

Characteristics for 200 words rated by young and older adults: Age-dependent evaluations of German adjectives (AGE)

DANIEL GRÜHN

University of Geneva, Geneva, Switzerland

AND

JACQUI SMITH

University of Michigan, Ann Arbor, Michigan

We describe the Age-Dependent Evaluations of German Adjectives (AGE). This database contains ratings for 200 German adjectives by young and older adults (general word-rating study) and graduate students (self-other relevance study). Words were rated on emotion-relevant (valence, arousal, and control) and memory-relevant (imagery) characteristics. In addition, adjectives were evaluated for self-relevance (Does this attribute describe you?), age relevance (Is this attribute typical for young or for older adults?), and self-other relevance (Is this attribute more relevant for the possessor or for other persons?). These ratings are included in the AGE database as a resource tool for experiments on word material. Our comparisons of young and older adults' evaluations revealed similarities but also significant mean-level differences for a large number of adjectives, especially on the valence dimension. This highlights the importance of age in the perception of emotional words. Data for all the words are archived at www.psychonomic.org/archive/.

In the social cognitive aging and cognitive neuroscience literature, the role of age in modulating the processing of emotional information has become a focus of research. In particular, the impact of emotional material on age-related differences in attention (e.g., Mather & Carstensen, 2003), memory (e.g., Charles, Mather, & Carstensen, 2003; Grünh, Scheibe, & Baltes, 2007; Grünh, Smith, & Baltes, 2005), and decision making (Löckenhoff & Carstensen, 2004) has been the focus of lively research and passionate debate. Despite this growing interest, we know rather little about age-related differences in the perception and meaning of emotional material (Grünh & Scheibe, 2008). Does our perception of certain emotional traits or states change as we age? Is *punctual* perceived as more positive in old age than in young adulthood? Is *amused*, in contrast, perceived as less positive?

The question of age differences in the meaning of emotional material is important in two respects. On the one hand, research on the meaning of emotional information is an opportunity to examine age differences in the salience of information. Some stimuli and their associated meanings may be more negative or positive for certain age groups than for others. These age-related differences may be a function of life experience, lifetime exposure, and age-related changes in psychological, biological, and social functioning. On the other hand, age differences in one domain (e.g., memory) may be modulated by age dif-

ferences in the connotations of the emotional material used. If young and older adults differ in the meanings they associate with material, age differences in processing this material are likely. In fact, there is some empirical evidence that age differences in emotional-cognitive tasks are moderated by age differences in perceiving the material used. For example, Kensinger, Brierley, Medford, Growdon, and Corkin (2002) investigated age differences in recalling positive and negative words from the Affective Norms for English Words (ANEW; Lang, Bradley, & Cuthbert, 1998). When subjective valence categorization (based on each person's own evaluations) was used, an interaction between valence and age was marginally significant: Older adults recalled more negative than positive words, whereas young adults remembered more positive than negative words. In contrast, when a priori categorization (from young adults' ratings in the ANEW) was used, no significant interaction (and nearly the opposite pattern) was found. This differential pattern implies that a considerable amount of variance in remembering emotional words can be explained by age-related differences in the emotional evaluation of the material. Consequently, there is a need for research on the perception and meaning of emotional information in different age groups.

Empirical evidence for age differences in evaluations of emotional material is scarce and available mainly for pictorial material. For example, Grünh and Scheibe (2008)

D. Grünh, dgruehn@ncsu.edu

investigated young and older adults' evaluations of 504 emotional pictures. A large number of pictures showed age-related differences in valence (30%) and arousal ratings (18.7%). In general, older adults tended to rate emotional pictures more extremely than young adults did: Older adults rated positive pictures as more positive and negative pictures as more negative than young adults did. Similarly, older adults rated negative pictures as more arousing and positive pictures as less arousing than young adults did (Grühn & Scheibe, 2008). Regarding word material, we know relatively little about age differences in emotional evaluations. We are aware of only two studies that, in part, investigated age differences in emotional evaluations of word material. Mueller, Wonderlich, and Dugan (1986), for example, investigated young and older adults' judgments about age-specific attributes. Specifically, participants were asked whether a trait "describes you" or "describes most people." The attributes consisted of 40 young-specific and 40 elderly-specific adjectives; half of them were negative attributes, and half were positive attributes. Findings showed that older adults were more likely to endorse positive attributes and to reject negative attributes than young adults were. This pattern was evident irrespective of the proposed age specificity of the adjectives. Heckhausen, Dixon, and Baltes (1989) investigated the average age associated with changes in personal attributes. Young, middle-aged, and older adults were asked to rate the desirability of developmental changes (e.g., *becoming more honest* or *less arrogant*) for 148 adjectives. Age groups showed high similarity in ratings of the desirability of developmental changes.

These two word-rating studies (Heckhausen et al., 1989; Mueller et al., 1986) emphasized, on the one hand, the distinction between self-relevant and other-relevant attributes and, on the other hand, the age relevance of these attributes (Is this typical for a specific age group?). There is a growing literature on the distinction between self and others, or between self-relevant and other-relevant information (e.g., Andersen & Chen, 2002; Batson et al., 1997; Labouvie-Vief, 2005). Attributes of others are often highly relevant. For example, the perception that someone is *sympathetic* or *aggressive* can make a difference for our own attitudes and behaviors toward this person. Following a distinction by Wentura, Rothermund, and Bak (2000), characteristics can be more relevant for the possessor (i.e., self) or for other persons interacting with someone who shows this characteristic. This distinction can be easily made by asking two questions: Is it good or bad for me when I am X? Or is it good or bad for me when I am interacting with someone who is X? Possessor-relevant characteristics reveal a clear response to the first question, but not to the second. In contrast, other-relevant characteristics show a clear answer to the second question, but not to the first. For example, it is clearly good for me to be intelligent, but whether it is good or bad for me to interact with someone who is intelligent depends on the context. In contrast, it is often bad for me to interact with someone who is aggressive; but whether it is good or bad for me to be aggressive depends on the context as well. Similar to the distinction between self and others is the

distinction of whether an attribute is typical for one's own age group or for other age groups. Thus, attributes might differ in their salience for different age groups. Some attributes might be more typical for older adults (e.g., *wise*), whereas other attributes are more typical for young adults (e.g., *spontaneous*). These differences in the age relevance of the material might influence the evaluations of people of different ages. For example, older adults might evaluate typical characteristics of older adults as more positive than young adults do.

To address the research gap on age differences in evaluations of emotional attributes, we conducted two word-rating studies on 200 German adjectives. The goals of the present studies were (1) to provide detailed evaluations of word material relevant for experimental settings, (2) to investigate age-related differences in the evaluations of emotional attributes, and (3) to make the age-dependent evaluations available to a greater public.

