Words high and low in pleasantness as rated by male and female college students

FRANCIS S. BELLEZZA
Ohio University, Athens, Ohio

and

ANTHONY G. GREENWALD and MAHZARIN R. BANAJI
Ohio State University, Columbus, Ohio

No catalog of words currently available contains normative data for large numbers of words rated low or high in affect. A preliminary sample of 1,545 words was rated for pleasantness by 26-33 college students. Of these words, 274 were selected on the basis of their high or low ratings. These words, along with 125 others (Rubin, 1981), were then rated by additional groups of 62-76 college students on 5-point rating scales for the dimensions of pleasantness, imagery, and familiarity. The resulting mean ratings were highly correlated with the ratings obtained by other investigators using some of the same words. However, systematic differences in the ratings were found for male versus female raters. Females tended to use more extreme ratings than did males when rating words on the pleasantness scale. Also, females tended to rate words higher on the imagery and familiarity scales. Whether these sex differences in ratings represent cognitive differences between the sexes or merely differences in response style is a question that can be determined only by further research.

The role of affect, mood, and emotional state in cognitive processes has been, and remains, an active area of research (e.g., Isen, 1984). Studies on this topic have investigated memory for words that have affective or emotional connotation (e.g., Anisfeld & Lambert, 1966; Bousfield, 1950; Broadbent & Gregory, 1967; Cason & Lungren, 1932; Levinger & Clark, 1961; Teasdale & Russell, 1983; White, 1936). A. S. Brown (1976) has provided a catalog of published studies reporting scaled verbal material. Words have been rated on such affective dimensions as pleasantness, desirability, goodness, and likability. Although these evaluative terms have slightly different meanings, the ratings of words on these scales are highly intercorrelated. For example, W. P. Brown and Ure (1969) found a correlation of .96 between the goodness and pleasantness scales, and Rubin (1980) found a correlation of .95.

Each of the existing sets of normative data on pleasantness and related dimensions has significant limitations. The likability norms of Anderson (1968) are extensive but use only personality trait words. The norms of W. P. Brown and Ure (1969) include 650 words comprising different parts of speech, but few of them are rated as very pleasant or unpleasant. Also, W. P. Brown and Ure's words include sets of words with the same root, such as *beauty, beautiful,* and *friend, friendly.*

The lack of a large number of pleasant and unpleasant words is the major problem in most of the published norms of pleasantness-goodness. For example, Rubin's (1981) extensive analysis of the rated characteristics of nouns was performed on only 125 words. Other lists also contain too few words (Ledgerwood, 1932; Silverstein & Dienstbier, 1968). Toglia and Battig (1978) include pleasantness ratings for 2,854 words and letter strings. However, as with the W. P. Brown and Ure norms, few of these words receive high or low mean pleasantness ratings.

Words for which norms on various affective dimensions have been obtained have not been rated on other dimensions. This is significant, because the manner in which a word is processed in any cognitive task may depend not only on its affective characteristics but also on other dimensions, such as imagery, familiarity, and meaningfulness (Paivio, 1971; Rubin, 1980; Toglia & Battig, 1978).

The purpose of the present study was to establish norms for ratings of pleasantness, visual imagery, and familiarity for a sample of words that was selected to include many that should be rated low or high in pleasantness. Because the affective dimensions are highly intercorrelated, only pleasantness was rated. Pleasantness is the scale most commonly used by previous investigators (W. P. Brown & Ure, 1969; Rubin, 1980; Toglia & Battig, 1978).

METHOD

Phase 1

A total of 1,545 words from Thorndike and Lorge (1944) were judged by the experimenters as likely to be either pleasant or unpleasant for a large number of people. The list of words was arranged in alphabetical order. From each set of five adjacent words, one word was assigned to each of the five forms of test booklet. Hence, each test form was made up of 309 words. Each test form was administered to a group of college students who received course credit in introductory psychology courses for their participation. The number of subjects tested using each form varied from 26 to 33. The subjects rated each word on a 5-point scale of pleasantness, with a rating of 1 meaning *very unpleasant* and a rating of 5 meaning *very pleasant*. A 5-point rating scale was used so that the response sheets could be machine scored by the computer facilities available to the authors. The instructions were as follows:

