WdNum) identifying each word according to the original word number in the ANEW (Bradley & Lang, 1999a), as well as the original English word (E-word) and its German

**Table 3** Bivariate correlations between the dimensions of valence [VAL], arousal [ARO] (ANEW), arousal (BAWL), dominance [DOM], potency [POT], and imageability [IMA], as well as written wordfrequency from the Leipzig Wortschatz Projekt (log10 Leip; *Wortschatz Universität Leipzig*, 2013) or spoken word frequency from SUBTLEX (log10 SUBT; Brysbaert et al., 2011)

	VAL	ARO (ANEW)	ARO (BAWL)	DOM	POT	IMA	log10 Leip	log10 SUBT
a) For the Whole Data Set								
VAL	1***	-.58***	-.56***	.65***	-.25***	.12***	.26***	.14***
ARO (ANEW)	-.58***	1***	.88***	-.47***	.64***	-.23***	-.14***	-.05
ARO (BAWL)	-.56***	.88***	1***	-.47***	.65***	-.18***	-.15***	-.06*
DOM	.65***	-.47***	-.47***	1***	-.35***	.02	.1***	.06
POT	-.25***	.64***	.65***	-.35***	1***	-.19***	.1**	.06*
IMA	.12***	-.23***	-.18***	.02	-.19***	1***	.04	.15***
log10 Leip	.26***	-.14***	-.15***	.1**	.1**	.04	1***	.74***
log10 SUBT	.14***	-.05	-.06*	.06	.06*	.15***	.74***	1***
b) For Words of Positive Valence								
VAL	1***	.21***	.2***	.27***	.33***	-.19***	.17***	.14***
ARO (ANEW)	.21***	1***	.81***	.04	.59***	-.24***	-.01	-.02
ARO (BAWL)	.2***	.81***	1***	.01	.58***	-.17***	-.08	-.08
DOM	.27***	.04	.01	1***	-.07	-.16***	-.02	-.01
POT	.33***	.59***	.58***	-.07	1***	-.31***	.18***	.06
IMA	-.19***	-.24***	-.17***	-.16***	-.31***	1***	-.04	.11*
log10 Leip	.17***	-.01	-.08	-.02	.18***	-.04	1***	.73***
log10 SUBT	.14***	-.02	-.08	-.01	.06	.11*	.73***	1***
c) For Words of Negative Valence								
VAL	1***	-.68***	-.65***	.55***	-.48***	.09	-.02	-.1*
ARO (ANEW)	-.68***	1***	.85***	-.48***	.67***	-.02	.01	.12**
ARO (BAWL)	-.65***	.85***	1***	-.49***	.71***	0	.07	.15**
DOM	.54***	-.48***	-.49***	1***	-.46***	-.03	-.08	-.05
POT	-.48***	.67***	.71***	-.46***	1***	.05	.16***	.16***
IMA	.09	-.02	0	-.03	.05	1***	.05	.17***
log10 Leip	-.02	.01	.07	-.08	.16***	.05	1***	.75***
log10 SUBT	-.1*	.12**	.15**	-.05	.16***	.17***	.75***	1***

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ 

translation (G-word). Next, the means (Mean) and standard deviations (SD) for each evaluative dimension of valence (VAL), arousal BAWL [ARO (BAWL)], arousal ANEW [ARO (ANEW)], dominance (DOM), and potency (POT) are provided, followed by the imageability ratings for the German words (IMA). In addition, the variables of word frequency—from both written language, taken from the Leipzig Wortschatz Projekt (*Wortschatz Universität Leipzig*, 2013; freq\_Leipzig), and movie subtitles reflecting the use of spoken language, taken from SUBTLEX (Brysbaert et al., 2011; freq\_SUBTLEX)—number of letters (#\_letters), number of syllables (#\_syllables), word class (word\_class), and number of orthographic neighbors (Colthearts  $N$ :  $N_{orth}$ ) are included. All of the latter variables are based on CELEX (Baayen, Piepenbrock, & Gulikers, 1995). The database can be downloaded as supplemental materials accompanying this article.

## General discussion

The ANEW (Bradley & Lang, 1999a) represents the best-known affective dictionary for the English language, providing US American ratings on the three dimensions of valence, arousal, and dominance. To facilitate cross-cultural emotion research with verbal materials, we have presented an adaptation for the German language, offering competing operationalizations of the affective measurements, along with additional psycholinguistic variables. For the purpose of a better understanding of the alternative scales measuring similar aspects of affective experience, the different operationalizations are here examined in more detail.