In Study 1, the general word-rating study, young and older adults were asked to rate 200 German adjectives on six dimensions: valence, arousal, control, imagery, age relevance, and self-relevance. The rating dimensions were selected to cover three aspects: The dimensions point to emotional meaning (i.e., valence, arousal, and control), to memory-relevant characteristics (i.e., imagery), or to the self-relevance and age relevance that may influence the processing of words differently for young and older adults (i.e., self-relevance and age relevance).

In Study 2, the self-other relevance study, we extended this procedure by asking psychology graduate students to rate self-other relevance. Self-other relevance involves judging whether an attribute is more relevant for a possessor of this attribute or more relevant for a person interacting with a possessor of this attribute (Wentura et al., 2000). Because the differentiation between a more possessor-relevant and a more other-relevant characteristic is conceptually difficult, we asked graduate students majoring in psychology to complete these ratings.

The ratings obtained in these two studies are included in the Age-Dependent Evaluations of German Adjectives (AGE) database. These evaluations are available in an archived file as a resource for researchers interested in selecting word material for future studies. To facilitate the description of the word material, we will present the findings from both studies together.

METHOD

Word Material

In a first step, a pool of German adjectives was established with corresponding rating data obtained in previous studies. These data were compiled predominantly from a book by Hager and Hasselhorn (1994), which reviewed the material and findings of several German rating studies (Angleitner, Ostendorf, & John, 1990; Busz et al., 1972; Hager, Mecklenbräuker, Möller, & Westermann, 1985; Heydecke, 1984; Klapprott, 1972; Mecklenbräuker, Hager, & Möller, 1994; Möller & Hager, 1991; Ostendorf, 1994; Schwibbe, Räder, Schwibbe, Borchardt, & Geiken-Pophanken, 1981; Wippich & Bredenkamp, 1977). Other sources were checked to determine whether relevant adjectives were missing. These other sources were (1) emotion adjectives of the PANAS-X (Watson & Clark, 1994), (2) emotion adjectives of the MDBF scales (Steyer, Schwenkmezger,

Notz, & Eid, 1997), (3) marker adjectives for the five-factor model (Goldberg, 1992), and (4) adjectives used in a study by Heckhausen et al. (1989). This procedure resulted in a total of 5,432 adjectives.

The word pool was systematically reduced on the basis of seven selection criteria. We removed words that (1) had a complex word structure (e.g., *Mutter-Seelen-allein*, which translates as *all alone like a motherless child*), (2) were uncommon (e.g., *virile*), (3) were highly frequent or rare in written text ($7 \leq$ word frequency class¹ ≤ 17), (4) had fewer than 4 and more than 12 letters, (5) lacked clarity of meaning, (6) were low in imagery or concreteness, and (7) could not be used to describe a person (e.g., *endless*). Words were not excluded if they were marker words for a rating dimension. Marker words were words that could be used to describe the endpoints of a rating dimension. For example, as a treatment check for whether participants understood the instructions for the age-relevance dimension, the adjectives *young* (*jung*) and *old* (*alt*) were kept in the word list. These two adjectives are prototypical examples of the endpoints of age relevance—that is, for a very low or very high score on a bipolar scale. The final set of words contained 200 German adjectives. The archived file contains English translations.

Rating Dimensions

In Study 1, the general word-rating study, words were rated on six dimensions: valence, arousal, control, imagery, age relevance, and self-relevance. Instructions for each dimension were adapted from instructions given by Paivio, Yuille, and Madigan (1968).² With one exception (i.e., age relevance), all the dimensions were rated on 7-point scales. The participants were asked to indicate the following: (1) for valence, the feeling of pleasantness elicited by each word, from *very unpleasant* (1) to *very pleasant* (7); (2) for arousal, the feeling of tension elicited by each word, from *very relaxed* (1) to *very tensed* (7); (3) for control, the feeling of control elicited by each word, from *no control* (1) to *high control* (7); (4) for imagery, how easily each word elicited a visual image, from *very difficult* (1) to *very easily* (7); and (5) for self-relevance, how accurately each word describes oneself, from *not at all accurate* (1) to *very accurate* (7). Age relevance, in contrast, was rated on a 5-point scale. The participants were asked to indicate whether a word was *very typical for young adults* (1), *more typical for young adults* (2), *more typical neither for young nor for older adults* (3), *more typical for older adults* (4), or *very typical for older adults* (5).

In Study 2, the self–other relevance study, words were rated on three scales: (1) self-valence, (2) other-valence, and (3) self–other relevance. Self-valence (i.e., possessor relevance) was assessed with the question, “How good or bad is it for me, when I am X?” Other-valence (i.e., interacting with possessor) was measured with “How good or bad is it for me, when I interact with someone who is X?” Both scales ranged from *very bad/unpleasant* (1) to *very good/pleasant* (7). Finally, the participants were asked to decide whether a word was other-relevant (1) or self-relevant (2).

Participants

General word-rating study. The sample for the general word-rating study comprised 24 young adults (20–30 years of age; $M = 24.3$, $SD = 2.7$; 50% female) and 24 older adults (65–76 years of age; $M = 70.8$, $SD = 3.3$; 50% female) recruited in the local area of Berlin. For the 2-h session, the participants received €20 as compensation. To facilitate comparison with other studies, we assessed indicators of subjective well-being and intellectual functioning that are frequently reported in age-comparative research. Subjective well-being was assessed with two single items tapping life satisfaction (“How satisfied are you with your present life?”) and subjective health (“How good is your physical health at present?”). Responses were made on 5-point scales ranging from *very unsatisfied* (1) to *very satisfied* (5) and from *very poor* (1) to *excellent* (5) for life satisfaction and subjective health, respectively. Young and older adults reported similar levels of life satisfaction [$F(1,46) = 0.33$, $p = .57$, $\eta^2 < .01$ ($M_{\text{young}} = 4.04$, $SD_{\text{young}} = 0.75$, $M_{\text{old}} = 4.17$, $SD_{\text{old}} = 0.76$)] and subjective health [$F(1,46) = 2.42$, $p = .13$, $\eta^2 = .05$ ($M_{\text{young}} = 4.0$; $SD_{\text{young}} = 0.66$,

$M_{\text{old}} = 3.67$, $SD_{\text{old}} = 0.82$]. Young adults scored higher than older adults in perceptual speed (Wechsler, 1981) [$F(1,46) = 56.08$, $p < .01$, $\eta^2 = .55$ ($M_{\text{young}} = 65.6$, $SD_{\text{young}} = 12.0$, $M_{\text{old}} = 43.3$, $SD_{\text{old}} = 8.4$)], and on a vocabulary test (Wechsler, 1981) [$F(1,46) = 5.6$, $p < .05$, $\eta^2 = .11$ ($M_{\text{young}} = 24.7$, $SD_{\text{young}} = 2.9$, $M_{\text{old}} = 21.7$, $SD_{\text{old}} = 5.4$)]. On average, young adults had more years of education ($M = 16.1$ years, $SD = 2.2$) than did older adults ($M = 13.6$ years, $SD = 3.5$) [$F(1,46) = 8.71$, $p < .01$, $\eta^2 = .16$].