> The purpose of this experiment is to find out whether or not college students have positive or negative feelings about different words. Words differ in the kinds of emotions that they can make people feel. The purpose of this experiment is to have you rate a list of approximately 300 words with regard to how pleasant or unpleasant they are; that is, how pleasant or unpleasant they make you feel. You should read each word very carefully. Then after you read it, fill in one of the circles on the response sheet that has the same identification number as the word you are rating. Make sure that you fill in each circle completely. Use the 5-point scale on the front page of the booklet. If the word has a *very pleasant* meaning for you, then rate if a 5. If the word has a *somewhat pleasant* meaning for you, then rate it a 4. If the word has *no pleasant or unpleasant* meaning for you, then rate it a 3. If the word has a *somewhat unpleasant* meaning for you, then rate it a 2. Finally, if the word has a *very unpleasant* meaning, then rate it a 1. Try to use all 5 points on the rating scale. Remember to read each word carefully before you rate it. Make your rating on the basis of how you *feel* about the word, not on how you thi..k people in general would rate the word. If you come across a word that you do not know, then do not rate it. Leave the spaces blank on the response sheet. Also, if you come across a word that you think is spelled wrong, do not rate it. However, try to rate as many words as you can. You have to spend at least 30 min doing this task. So try to take at least 5 sec to rate each word. Any questions?

Subjects were instructed not to rate words that seemed to be spelled incorrectly in order to eliminate misinterpretations of words. This happened rarely, and most subjects rated all the words.

After the mean ratings were computed, a number of words were eliminated in the following manner. First, for each group of words with the same root, such as *addict*, *addicted*, and *addiction*, only that word with the most extreme rating was included. Second, to obtain approximately equal numbers of pleasant and unpleasant words, those words with mean ratings above 4.25 ($n = 129$) and

below 2.00 ($n = 145$) were retained. To these 274 words were added Rubin's (1981) 125 nouns, permitting the total set of 399 words to include words that would not be perceived as either pleasant or unpleasant. Hence, in the testing done during Phase 2, the subjects would not infer that all the words in the list had to be judged as being either pleasant or unpleasant.

Phase 2

The final list of 399 words was randomized five times to create five different test forms. These words were again rated on a 5-point scale of pleasantness by 40 males and 36 females. The instructions were the same as those used in Phase 1, except that the subjects were told to rate approximately 400 words rather than 300. In addition, 27 males and 35 females rated the words on a 5-point scale of visual imagery, and 29 males and 34 females rated the words on a 5-point scale of familiarity. The subjects were tested in groups ranging from 15 to 25 subjects, and no subject rated the words on more than one scale. The subjects in the imagery condition were given instructions for the imagery rating task as follows:

> The purpose of this experiment is to find out whether or not college students can form mental pictures or visual images for different words. Words differ in how much visual imagery they cause people to have. The purpose of this experiment is to have you rate a list of approximately 400 words with regard to how easily you can form a visual image or mental picture of what the word means or represents. You should read each word very carefully. Then after you read it, fill in one of the circles on the response sheet that has the same identification number as the word you are rating. Make sure that you fill in each circle completely. Use the 5-point scale on the front page of the booklet. If the word creates a *very clear and vivid visual image* in your mind, then you should give the word a rating of 5. If the word creates a *clear mental picture* for you, then rate it a 4. If the image created by the word is *somewhat unclear*, then rate it a 3. If your mental picture is *very unclear*, then give the word a rating of 2. Finally, if you can form *no image at all*, give the word a rating of 1. Try to use all 5 points on the rating scale. Remember to read each word carefully before you rate it. Make your rating on the basis of how easily you can form a visual image for the word, not on how you think people in general would rate the word. If you come across a word that you do not know, then do not rate it. Leave the spaces blank on the response sheet. Also, if you come across a word that you think is spelled wrong, do not rate it. However, try to rate as many words as you can. You have to spend at least 30 min doing this task. Try to take at least 5 sec to rate each word. Any questions?

The subjects rating the words for familiarity were given the following instructions:

> The purpose of this experiment is to find out how frequently college students come across various English words. Words differ in how frequently they are used and encountered in language. Some words such as the words *go*, *eat*, *drink*, *table*, and *want* are very frequently encountered by col-

lege students, because you use these words in your speech, hear other people use these words, use these words in writing, and also read the words in books and newspapers. Other words such as *spanner*, *warlock*, *surtax*, and *archive* are less frequently encountered by college students, because these words are not often spoken, heard, read, or written. The purpose of this experiment is to have you rate a list of approximately 400 words with regard to how frequently you either use these words or come across them in speech and writing. You should read each word very carefully. Then *after* you *read* it, fill in one of the circles on the response sheet that has the same identification number as the word you are rating. Make sure that you fill in each circle completely. Use the 5-point scale on the front page of the booklet. If the word is one that you encounter *very frequently*, that is, one you often read, hear, and use, then you should give the word a rating of 5. If the word is one you encounter *fairly frequently*, then rate it a 4. If the word is encountered *often*, then rate it a 3. If the word appears *somewhat infrequently*, then give the word a rating of 2. Finally, if the word is one that appears *very infrequently*, then give the word a rating of 1. Try to use all 5 points on the rating scale. Remember to read each word carefully before you rate it. If you come across a word that you do not know, then do not rate it. Leave the spaces blank on the response sheet. Also, if you come across a word that you think is spelled wrong, do not rate it. However, try to rate as many words as you can. You have to spend at least 30 min doing this task. Try to take at least 5 sec to rate each word. Any questions?