Concerning the two variables of valence and arousal, defining the two-dimensional affective space proposed by some theoretical accounts (Russell, 1980), a closer look at the previously available large-scale databases already makes the

assumption of two independent components seem less plausible: The normative databases have classically been characterized by a U- or boomerang-shaped distribution of the two dimensions, and are thus only weakly correlated across the whole range of the valence scale, but closely correlated within the more constrained ranges of either positive or negative valence (see Table 2).

In the case of our German data, even across the whole range of the valence scale, a relatively high negative correlation between the two is given, with arousal decreasing with increasing valence—mostly due to the very tight correlation of the two within the range of negative valence, a pattern consistent across operationalizations of the arousal scale according to both ANEW (Bradley & Lang, 1999a) and BAWL (Võ et al., 2009).

Apparently, though this is rarely emphasized, the issue of high linear correlations between valence and arousal seems to be a rather widely spread phenomenon. A comparable pattern also appears in the Portuguese sample (Soares et al., 2012), where we calculated an overall correlation coefficient for both dimensions of  $r = -.49$ . A similar misalignment has also been reported in a cross-cultural comparison of the IAPS pictures (Lang, Bradley, & Cuthbert, 2008) to a Hungarian sample, where the dimensions of valence and arousal correlated to  $r = -.54$  (Deák, Csenski, & Révész, 2010). Ribeiro, Pompéia, and Bueno (2005) found a correlation of  $r = -.82$  for valence and arousal on the basis of evaluation of the IAPS pictures for a Brazilian sample.

As Ribeiro et al. (2005) argued, this might be explained by “the absence of a clearly defined concept of ‘arousal’ ” (p. 214). Whereas the left extreme of the SAM, in the original study by Bradley and Lang (1999a) anchored as *calm, relaxed*, might be understood as the “absence of alteration of the participant’s normal state”, this judgment mostly holds true for stimuli that are neutral in valence, as has clearly been the case for the American and Spanish sample.

Our replications of arousal ratings with the original ANEW instructions and anchors shows that these difficulties to replicate the data from Bradley and Lang (1999a) concerning the symmetric U-shape of the two-dimensional affective space clearly do not result from methodological issues such as the rating instructions emphasizing the bipolar character of the scale or differences between a 5- and a 9-point scale.

On the other hand, it needs to be acknowledged that the positive emotional connotations of the English words “exciting” or “excited” might in part explain why a more pronounced positive correlation between valence and arousal for words of a general positive valence is given for the US American data, as compared to the data from other languages in which the translations of these verbal anchors have less positive connotations. However, these potential emotional connotation differences across

languages at the level of one single word may also point to a more general cultural bias determining the specific way that arousal ratings evolve in affective databases across languages: Whereas correlations of arousal values from an American sample are quite comparable to those of Spanish subjects, the values from Portuguese subjects seem to correlate higher with those from a German sample, and both the Portuguese and German values differ from those in the American and Spanish samples.

We suggest that these specific differences between the four languages involved match with differences characterizing them in terms of cultural norms concerning the personality construct of extraversion. As has been shown in previous studies (Ramírez-Esparza, Gosling, Benet-Martínez, Potter, & Pennebaker, 2006; Veltkamp, Recio, Jacobs, & Conrad, 2013), the construct of extraversion in bilingual individuals, among other things, is subject to change depending on cultural framing. Bilinguals, for instance, reached higher scores of extraversion when filling out a personality inventory in Spanish than when doing so in German (Veltkamp et al., 2013). No such results are available so far for Portuguese, but we believe that labeling the Portuguese culture as being less extraverted than the Spanish can be considered a “common ground” and the same would hold true when comparing German to American culture. As for the present data, our tentative proposal is that cultures favoring extraversion over introversion—as we believe is the case if we compare English and Spanish, on one side of the extra-/introversion dimension, with German and Portuguese, on the other—tend to assign a higher arousal to positive events, thus making them a focus of overt verbal communication. On the other side, more introverted cultures tend to favor silently enjoying positive events, which, in turn, are perceived as being less arousing.