Self–other relevance study. Nineteen young psychology graduate students (22–30 years of age; $M = 25.2$, $SD = 2.3$; 78.9% female) from the Free University Berlin and Humboldt University of Berlin were recruited for the self–other relevance study. The students’ participation was voluntary and unrewarded.

Procedure

General word-rating study. The participants were tested at the Max Planck Institute for Human Development, Berlin, in small age-homogeneous groups of 2–5 persons. After completing some demographic questions and cognitive tasks (i.e., vocabulary and perceptual speed), the participants were introduced to the rating procedure and were asked to complete three booklets: The first booklet contained material for the dimensions of valence and arousal; the second booklet for control and imagery; and the third booklet contained material for age relevance and self-relevance. Across participants, the order of dimensions was counterbalanced within each booklet. Each dimension was treated separately in one section of the booklets. Each section contained an instructions page for this dimension, followed by eight pages of 25 words for the ratings. The order of words varied within the different rating dimensions.

Self–other relevance study. Graduate psychology students were instructed in the self–other evaluation procedure at the Max Planck Institute for Human Development, Berlin. The order of the three rating dimensions in this study was constant across participants: (1) self-valence, (2) other-valence, and (3) self–other relevance. As for the design of Study 1, each rating dimension was treated separately in one section of the booklet. Each section contained an instructions page for this dimension, followed by eight pages, each listing 25 words. Word order was randomized across participants and between dimensions.

RESULTS

We will first report comprehensive information about word evaluations, their intercorrelations, and the correlations between ratings obtained in the present study and in previous research. The latter analyses were done in order to provide information about the generalizability of our data. Second, we will report age-related differences in these word characteristics. All the ratings are available in the archived file.

General Word Characteristics of the 200 German Adjectives

Marker adjectives. In order to check whether the participants understood the rating procedure employed, we examined words at the bipolar ends of each dimension. Table 1 provides the three adjectives with the highest scores and the three adjectives with the lowest scores on each dimension for the total sample. For all the dimensions, the participants responded in expected ranges and directions. Words generally associated with a negative tone (e.g., *brutal*, *mendacious*) were rated as unpleasant, whereas positive words (e.g., *happy*, *healthy*) were rated as pleasant. Words that connote a high degree of tension (e.g., *aggressive*, *belligerent*) were rated high on arousal,

Table 1
Adjectives With Highest and Lowest Scores on Each Dimension

Dimension	Words With Highest Scores	Words With Lowest Scores
Valence	happy (6.79), honest (6.56), delighted (6.53)	mendacious (1.33), cruel (1.10), brutal (1.10)
Arousal	brutal (6.67), aggressive (6.66), cruel (6.56)	content (1.54), relaxed (1.44), comfortable (1.44)
Control	determined (6.38), concentrating (6.21), active (6.17)	spoiled (1.92), depressive (1.85), helpless (1.55)
Imagery	old (6.48), ugly (6.33), attractive (6.32)	neutral (2.54), liberal (2.50), subjective (1.60)
Self-relevance	tolerant (6.23), honest (6.19), interested (6.10)	cruel (1.38), brutal (1.33), dumb (1.31)
Age relevance	old (4.79), experienced (4.40), wise (4.38)	lively (1.72), spontaneous (1.65), young (1.25)
Self-valence	happy (6.94), healthy (6.88), balanced (6.81)	dumb (1.19), depressive (1.19), brutal (1.13)
Other-valence	happy (6.94), content (6.69), interested (6.69)	malicious (1.06), cruel (1.06), brutal (1.00)

Note—Scores ranged (1) for valence, from *very unpleasant* (1) to *very pleasant* (7); (2) for arousal, from *very relaxed* (1) to *very tensed* (7); (3) for control, from *no control* (1) to *high control* (7); (4) for imagery, from *very difficult* to elicit a visual image (1) to *very easy* (7); (5) for self-relevance, from *not at all accurate* (1) to *very accurate* (7); (6) for age relevance, from *very typical for young adults* (1) to *very typical for older adults* (5); and (7) for self-valence (“How good or bad is it for me, when I am X?”) and other valence (“How good or bad is it for me, when I am interacting with someone who is X?”), from *very bad/unpleasant* (1) to *very good/pleasant* (7).

whereas the opposite was true for words involving a feeling of relaxation (e.g., *relaxed*, *content*). A feeling of control was associated with active words (e.g., *determined*, *active*), but not with words that connote some degree of helplessness (e.g., *helpless*, *bewildered*). Words considered easy to imagine were concrete (e.g., *old*, *ugly*), whereas words rated hard to imagine were abstract (e.g., *subjective*, *neutral*). For the self-rating, nearly all the participants indicated that they were *tolerant* but not *dumb*.

For age relevance, both marker adjectives, *young* and *old*, were rated as one would expect—namely, as the most typical characteristics of young and older adults, respectively. It is noteworthy that words rated as being typical descriptions of older adults were generally not negative. For example, *experienced* and *wise*, both rated as very pleasant ($M_{\text{experienced}} = 6.15$, $M_{\text{wise}} = 6.38$), were rated as the second and third most typical characteristics of older adults. However, the fourth and fifth ranked words, *ill* and *lonely*, were rated as negative characteristics ($M_{\text{ill}} = 2.08$, $M_{\text{lonely}} = 2.44$).

Which characteristics were possessor relevant or other relevant? Among adjectives consistently rated as possessor relevant were, for example, *sentimental*, *sleepy*, *depressed*, *strong*, *active*, *healthy*, and *playful*. In contrast, *authoritarian*, *dominant*, *brutal*, *gossipy*, *polite*, and *tolerant* were among the adjectives consistently rated as other relevant. As is suggested by these examples, possessor-relevant and other-relevant characteristics were both positive and nega-

tive. A related question was which characteristics showed the greatest valence difference between a possessor-relevant and an other-relevant perspective. In other words, some characteristics might be negative for others but positive for the possessor. Similarly, some characteristics might be positive for others but negative for the possessor. Among the characteristics that were rated as more positive for the self than for other people were *calculating* ($M_p = 3.63$, $M_o = 2.00$), *determining* ($M_p = 4.13$, $M_o = 2.69$), and *dominant* ($M_p = 3.81$, $M_o = 2.38$). In contrast, *compliant* ($M_p = 3.63$, $M_o = 5.00$), *lonely* ($M_p = 1.75$, $M_o = 3.00$), and *ill* ($M_p = 2.00$, $M_o = 3.19$) were more negative for the self than for others. The archived file contains these ratings as well.