RESULTS

Mean pleasantness, imagery, and familiarity ratings were computed for each of the 399 words. Using these 399 words as a sample, it was found that the product moment correlation between the mean pleasantness ratings and the mean imagery ratings was .26. The correlation between the pleasantness and familiarity ratings was .46, and between imagery and familiarity was .21. Hence, these three attributes of the words were not correlated with one another to the extent that the attributes were redundant.

Interstudy Reliabilities

The mean ratings obtained in this study were correlated with ratings found in previous studies for various subsets of the words. The present pleasantness ratings correlated .97 (155 words) with the goodness ratings reported by Rubin (1981) and correlated .87 (73 words) with the pleasantness ratings reported by Toglia and Battig (1978). The imagery ratings correlated .94 (148 words) with the imagery ratings of Paivio, Yuille, and Madigan (1968) and correlated .85 (73 words) with the imagery ratings reported by Toglia and Battig (1978). Finally, the familiarity ratings correlated .72 (73 words) with the familiarity ratings reported by Toglia and Battig (1978). All of these correlations are high except for the correlations between the collected familiarity ratings and those reported by

Toglia and Battig (1978). It appears that familiarity ratings are less reliable than ratings of pleasantness or imagery.

Male-Female Differences

Correlations were computed between the mean ratings provided by the males and females over the 399 words. The correlation was .99 on the pleasantness scale, .93 on the imagery scale, and .90 on the familiarity scale. Despite this general evidence of between-gender agreement, there were differences in the means and distributions of ratings for males and females on the three scales.

Overall Mean Differences

The mean pleasantness rating give by males was 3.10, and that given by females was 3.06. Computing a correlated t test over the 399 words, this difference was small but significant [$t(398) = 3.19, p < .002$]. The mean imagery rating given by males was 3.57 and by females was 3.85 [$t(398) = 20.95, p < .001$]. The mean familiarity rating given by males was 2.87 and by females was 3.03 [$t(398) = 8.41, p < .001$]. In summary, males rated words as slightly more pleasant than did females, but females rated words as higher in imagery and as more familiar than did males.

Sex Differences on Individual Words

The Appendix provides a sample of 36 words for which the mean male and female ratings differed at the $p < .05$ level on one or more of the three scales. The complete set of ratings for the 399 words can be obtained from Francis S. Bellezza, Department of Psychology, Ohio University, Athens, Ohio 45701. Of the 19 unpleasant words rated differently by males and females, all were rated higher (i.e., less unpleasant) by males. Of the 24 pleasant words rated differently by males and females, 6 were rated as more pleasant by males. Of the 58 words rated differently in visual imagery, only 1 was rated higher in imagery by males (*execution*). Of the 59 words rated differently in familiarity by males and females, only 7 were rated higher by males.

Sex Differences in Extremity of Pleasantness Ratings

The distributions of ratings on the three scales for males and females are given in Table 1. The distributions of ratings for imagery and familiarity support the result that females rated the words higher than males on both scales. Ratings for pleasantness are interesting because they indicate that females rated unpleasant words as being more unpleasant and pleasant words as being more pleasant than did males. Therefore, although the means were comparable for males and females on the pleasantness scale, the females gave more extreme ratings. An overall test of this sex difference in extremity of pleasantness ratings is given by the correlation across the 399 words of each word's

Table 1
Distribution of Mean Ratings by Males and Females on the
Pleasantness, Imagery, and Familiarity Scales for the 399 Words

	Mean Rating Ranges				
	1.00-1.50	1.51-2.50	2.51-3.50	3.51-4.50	4.51-5.00
	Pleasantness				
Males	.07	.35	.09	.43	.06
Females	.19	.23	.09	.34	.15
	Imagery				
Males	.00	.05	.46	.36	.13
Females	.00	.01	.34	.39	.26
	Familiarity				
Males	.01	.34	.46	.18	.01
Females	.02	.28	.38	.28	.04

Note—Cell entries are proportions of the total of 399 words. The sum of entries in each row is 1.00.

pleasantness rating with its female-minus-male difference in mean pleasantness. This correlation was $r = .52$, $df = 397$, $p < .001$.