Relative to valence and arousal, the role of dominance and/or potency in language processing has long been ignored, possibly due to its indeterminate nature. From a formal point of view, the two dimensions simply differ concerning the perspective of the raters: whether they feel more or less “dominant” toward something, or whether they perceive something as being more or less “potent”. If a mere inversion of perspective characterized the relation between the two concepts of potency and dominance, a strong negative correlation between their respective values would be expected. The rather weak negative correlation between the two measures revealed by our data, thus, clearly contradicts the assumption of such a simple inversion. Bradley and Lang (1994) already suggested that the dominance scale, with its originally envisaged clinical use, highlights more the raters’ self-centered perspective. Therefore, raters of dominance might focus more on their own coping strategies towards an object, as compared to directly rating its potency. As our data show, this holds



implications for the relations of both variables to the valence and arousal dimensions. On the one hand, the particularly high positive correlation of dominance and valence suggests that more positive affect arises as a function of perceiving a higher internal locus of control with regard to a given stimulus. On the other hand, ratings of potency instead display particularly high positive correlations with arousal. Both phenomena are significantly more pronounced in the case of negative than of positive words, both for dominance correlating with valence ( $r = .27$  for positive words,  $r = .55$  for negative words,  $z = 26.83$ ;  $p < .001$ ) and for potency correlating with arousal (ANEW) ( $r = .59$  for positive words,  $r = .67$  for negative words,  $z = 18.61$ ;  $p < .001$ ) (see Table 3b and c). This pattern shows that both scales cannot be considered, and should not be used as if they were, interchangeable. The specific contrast observed in our data seems to align well with influential appraisal models of emotion (Scherer, 1999; Scherer, Schorr, & Johnstone, 2001), according to which stimuli or events of emotional relevance initially evoke unconscious reactions of the autonomous nervous system—for instance, in the noradrenergic system. The potential of stimuli to trigger such responses might, thus, strongly determine ratings of potency, which accordingly appear to be strongly related to arousal ratings—the dimension of affective space that more or less directly refers to changes in the autonomous nervous system. According to Scherer (1999), these initial reactions to emotional input should then be followed by more conscious evaluation processes involving the evaluation of available coping strategies—presumably assessed by dominance ratings—and leading to a final evaluation of concepts, stimuli, or situations as being pleasant or unpleasant, which would, then, explain the tight positive correlation between valence and dominance ratings.

In sum, clearly more research will have to be done to fully understand the interplay of the various dimensions constituting the affective space within and across different languages (see Conrad et al., 2011, who compared emotional event-related potential effects in response to German and Spanish words). The German adaptation of ANEW will be a further useful and valid measure to support future research on emotion in general and cross-linguistic cognitive and psychophysiological research; in particular, the new database extends the range of the classical ANEW dimensions with the newly added dimension of imageability, a variable that has been shown to influence word processing in a number of experimental studies (Altarriba & Bauer, 2004; Altarriba et al., 1999; Huang, Lee, & Federmeier, 2010; Kanske & Kotz, 2007; Kousta et al., 2011; Vigliocco et al., 2011) and that can be largely or entirely accounted for by two computable measures: the size and density of a word's context, and the emotional associations of the word (Westbury et al., 2013).

## Appendix: Verbal anchors

**Valence (7-point scale)** The verbal markers presented below the scale were *neutral* (“neutral”) for the middle position, *etwas positiv/negativ* (“slightly positive/negative”), *positiv/negativ* (“positive/negative”), and *sehr positiv/negativ* (“very positive/negative”).

**Arousal BAWL (5-point scale)** The verbal markers presented below the scale were *sehr beruhigend* (“very calming”), *beruhigend* (“calming”), *weder noch* (“neither nor”), *aufregend* (“exciting”), and *sehr aufregend* (“very exciting”).

**Arousal ANEW (9-point scale)** The verbal anchors used for the extreme ends of the scale were *ruhig* (“very calm”) and *aufregend* (“exciting”).

**Dominance (9-point scale)** The verbal anchors used for the extreme ends of the scale were *einflussreich/beeinflusst; kontrollierend/kontrolliert; dominant/gefällig; eigenständig/geführt* (“influential/influenced; controlled/controlling; dominant/submissive; autonomous/guided”). The verbal markers below were *neutral* (“neutral”), *etwas* (“slightly”), *ziemlich* (“rather”), *sehr* (“very”), and *äußerst* (“extremely”), with *neutral* being in the center of the scale.

**Potency (9-point scale)** The verbal anchors used for the extreme ends of the scale were *beeinflusst/einflussreich; kontrolliert/kontrollierend; gefällig/dominant; geführt/eigenständig* (“influenced/influential; controlled/controlling; submissive/dominant; guided/autonomous”). The verbal markers below were *neutral* (“neutral”), *etwas* (“slightly”), *ziemlich* (“rather”), *sehr* (“very”), and *äußerst* (“extremely”), with *neutral* being in the center of the scale.

**Imageability (7-point scale)** The verbal anchors used for the extreme ends of the scale were *gar nicht bildhaft* (“not imageable at all”) to *absolut klar bildhaft vorstellbar* (“perfectly imageable”).

## References

- Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology*, 117, 389–410.
- 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.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (Release 2, CD-ROM)*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001).

- Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1, (3), 276–298. doi:10.1037/1528-3542.1.3.276
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49–59.
- Bradley, M. M., & Lang, P. J. (1999a). *Affective Norms for English Words (ANEW): Stimuli, instruction manual and affective ratings (Technical Report No. C-1)*. Gainesville: University of Florida, Center for Research in Psychophysiology.
- Bradley, M. M., & Lang, P. J. (1999b). *International Affective Digitized Sounds (IADS): Stimuli, instruction manual and affective ratings*. Gainesville: University of Florida, Center for Research in Psychophysiology.
- Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011a). Discrete emotion norms for nouns: Berlin Affective Word List (DENN–BAWL). *Behavior Research Methods*, 43, 441–448. doi:10.3758/s13428-011-0059-y
- Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011b). Discrete emotion effects on lexical decision response times. *PLoS ONE*, 6, e23743. doi:10.1371/journal.pone.0023743
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, 58, 412–424.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1, 3–25.
- Citron, F. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, 122, 211–226.
- Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: Across cultural ERP study with German–Spanish bilinguals. *Frontiers in Psychology*, 2(351), 1–16. doi:10.3389/fpsyg.2011.00351
- Deák, A., Csenski, L., & Révész, G. (2010). Hungarian ratings for the International Affective Picture System (IAPS): A cross-cultural comparison. *Empirical Text and Culture Research*, 4, 90–101.
- Eilola, T. M., & Havelka, J. (2010). Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods*, 42, 134–140. doi:10.3758/BRM.42.1.134
- Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103, 518–565. doi:10.1037/0033-295X.103.3.518
- Grimm, S., Weigand, A., Kazzner, P., Jacobs, A. M., & Bajbouj, M. (2012). Neural mechanisms underlying the integration of emotion and working memory. *NeuroImage*, 61, 1188–1194.
- Hamann, S., & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *NeuroReport*, 13, 15–19.
- Heise, D. R. (2010). *Surveying cultures. Discovering shared conceptions and sentiments*. Hoboken: Wiley.
- Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., & Kissler, J. (2009). Amygdala activation during reading of emotional adjectives—An advantage for pleasant content. *Social Cognitive and Affective Neuroscience*, 4, 35–49.
- Hofmann, M. J., Kuchinke, L., Biemann, C., Tamm, S., & Jacobs, A. M. (2011). Remembering words in context as predicted by an associative read-out model. *Frontiers in Psychology*, 2, 252. doi:10.3389/fpsyg.2011.00252
- Hofmann, M. J., Kuchinke, L., Tamm, S., Vö, M. L.-H., & Jacobs, A. M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, & Behavioral Neuroscience*, 9, 389–397. doi:10.3758/9.4.389
- Huang, H. W., Lee, C. L., & Federmeier, K. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage*, 49, 1116–1123.
- Huckauf, A., Heller, D., & Gouzouli-Mayfrank, E. (2003). Ein Instrument zur Messung der Wirkung emotionaler Valenzen auf die Wortverarbeitung. *Zeitschrift für Psychologie*, 211, 129–137.
- Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, 1148, 138–148.
- Kanske, P., & Kotz, S. A. (2010). Leipzig affective norms for German: A reliability study. *Behavior Research Methods*, 42, 987–991. doi:10.3857/BRM.42.3.987
- Kayser, J., Fong, R., Tenke, C. E., & Bruder, G. (2003). Event-related brain potentials during auditory and visual word recognition memory tasks. *Cognitive Brain Research*, 16, 11–25.
- Kissler, J., Herbert, C., Peyk, P., & Junghofer, M. (2007). Buzzwords: Early cortical responses to emotional words during reading. *Psychological Science*, 18, 475–480.
- 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. doi:10.1037/a0021446
- Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, 112, 473–481. doi:10.1016/j.cognition.2009.06.007
- Kuchinke, L., Jacobs, A. M., Grubich, C., Vö, M. L.-H., Conrad, M., & Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: An fMRI study. *NeuroImage*, 28, 1022–1032.
- Kuchinke, L., Vö, M. L.-H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, 65, 132–140. doi:10.1016/j.ijpsycho.2007.04.004
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. H. J. B. Sidowski (Ed.), *Technology in mental health and delivery systems* (pp. 119–137). Norwood: Ablex.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual (Technical Report A-8)*. Gainesville: University of Florida, Center for Research in Psychophysiology.
- Mehrabian, A., & Russell, J. (1974). *An approach to environmental psychology*. Cambridge: MIT Press.
- Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273–315. doi:10.1037/0033-295X.114.2.273
- Ponz, A., Montant, M., Liegeois-Chauvel, C., Silva, C., Braun, M., Jacobs, A. M., & Ziegler, J. C. (2013). Emotion processing in words: A test of the neural re-use hypothesis using surface and intracranial EEG. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nst034. Advance online publication.
- Ramirez-Esparza, N., Gosling, S. D., Benet-Martinez, V., Potter, J. P., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of research in personality*, 40(2), 99–120.
- Redondo, J., Fraga, I., Comesaña, M., & Perea, M. (2005). Estudio normativo del valor afectivo de 478 palabras españolas. *Psicológica*, 26, 317–326.
- Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, 39, 600–605. doi:10.3758/BF03193031
- Ribeiro, R. L., Pompéia, S., & Bueno, O. F. (2005). Comparison of Brazilian and American norms for the International Affective