Intercorrelations between rating dimensions. Table 2 provides the intercorrelation matrix between subjective ratings and objective measures of word frequency, word frequency class, and word length.³ Valence was clearly related to word frequency. Positive words were more frequent than negative words. There was no significant correlation between valence and word length. Valence was highly related to arousal ($r = -.62$) and control ($r = .64$) ratings. Negative words involved a more intense feeling of tension/arousal than did positive words; and positive words involved a greater degree of control than did negative words. The ratings of arousal and control were, however, unrelated ($r = -.09$). Regarding the three dimensions of valence, arousal, and control to-

Table 2
Correlations Between Overall Word Characteristics

Word Characteristic	1	2	3	4	5	6	7	8	9	10	11
1. Frequency											
2. Frequency class	-.74**										
3. Length in letters	-.15**	.20**									
4. Valence	.29**	-.34**	.05								
5. Arousal	-.16*	.21**	.15*	-.62**							
6. Control	.25**	-.17*	.13	.64**	-.09						
7. Imagery	.09	-.14	-.24**	-.06	.09	-.14					
8. Age relevance	-.01	-.08	-.05	-.06	-.29**	-.14	-.24**				
9. Self-relevance	.24**	-.26**	.15*	.91**	-.47**	.70**	-.15*	-.04			
10. Self-valence	.27**	-.29**	.12	.95**	-.51**	.75**	-.06	-.14	.92**		
11. Other valence	.28**	-.33**	.08	.96**	-.58**	.64**	-.08	-.04	.92**	.96**	
12. Self-other relevance	.18*	-.33**	.00	.16*	-.19**	-.19**	.11	.04	.05	.08	.13

* $p < .05$. ** $p < .01$.

gether, it is noteworthy that words formed approximately a two-dimensional disk skewed in the three-dimensional space. Thus, the three rating dimensions could probably be reduced to two latent dimensions. This suggests that valence, arousal, and dominance are nonorthogonal dimensions.

Valence ratings showed a high correlation with ratings of self-relevance ($r = .91$) as well, signifying that the more positive a word was evaluated to be, the more typical it was for the participants, on average. Although we had expected a high correlation between valence and self-relevance, this high a correlation was somewhat surprising. The correlations between valence and imagery and between valence and age relevance were not significantly different from zero. Positive and negative attributes were equally easy to imagine, and attributes assigned to older adults were not perceived as more negative than the ones assigned to young adults.

Arousal showed small but significant associations with word frequency and word length: High-arousing words were less frequent and shorter than low-arousing words. High-arousing words were also regarded as more typical for young adults than for older adults. In contrast, high-arousing words were used less often to describe the self. Words associated with an intense feeling of control were considered more frequent, more pleasant, and more typical for oneself. Another interesting aspect was the high correlation between control and self-relevance ($r = .70$). Personality characteristics that were rated as very typical for oneself were also rated as involving a strong feeling of control. In addition, easy-to-imagine words, in contrast to hard-to-imagine words, were shorter, were more typical for young adults, and were less often used to describe the self.

Possessor-relevant and other-relevant adjectives were rated similarly good or bad ($r = .96$). The rank order for self-valence and other valence corresponded highly with the normal valence rating. In contrast, both valence dimensions were unrelated to the distinction of whether a word is more self-relevant or other relevant. Self-relevant words, in contrast to other-relevant words, were more frequent, more positive, and less arousing, and involved less

often a feeling of control. However, these associations were generally small. The small correlations between self-other relevance and other word characteristics provide the opportunity to find sets of highly self-relevant and highly other-relevant words that are matched with respect to these other characteristics.

Consistency between present and previous ratings.

To examine the generalizability of the present ratings, we compared them with ratings available from previous German rating studies (Hager et al., 1985; Möller & Hager, 1991; Ostendorf, 1994; Schwibbe et al., 1981; Wippich & Bredenkamp, 1977).⁴ This comparison was possible for the dimensions of valence, control, imagery, arousal, and self-other relevance. In general, our ratings were highly consistent with previous ratings. In particular, judged valence showed very high consistency across studies ($.94 \leq r \leq .97$). Control showed high correlations with past ratings of potency or dominance ($.94 \leq r \leq .97$). Imagery was also highly related to previous imagery ($.86 \leq r \leq .89$) and concreteness ($.69 \leq r \leq .79$) ratings. One exception was the arousal dimension, which showed only moderate to high correlations with the arousal ratings from previous word-rating studies ($.30 \leq r \leq .75$). Wentura (personal communication, March 31, 2004) provided self-other relevance ratings for 52 words in the present set of 200 words. Correlations revealed a high consistency in self-other relevance between the present and Wentura's ratings ($r = .94$).

Overall word characteristics. Table 3 provides the overall means for the word characteristics across all 200 adjectives. In addition, we grouped the adjectives into positive, neutral, and negative words. This was done because many researchers might be interested in selecting words from these categories. On the basis of the overall valence score (combined across young and older adults), we classified words as negative ($M_{\text{valence}} < 3$), neutral ($3 \leq M_{\text{valence}} \leq 5$), and positive ($M_{\text{valence}} > 5$).⁵ This procedure resulted in 68 negative, 56 neutral, and 76 positive adjectives. Table 3 provides overall means separately for negative, neutral, and positive adjectives. For each word characteristic, we conducted an ANOVA with valence as

Table 3
Overall Means Across the Total of 200 Adjectives (T) and for Subgroups of 68 Negative (N), 56 Neutral (O), and 76 Positive (P) Adjectives

Word Characteristic	Means				Standard Deviations				n^2
	T	N	O	P	T	N	O	P	
Frequency	11.55	5.55	11.82	16.71	18.08	6.51	17.58	23.31	.07**
Frequency class	12.41	13.13	12.41	11.75	1.96	1.72	1.98	1.94	.07**
Length in letters	8.02	7.66	8.50	8.00	2.02	1.97	1.99	2.05	.03
Valence	4.09	2.23	4.01	5.86	1.28	0.50	0.60	0.44	.90**
Arousal	4.11	5.02	4.03	3.29	1.62	1.12	1.01	1.03	.33**
Control	4.08	3.17	3.95	4.98	1.23	1.04	1.12	0.73	.39**
Imagery	4.46	4.67	4.12	4.52	0.91	0.08	0.94	0.92	.06**
Age relevance	3.05	3.03	3.21	2.96	0.69	0.63	0.68	0.75	.02
Self-relevance	4.02	2.61	4.03	5.28	1.29	0.60	0.84	0.47	.76**
Self-valence	4.17	2.21	4.13	5.94	1.76	0.75	1.03	0.53	.81**
Other valence	4.11	2.24	4.09	5.80	1.67	0.68	0.90	0.54	.82**
Self-other relevance	1.62	1.55	1.64	1.65	0.34	0.38	0.29	0.32	.02

Note—The corresponding F tests for the main effect of valence had (2,197) degrees of freedom. ** $p < .01$.