GENERAL DISCUSSION

The reason for collecting the normative data reported here was to provide a large number of words rated as being positive or negative in pleasantness for use in studying the role of affect in the cognitive processing of words. The mean pleasantness ratings for our sample of 399 words are presented for males and females separately and for males and females combined. Imagery and familiarity ratings for these words are also presented so that these factors can be controlled or otherwise taken into account. Females tended to rate words as more extreme on the pleasantness scale than did males, even though the correlation between the male and female mean pleasantness ratings over the 399 words was .99. Also, when using the imagery and familiarity scales, females tended to rate the words higher than did males.

Whether these differences represent underlying cognitive and affective differences between males and females or represent differences in response style is a question that cannot be answered here. It seems that the results cannot be attributed to any simple gender-associated response bias, because the gender difference patterns for pleasantness were different from those of imagery and familiarity. However, if the pleasantness scale is considered as a composite of two dimensions, pleasantness and unpleasantness, then it is reasonable to assume that females tend to give higher ratings both on pleasantness and on unpleasantness.

In an initial study using the normative data reported here, Banaji, Bellezza, and Greenwald (1985) asked college students to rate the 399 words on scales of pleasantness, emotionality, and imagery. This rating task was followed by a surprise test of free recall. In their rating data, Banaji et al. found that females rated words more extremely pleasant or unpleasant than did males, and that more extremely pleasant and unpleasant words were bet-

ter remembered, replicating previous findings by Kanungo and Dutta (1966). Importantly, however, Banaji et al. found no difference in the effect of affective intensity on recall for males versus females. If the ratings actually represented true sex differences in affective experience of the words, then females should have recalled relatively more of the extremely pleasant and unpleasant words. This did not occur. Further research is needed to determine whether the difference in ratings found between males and females represents response bias or some more basic cognitive mechanism.

REFERENCES

ANDERSON, N. H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality & Social Psychology*, **9**, 272-279.
ANISFELD, M., & LAMBERT, W. E. (1966). When are pleasant words learned faster than unpleasant words? *Journal of Verbal Learning & Verbal Behavior*, **5**, 132-141.
BANAJI, M. R., BELLEZZA, F. S., & GREENWALD, A. G. (1985, May). *Are women more emotional? Gender differences in reported emotional response do not translate to recall.* Paper presented at the meeting of the Midwestern Psychological Association, Chicago, IL.
BOUSFIELD, W. A. (1950). The relationship between mood and the production of affectively toned associations. *Journal of General Psychology*, **42**, 67-85.
BROADBENT, D. E., & GREGORY, M. H. P. (1967). The perception of emotionally toned words. *Nature*, **215**, 581-584.
BROWN, A. S. (1976). Catalog of scaled verbal material. *Memory & Cognition*, **4(1B)**, 1S-45S.
BROWN, W. P., & URE, D. M. J. (1969). Five rated characteristics of 650 word association stimuli. *British Journal of Psychology*, **60**, 233-249.
CASON, H., & LUNGREN, F. C. (1932). Memory for pleasant, unpleasant, and indifferent words. *Journal of Experimental Psychology*, **15**, 728-732.
ISEN, A. (1984). Toward understanding the role of affect in cognition. In R. Wyer & T. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179-236). Hillsdale, NJ: Erlbaum.
KANUNGO, R. N., & DUTTA, S. (1966). Retention of affective material: Frame of reference or intensity? *Journal of Personality & Social Psychology*, **4**, 27-35.
LEDGERWOOD, R. (1932). A comparison of methods in determining the affective value of words. *American Journal of Psychology*, **44**, 796-797.
LEVINGER, G., & CLARK, J. (1961). Emotional factors in the forgetting of word associations. *Journal of Abnormal & Social Psychology*, **62**, 99-105.
PAIVIO, A. (1971). *Imagery and verbal processes.* New York: Holt.
PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, **76**(1, Pt. 2).
RUBIN, D. C. (1980). 51 properties of 125 words: A unit analysis of behavior. *Journal of Verbal Learning & Verbal Behavior*, **19**, 736-755.
RUBIN, D. C. (1981). Norms for 34 properties of 125 words. *JASA Catalog of Selected Documents in Psychology*, **11**, 19. (Ms. No. 2213)
SILVERSTEIN, A., & DIENSTBIER, R. A. (1968). Rated pleasantness and association value of 101 English nouns. *Journal of Verbal Learning & Verbal Behavior*, **7**, 81-86.
TEASDALE, J. D., & RUSSELL, M. L. (1983). Differential aspects of induced mood on the recall of positive, negative, and neutral words. *British Journal of Clinical Psychology*, **22**, 163-171.
THORNDIKE, E. L., & LORGE, I. (1944). *The teacher's word book of 30,000 words.* New York: Teachers College Press, Columbia University.
TOGLIA, M. P., & BATTIG, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.
WHITE, M. M. (1936). Some factors influencing the recall of pleasant and unpleasant words. *American Journal of Psychology*, **48**, 134-139.