- Picture System (IAPS). *Revista Brasileira de Psiquiatria*, 27, 208–215.
- Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.
- Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience*, 9, 28–43. doi:[10.3758/CABN.9.1.28](https://doi.org/10.3758/CABN.9.1.28)
- Scherer, K. R. (1999). Appraisal theory. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). Chichester: Wiley.
- Scherer, K. R., Schorr, A. E., & Johnstone, T. E. (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford: Oxford University Press.
- Schröder, T. (2011). A model of language-based impression formation and attribution among Germans. *Journal of Language and Social Psychology*, 30, 82–102.
- Silva, C., Montant, M., Ponz, A., & Ziegler, J. C. (2012). Emotions in reading: Disgust, empathy and the contextual learning hypothesis. *Cognition*, 125, 333–338.
- Soares, A., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods*, 44, 256–269. doi:[10.3758/s13428-011-0131-7](https://doi.org/10.3758/s13428-011-0131-7)
- Stevenson, R. A., Mikels, J. A., & James, T. W. (2007). Characterization of affective norms for english words by discrete emotional categories. *Behavior Research Methods*, 39, 1020–1024. doi:[10.3758/BF03192999](https://doi.org/10.3758/BF03192999)
- Tabert, M. H., Borod, J. C., Tang, C. Y., Lange, G., Wei, T. C., Johnson, R., & Buchsbaum, M. S. (2001). Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: An fMRI study. *Neuropsychologia*, 39, 556–573.
- Trauer, S. M., Andersen, S. K., Kotz, S. A., & Müller, M. M. (2012). Capture of lexical but not visual resources by task-irrelevant emotional words: A combined ERP and steady-state visual evoked potential study. *NeuroImage*, 60, 130–138.
- Veltkamp, G. M., Recio, G., Jacobs, A. M., & Conrad, M. (2013). Is personality modulated by language? *International Journal of Bilingualism*, 17, 496–504. doi:[10.1177/1367006912438894](https://doi.org/10.1177/1367006912438894)
- Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews*, 35, 407–426. doi:[10.1016/j.neubiorev.2010.04.007](https://doi.org/10.1016/j.neubiorev.2010.04.007)
- Vö, M. L.-H., Conrad, M., Kuchinke, L., Hartfeld, K., Hofmann, M. F., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, 41, 534–538. doi:[10.3758/BRM.41.2.534](https://doi.org/10.3758/BRM.41.2.534)
- Vö, M. L.-H., Jacobs, A. M., & Conrad, M. (2006). Crossvalidating the Berlin affective word list (BAWL). *Behavior Research Methods*, 38, 606–609. doi:[10.3758/BF03193892](https://doi.org/10.3758/BF03193892)
- Vö, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, 45, 130–140. doi:[10.1111/j.1469-8986.2007.00606.x](https://doi.org/10.1111/j.1469-8986.2007.00606.x)
- Weigand, A., Grimm, S., Astalosch, A., Guo, J. S., Briesemeister, B. B., Lisanby, S. H., & Bajbouj, M. (2013). Lateralized effects of prefrontal repetitive transcranial magnetic stimulation on emotional working memory. *Experimental Brain Research*, 227, 43–52.
- Westbury, C. F., Shaoul, C., Hollis, G., Smithson, L., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2013). Now you see it, now you don't: On emotion, context, & the algorithmic prediction of human imageability judgments. *Frontiers in Psychology*.
- Whissell, C., Fournier, M., Pelland, R., Weir, D., & Makarec, K. (1986). A dictionary of affect in language: IV. Reliability, validity, and applications. *Perceptual and Motor Skills*, 62, 875–888.
- Wortschatz Universität Leipzig. (2013). Retrieved May 7, 2013, from <http://wortschatz.uni-leipzig.de/>
- Wundt, W. (1896). *Grundriss der Psychologie*. Leipzig: Engelmann.