Table 4
Intercorrelations Between Rating Dimensions for Young Adults
(Below Diagonal) and Older Adults (Above Diagonal)

Dimension	Dimension					
	V	A	C	I	AR	SR
Valence (V)	.91**	-.52**	.66**	-.03	.07	.90**
Arousal (A)	-.70**	.90**	.01	.12	-.39**	-.40**
Control (C)	.60**	-.18*	.92**	-.07	-.07	.71**
Imagery (I)	-.04	.05	-.16*	.80**	-.30**	-.10
Age relevance (AR)	-.16*	-.17*	-.29**	-.16*	.91**	.16*
Self-relevance (SR)	.88**	-.51**	.61**	-.10*	-.26*	.79**

Note—Values on the diagonal were correlations between young and older adults' ratings. * $p < .05$. ** $p < .01$.

a between-words factor. The analyses revealed, for practically all word characteristics, significant differences between positive, negative, and neutral words (see Table 3). Positive words were more frequent, less arousing, and higher in control than negative words were. Similarly, positive words were higher in self-relevance, self-valence, and other valence than negative words were. Exceptions were only word length, age relevance, and self-other relevance. Thus, positive, negative, and neutral words were equally long, were equally relevant for young and older adults, and were equally salient for the self and other people.

Young and Older Adults' Evaluations

Associations between young and older adults' ratings. We examined the correlations between young and older adults' ratings for the 200 words. Table 4 provides these correlations, as well as the intercorrelation matrix, for young and older adults separately. The correlations were extremely high for valence ($r = .91$), arousal ($r = .90$), control ($r = .92$), and age relevance ($r = .91$). The correlations between young and older adults' ratings for imagery ($r = .80$) and self-relevance ($r = .79$) were also high but somewhat smaller than those for the other dimensions. Older adults' imagery responses were distributed in a smaller range (3–6) than were young adults' responses (2–7). Self-relevance, in contrast, may involve more individual and heterogeneous responses than do other dimensions.

Age differences in mean evaluations. To address the question of age differences in emotional evaluations, we examined age differences in means across subsets of positive, negative, and neutral words. For the rating dimen-

sions, we conducted ANOVAs with age (young vs. old) as a within-words factor and valence (negative vs. neutral vs. positive) as a between-words factor. Please note that these analyses were done on the level of words and not on the level of persons. Table 5 provides means for young and older adults across all 200 adjectives and for subsets of 68 negative, 56 neutral, and 76 positive adjectives. In addition, effect sizes from the ANOVAs are displayed in Table 5. The main effects of age were generally small or nonsignificant. One exception was age relevance, with a large effect size: Older adults rated adjectives, on average, as more typical for older people than young adults did. The interactions between age group and word valence were all significant. Older adults rated positive adjectives as more positive, more arousing, less controllable, more easily imaginable, and more typical for older adults than young adults did. In contrast, older adults rated negative adjectives as less arousing, more controllable, and less typical for the self than young adults did.

Age differences for individual adjectives. To select individual words for experimental settings, it is important to know which adjectives showed age-related differences. We conducted multivariate and univariate ANOVAs for individual words. Age (young vs. old) functioned as a between-subjects factor. This procedure resulted in 6 (dimension) \times 200 (words) = 1,200 analyses on the univariate level and 200 analyses on the multivariate level. The large number of analyses resulted in significant effects by chance (approximately 5%). An overall MANOVA with the six rating dimensions as dependent variables (i.e., valence, arousal, control, imagery, self-relevance, and age

Table 5

Young and Older Adults' Ratings Across All 200 Adjectives (T) and for 68 Negative (N), 56 Neutral (O), and 76 Positive (P) Adjectives

Dimension	Means								Standard Deviations								η_p^2		
	Young Adults				Older Adults				Young Adults				Older Adults						
	T	N	O	P	T	N	O	P	T	N	O	P	T	N	O	P	A	V	A \times V
Valence	4.04	2.25	3.92	5.74	4.17	2.20	4.11	5.98	1.61	0.58	0.73	0.61	1.71	0.56	0.80	0.45	.03**	.90**	.03*
Arousal	4.08	5.11	4.09	3.15	4.09	4.93	3.98	3.43	1.36	1.13	1.05	1.08	1.27	1.16	1.07	1.06	.00	.33**	.12**
Control	4.08	3.06	3.90	5.13	4.07	3.28	4.01	4.83	1.40	1.21	1.30	0.76	1.10	0.93	0.99	0.76	.00	.40**	.16**
Imagery	4.41	4.73	4.05	4.39	4.51	4.61	4.19	4.65	1.10	0.94	1.15	1.12	0.82	0.76	0.83	0.80	.02*	.06**	.06**
Age relevance	3.00	3.00	3.18	2.85	3.11	3.08	3.23	3.06	0.73	0.64	0.72	0.79	0.69	0.64	0.67	0.75	.12**	.02	.05**
Self-relevance	4.11	2.90	4.06	5.23	3.93	2.33	4.00	5.32	1.23	0.74	0.87	0.60	1.51	0.63	1.13	0.69	.04**	.76**	.09**

Note— η_p^2 values are for the main effects of age group (A) and valence category (V), as well as for the interaction between age group and valence (A \times V). Adjectives were classified on the basis of the overall valence score. Thus, some words might be classified differently when young and older adults' evaluations are considered separately. * $p < .05$. ** $p < .01$.

relevance) revealed significant main effects of age for approximately half of the words (93; 46.5% of all 200 words showed significant main effects for age).