APPENDIX
Mean Ratings of Pleasantness, Imagery, and Familiarity by
Males and Females for a Sample of 36 Words

	Pleasantness			Imagery			Frequency		
	Male	Female	Combined	Male	Female	Combined	Male	Female	Combined
adorable	4.33	4.47	4.40	3.33	3.63	3.50	2.66	3.76	3.25*
agility	4.28	3.67	3.99*	3.22	2.97	3.08	2.38	2.21	2.29
breast	4.21	3.57	3.91*	4.78	4.66	4.71	3.79	3.18	3.46
capacity	3.47	3.03	3.27*	2.65	2.69	2.67	2.66	2.53	2.59
cemetery	2.00	1.31	1.67*	4.78	4.74	4.75	2.38	1.91	2.13
cheer	4.34	4.69	4.51*	3.33	3.54	3.45	2.90	3.35	3.14
clothing	3.95	4.31	4.12*	4.26	4.91	4.62*	3.90	4.06	3.98
corpse	1.57	1.44	1.51	4.07	4.14	4.11	2.45	1.79	2.10*
cozy	4.38	4.53	4.45	3.15	3.69	3.45	2.17	3.26	2.76*
cuddle	4.32	4.64	4.47	3.37	3.97	3.70*	2.17	3.53	2.90*
dagger	2.29	1.57	1.95*	4.33	4.00	4.15	1.79	1.56	1.66
engine	3.38	2.75	3.08*	4.41	4.29	4.34	3.34	2.38	2.83*
execution	1.56	1.75	1.65	4.19	3.49	3.79*	2.31	1.97	2.13
fragrance	4.16	4.51	4.33	3.04	3.62	3.36	2.34	3.50	2.97*
girl	4.52	4.17	4.36*	4.73	4.59	4.65	4.59	4.09	4.32
green	3.67	3.43	3.55	4.48	4.66	4.58	4.21	3.50	3.82*
handsome	4.26	4.49	4.37	3.37	4.34	3.92*	3.21	3.76	3.52
hurt	1.95	1.72	1.84	2.80	3.74	3.35*	3.28	4.26	3.81*
jealousy	2.07	1.53	1.82*	2.62	3.37	3.05*	2.86	3.53	3.23*
lake	4.16	4.25	4.21	4.70	4.77	4.74	3.52	3.00	3.24*
love	4.69	4.97	4.83*	3.89	4.29	4.11	4.41	4.56	4.49
maggot	1.44	1.39	1.41	3.96	3.69	3.80	2.14	1.47	1.78*
nazi	1.47	1.33	1.41	3.78	3.79	3.78	2.21	1.55	1.85*
pervert	1.92	1.36	1.65*	3.37	3.46	3.42	2.79	2.88	2.84
rattle	3.02	2.51	2.79	3.73	4.57	4.21*	2.31	2.12	2.21
romantic	4.57	4.89	4.72*	3.44	3.89	3.69	3.34	4.12	3.76*
sapphire	4.34	4.47	4.41	3.37	4.41	3.95*	1.54	1.74	1.65
sentimental	4.27	4.78	4.51*	2.63	3.23	2.97*	2.62	3.09	2.87
slaughter	1.47	1.31	1.39	3.59	3.40	3.48	2.14	1.53	1.81*
stress	1.90	1.81	1.86	2.74	3.03	2.90	3.17	4.09	3.67*
thoughtful	4.03	4.64	4.32*	2.52	3.06	2.82	2.79	3.71	3.29*
ugly	1.56	1.67	1.61	3.30	4.11	3.76*	3.72	3.97	3.86
vomit	1.44	1.11	1.28*	4.15	4.56	4.38	3.00	3.03	3.02
wedding	4.08	4.72	4.39*	4.54	4.66	4.61	2.72	3.26	3.02
whore	2.00	1.19	1.62*	3.89	4.11	4.02	3.21	2.65	2.90
woman	4.47	4.00	4.24*	4.74	4.63	4.68	4.31	4.50	4.41

*Indicates a difference between male and female mean ratings at $p < .05$.