As was expected from the multivariate analyses, the univariate analyses (for the six dimensions separately) revealed a substantial number of significant main effects of age for valence (62 words with significant main effects of age; 31% of 200 words), arousal (42; 21%), control (31; 15.5%), imagery (27; 13.5%), self-relevance (72; 36%), and age relevance (26; 13%). Thus, despite the extremely high correlations between the young and the older adults' ratings, both differed dramatically in their mean ratings for these words. Figure 1 shows scatterplots between young and older adults' ratings. The archived file contains these wordwise analyses as well.

DISCUSSION

The goal of the present studies was to establish the AGE database and to make these age-dependent ratings available in the public domain. We conducted this research because there is an increasing interest in age differences in the social cognitive and cognitive neuroscience fields. To date, there are few published word stimulus databases that have been evaluated by young and older raters. We obtained extensive ratings on nine dimensions for 200 well-selected adjectives. In particular, we asked young experts (psychology graduate students) to provide information on the self-relevance, other relevance, and self-other relevance of these adjectives. In addition, young and older adults rated adjectives on six dimensions: valence, arousal, control, imagery, age relevance, and self-relevance. These data were collected to form the backbone of the AGE database.

Associations Between Dimensions

Consistent with past findings (Ortony, Clore, & Foss, 1987), frequent words were more likely to be rated as positive than as negative. This signals that we more often use positive than negative words in our language (e.g., Mehl & Pennebaker, 2003). Valence and self-relevance were highly correlated, indicating that positive words were very typical and negative words were very untypical for participants' self-concept, on average. It is not surprising that we generally use positive words to describe ourselves and not negative words. However, the size of this correlation was as high as the correlation between young and older adults' valence ratings. That said, it seems to be quite difficult to identify negative attributes of oneself. Positive words were less arousing and were associated with more control than negative words were. High-arousing words were longer and less frequent than low-arousing words. In addition, difficult-to-imagine words were longer than easy-to-imagine words. However, imagery was not related to any emotional facet (i.e., valence, arousal, and control). This provides the opportunity to find positive and negative words that are similarly imaginable.

Interestingly, valence and age relevance were not correlated: Attributes assigned to older adults were not more negative or positive than attributes assigned to young

adults. This finding contrasts with the negative stereotypes about older adults and highlights the potential of older adults. Despite the loss of functioning in some domains, older adults are considered able to learn new skills and to adapt to new challenges (Heckhausen et al., 1989). Thus, attributes of older adults are different from young adults' attributes but are equally positive. In contrast, attributes assigned to young adults were more arousing and more easily imaginable than attributes assigned to older adults. These ratings might be helpful for researchers interested in choosing attributes for experimental investigations in age stereotype activation. In addition, ratings of self-other relevance will be useful for research on processing social information (e.g., Wentura et al., 2000).

The Impact of Age: Consensus and Differences

With regard to participants' age, the AGE ratings revealed two major findings. First, age groups agreed on the rank order of attributes, as is evident from the high correlation between young and older adults' evaluation. The high consistency maps onto findings obtained in previous experiments with word material. For example, Wurm, Labouvie-Vief, Aycock, Rebusal, and Koch (2004) found a high correlation between young and older adults' ratings for valence ($r = .93$) and arousal ($r = .88$). This indicates that young and older adults agree on whether a word is more positive or more negative than another word (see also Heckhausen et al., 1989). High correlations between young and older adults' ratings were also found for the other dimensions.

Second, despite the high correlations between age groups, age-related mean differences were evident for all six rating dimensions. In particular, main mean differences between young and older adults were generally small. This may indicate that age groups used similar mean anchors for the scales. Age differences for positive, negative, and neutral adjectives were, however, more pronounced. Older adults rated positive attributes as more positive, more arousing, less controllable, more easily imaginable, and more relevant for older adults than young adults did. Older adults also rated negative attributes as less typical for themselves than young adults did. These differences showed that young and older adults perceived positive and negative attributes differently. This pattern of findings may also suggest that young and older adults process positive and negative attributes differently.

These findings of overall mean-level differences between young and older adults were also supported by analyses of individual words. In particular, young and older adults showed significant mean-level differences in valence, arousal, control, imagery, self-relevance, and age relevance for 31%, 21%, 15.5%, 13.5%, 36%, and 13% of all 200 words, respectively. These findings emphasize the importance of assessing young and older adults' perceptions of stimulus material and the need for systematic investigations of age-related differences in the perception of emotional material. The AGE database provides these ratings, as well as ANOVAs for individual words.

In sum, in general, young and older adults agree on whether a word is more or less positive than another word,

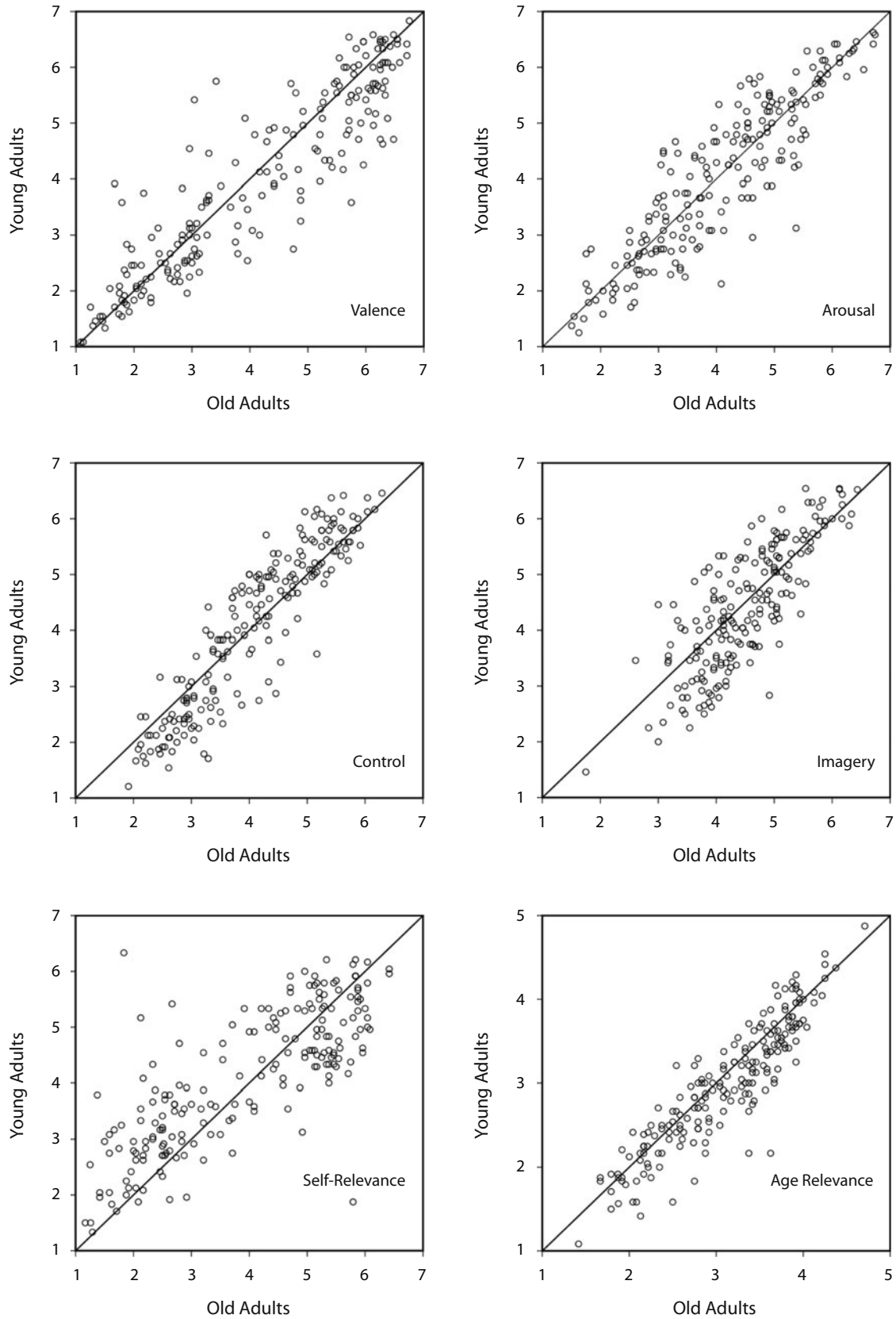


Figure 1. Scatterplots between young and older adults' ratings for each dimension. Each dot represents one word. The diagonal stands for perfect agreement between age groups ($r = 1.00$).

but they differ for a large number of words in their exact evaluations of how positive or negative these words are. A high correlation between young and older adults' ratings does not guarantee that material is perceived similarly at all levels. These mean differences have practical implications: For example, some words may be more positive, more arousing, or more relevant for older adults than for young adults, resulting in biased age differences in other domains (e.g., memory).

Summary and Conclusions

The present study provides data about age-related differences in word ratings. Although ratings were highly consistent between young and older adults, there were age differences in ratings for a large number of words. In future work, we want to expand the AGE database to include (1) a larger number of adjectives collected from (2) all periods of the life span. In order to provide an appropriate number of words for many experimental designs, we carefully selected the present item pool. Some researchers might simply need a larger number of items than that which is currently available. Similarly, we provided ratings by young and older adults. It would be desirable to have ratings from other age groups, especially from adolescents and middle-aged adults. In addition, it would be desirable to add ratings from other languages. The AGE database consists of German adjectives. It is an open question whether age differences would be similar in other languages or other cultural environments. Given that the semantic meanings of words are mainly culturally based and influenced by historical changes, different age patterns are likely in different cultural contexts.

To conclude, our data provide a useful source for experiments in which the effects of aging are considered and affective word material is used. Although many adjectives were rated differently by young and older adults, many words were not. Thus, using the AGE word pool, researchers will be able to select words that are matched across age groups for future research.

AUTHOR NOTE

Correspondence concerning this article should be addressed to D. Grün, Department of Psychology, North Carolina State University, Box 7650, 640 Poe Hall, Raleigh, NC 27695-7650, or to J. Smith, Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043 (e-mail: dgruehn@ncsu.edu or smitjacq@isr.umich.edu).

REFERENCES

- ANDERSEN, S. M., & CHEN, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review*, *109*, 619-645.
- ANGLEITNER, A., OSTENDORF, F., & JOHN, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality*, *4*, 89-118.
- BATSON, C. D., SAGER, K., GARST, E., KANG, M., RUBCHINSKY, K., & DAWSON, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality & Social Psychology*, *73*, 495-509.
- BUSZ, M., COHEN, R., POSER, U., SCHÜMER, A., SCHÜMER, R., & SONNENFELD, C. (1972). Die soziale Bewertung von 880 Eigenschaftsbegriffen sowie die Analyse der Ähnlichkeitsbeziehungen zwischen einigen dieser Begriffe [The social evaluation of 880 attribute concepts and the analysis of similarity relationships between some of these concepts]. *Zeitschrift für Experimentelle & Angewandte Psychologie*, *19*, 282-308.
- CHARLES, S. T., MATHER, M., & CARSTENSEN, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, *132*, 310-324.
- GOLDBERG, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*, 26-42.
- GRÜHN, D., & SCHEIBE, S. (2008). Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age? *Behavior Research Methods*, *40*, 512-521.
- GRÜHN, D., SCHEIBE, S., & BALTES, P. B. (2007). Reduced negativity effect in older adults' memory for emotional pictures: The heterogeneity-homogeneity list paradigm. *Psychology & Aging*, *22*, 644-649.
- GRÜHN, D., SMITH, J., & BALTES, P. B. (2005). No aging bias favoring memory for positive material: Evidence from a heterogeneity-homogeneity list paradigm using emotionally toned words. *Psychology & Aging*, *20*, 579-588.
- HAGER, W., & HASSELHORN, M. (1994). *Handbuch deutschsprachiger Wortnormen* [Handbook of German word norms]. Göttingen: Hogrefe.
- HAGER, W., MECKLENBRÄUKER, S., MÖLLER, H., & WESTERMANN, R. (1985). Emotionsgehalt, Bildhaftigkeit, Konkretheit und Bedeutungshaltigkeit von 580 Adjektiven: Ein Beitrag zur Normierung und zur Prüfung einiger Zusammenhangshypothesen. [Emotionality, imagery, concreteness and meaningfulness for 580 adjectives: Norm values and examination of some hypotheses concerning the different variables]. *Archiv für Psychologie*, *137*, 75-97.
- HECKHAUSEN, J., DIXON, R. A., & BALTES, P. B. (1989). Gains and losses in development throughout adulthood as perceived by different adult age groups. *Developmental Psychology*, *25*, 109-121.
- HEYDECKE, J. (1984). *Kontentanalytischer Versuch zur Messung affektiver Bedeutung* [Content-analytical attempt of measuring affective meaning]. Unpublished thesis, Georg-August-Universität, Göttingen.
- KENSINGER, E. A., BRIERLEY, B., MEDFORD, N., GROWDON, J. H., & CORKIN, S. (2002). Effects of normal aging and Alzheimer's disease on emotional memory. *Emotion*, *2*, 118-134.
- KLAPPROTT, J. (1972). Erwünschtheit und Bedeutung von 338 alltagspsychologischen Eigenschaftsbegriffen [Desirability and meaning of 338 everyday psychological trait names]. *Psychologische Beiträge*, *14*, 496-520.
- LABOUVIE-VIEF, G. (2005). Self-with-other representations and the organization of the self. *Journal of Research in Personality*, *39*, 185-205.
- LANG, P. J., BRADLEY, M. M., & CUTHBERT, B. N. (1998). *International Affective Pictures System (IAPS): Digitized photographs, instruction manual and affective ratings* (Tech. Rep. A-6). Gainesville: University of Florida.
- LÖCKENHOFF, C. E., & CARSTENSEN, L. L. (2004). Socioemotional selectivity theory, aging, and health: The increasingly delicate balance between regulating emotions and making tough choices. *Journal of Personality*, *72*, 1395-1424.
- MATHER, M., & CARSTENSEN, L. L. (2003). Aging and attentional biases for emotional faces. *Psychological Science*, *14*, 409-415.
- MECKLENBRÄUKER, S., HAGER, W., & MÖLLER, H. (1994). Erwünschtheit von 908 Eigenschaften bei Frauen und bei Männern aus der Sicht von Männern und Frauen [Desirability of 908 traits of women and men from the perspective of men and women]. In W. Hager & M. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen* (pp. 310-328). Göttingen: Hogrefe.
- MEHL, M. R., & PENNEBAKER, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality & Social Psychology*, *84*, 857-870.
- MÖLLER, H., & HAGER, W. (1991). Angenehmheit, Bedeutungshaltigkeit, Bildhaftigkeit und Konkretheit-Abstraktheit von 452 Adjektiven [Pleasantness, meaningfulness, imagery, and concreteness-abstractness of 452 adjectives]. *Sprache & Kognition*, *10*, 39-51.
- MUELLER, J. H., WONDERLICH, S., & DUGAN, K. (1986). Self-referent processing of age-specific material. *Psychology & Aging*, *1*, 293-299.
- ORTONY, A., CLORE, G. L., & FOSS, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*, 341-364.

- OSTENDORF, F. (1994). Zur Taxonomie deutscher Dispositionsbegriffe [A taxonomy of German dispositional terms]. In W. Hager & M. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen* (pp. 382-441). Göttingen: Hogrefe.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, **76**, 1-25.
- SCHWIBBE, M., RÄDER, K., SCHWIBBE, G., BORCHARDT, M., & GEIKEN-POPANKEN, G. (1981). Zum emotionalen Gehalt von Substantiven, Adjektiven und Verben [The emotional content of substantives, adjectives, and verbs]. *Zeitschrift für Experimentelle & Angewandte Psychologie*, **28**, 486-501.
- STEYER, R., SCHWENKMEZGER, P., NOTZ, P., & EID, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)* [The multidimensional mood questionnaire]. Göttingen: Hogrefe.
- WATSON, D., & CLARK, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule—Expanded Form*. Iowa City: University of Iowa, Department of Psychology.
- WECHSLER, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised (WAIS-R)*. New York: Psychological Corp.
- WENTURA, D., ROTHERMUND, K., & BAK, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality & Social Psychology*, **78**, 1024-1037.
- WIPPICH, W., & BREDEKAMP, J. (1977). Bestimmung der Bildhaftigkeit, Konkretheit und der Bedeutungshaltigkeit von 498 Verben und 400 Adjektiven [Identification of imagery, concreteness, and meaningfulness of 498 verbs and 400 adjectives]. *Zeitschrift für Experimentelle & Angewandte Psychologie*, **24**, 671-680.
- WURM, L. H., LABOUIE-VIEF, G., AYCOCK, J., REBUCAL, K. A., & KOCH, H. E. (2004). Performance in auditory and visual emotional Stroop tasks: A comparison of older and younger adults. *Psychology & Aging*, **19**, 523-535.
- ZIPF, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Boston: Houghton-Mifflin.

NOTES

1. The distribution of simple word frequencies is highly skewed and follows a function called *Zipf's law* (Zipf, 1935): There are only a few very frequent words and numerous very rare words. Word frequency classes are derived by considering the frequency of the word of interest (f_{word}) and the frequency of the most frequent word in a language (f_{der}). In German, the most frequent word is *der* (i.e., masculine form of *the*), which accounts for approximately 2%–3% of all written text. The formula is: $WFC = \log_2(f_{\text{word}}/f_{\text{der}})$. Only the whole-numbered part is taken from the exact result. Frequency classes are in reversed order:

High-frequent words have low-frequency classes (e.g., $WFC = 7$), whereas very rare words have high-frequency classes (e.g., $WFC = 18$). Whereas simple word frequencies are highly skewed, word frequency classes follow approximately a normal distribution. These superior distribution properties are especially relevant for many test statistics that assume a normal distribution of the considered variables. Word frequency data were obtained from the Web-based project *Deutscher Wortschatz* (wortschatz.uni-leipzig.de/).

2. The exact instructions for each dimension in Study 1 and Study 2 are available from the first author.

3. Word frequency and word frequency class were highly correlated ($r = -.74$). Both measures of word frequency also showed correlation patterns similar to other measures. However, the correlations seemed to be a little higher for word frequency class than for simple word frequency. This may be due to the superior distribution properties of word frequency classes over word frequencies. Moreover, word frequency (and word frequency class) was slightly related to word length, indicating that high-frequent words were typically shorter than low-frequent words.

4. We only considered rating studies that had at least 40 words (20% of all words) in common with the present item pool.

5. Other classifications are possible as well. For example, words might be classified only as negative (<5) or positive (>5). It is also noteworthy that this classification was based on the overall valence score by young and older adults. When young and older adults' ratings were considered separately, some words were classified in different valence groups. These were mainly words close to the valence boundaries.

ARCHIVED MATERIALS

The following materials may be accessed through the Psychonomic Society's Norms, Stimuli, and Data Archive, www.psychonomic.org/archive/.

To access these files, search the archive for this article, using the journal (*Behavior Research Methods*), the first author's name (Gruhn), and the publication year (2008).

FILE: Gruhn-BRM-2008.zip

DESCRIPTION: The compressed archive file contains two files:

Gruhn_AGE.txt, containing word characteristics of the age-dependent evaluations of German adjectives (AGE) developed by Grünh and Smith (2008) as a tab-delimited text file.

Gruhn_ReadMe.txt, containing a brief description of variables in the data file (Gruhn_AGE) as a text file.

AUTHOR'S E-MAIL ADDRESS: dgruehn@ncsu.edu.

(Manuscript received March 17, 2008;
revision accepted for publication June 19, 2008.)