Bulletin of the Psychonomic Society
1981, 17(2), 85-88

Self-reference in facial recognition

JOHN H. MUELLER, MICHAEL R. COURTOIS, and KAREN L. BAILIS
University of Missouri, Columbia, Missouri 65211

Subjects made decisions about facial photographs and were tested later for recognition memory of the pictures. The study decisions involved judgments about abstract personality traits (e.g., friendliness) or physical features (e.g., lip thickness) relative to either self-comparisons or some nonself standard. The expected abstract-physical feature difference emerged, but there was no evidence for a self-other difference for either type of feature. A molar self-reference task specifying no particular attribute produced good performance, but no better than the abstract-nonself task or intent-to-learn instructions. For the molar-self task, faces judged to be similar to one's own were easier to recognize than were dissimilar faces. The self-comparison task yields good retention with face stimuli, but apparently no better than other tasks that require examining many features during encoding.

Several recent studies have shown a positive effect on verbal memory when the subject's self-concept is involved during study (e.g., Bower & Gilligan, 1979; Keenan & Baillet, 1980; Rogers, Kuiper, & Kirker, 1977). To induce self-reference, subjects are required to make decisions about the study items that entail self-comparisons, such as "Does this word describe you?" Two outcomes are of interest here. First, it has been found that self-reference processing yields retention levels at least as high as following semantic processing (e.g., "Does this word mean the same as X?"), and self-reference often leads to better retention than does semantic processing (e.g., Rogers et al., 1977). Second, it also appears that items judged to be self-descriptive are recalled better than items judged nondescriptive, but this result is not always robust (Kuiper & Rogers, 1979).

The present experiments examined self-comparisons in face memory, to test the generality of the effect. Compared with the research on verbal memory, the findings to date using face stimuli have been somewhat disappointing in terms of the superiority of self-processing effects on retention. Mueller, Bailis, and Goldstein (1979) manipulated self-reference with regard to a particular feature, either some abstract personality trait, such as friendliness, or some physical feature of the person, such as weight. For facial photographs, subjects made one of four types of study decisions: abstract-self ("Does this person look more intelligent than you?"), abstract-nonself ("Does this person look intelligent?"), physical-self ("Does this person weigh more than you?"), physical-nonself ("Does this person weigh more than 150 lb?"). The usual (Bower & Karlin, 1974) superiority of abstract over physical feature processing was observed, but there was no evidence for a benefit due to self-reference, nor any interaction between type of feature and self-reference. That is, feature "depth" (Craik, 1977) seemed more potent than self-reference. Bailis and Mueller (in press) reported similar null effects for self-reference, examining only abstract features. Courtois and Mueller (1979) considered self-reference on a molar level, without regard to any specific feature ("Does this person look like you?"), and found global self-reference processing to be as good as abstract-nonself processing, but not superior to it.

The two studies reported here combined molar self-reference with self-reference tasks that focus on specific attributes of the self. The studies were similar in methodology and results, so they will be presented together. Experiment 1 was conducted to replicate the Mueller et al. (1979) study, using a different nonself comparison. Mueller et al. (1979) used an absolute value (e.g., 150 lb), whereas Experiment 1 here used "average" as the standard. Experiment 1 also changed to a physical feature that was entirely apparent in the face (e.g., lip thickness) rather than a physical feature that actually is reflected in the whole body (e.g., weight), because some research has indicated differences for the two types of physical features (Mueller, Carlomusto, & Goldstein, 1978). Experiment 2 changed the nonself comparison from such neutral terms as "average" to a personalistic basis by defining nonself as some specific other person. This self-nonself comparison thus involves actual people in each case, whereas the self-nonself comparison in previous studies of faces has in fact been a person-nonperson comparison.

Reprint requests should be sent to J. H. Mueller, Psychology Department, University of Missouri, Columbia, Missouri 65211. The authors would like to acknowledge helpful comments on an earlier draft by Richard Petty and Donald Kausler. This research was supported by funds from the Research Council of the Graduate School of the University of Missouri.

METHOD

Experiment 1

Subjects and Design. The conditions were defined by type of orienting task during study. There were five incidental learning

0090-5054/81/020085-04$00.65/0

groups: abstract-self, face-self, abstract-nonself, face-nonself, and molar-self, plus an intentional learning control group. The design can be summarized as a 2 by 2 factorial, for task "level" (abstract, physical) by criterion (self, average), with the molar-self and intentional groups outside the factorial. Study decision outcome (i.e., similar or dissimilar to self or nonself) was added to all analyses (except for false alarms, for which it is not meaningful). Hit rates were computed separately for these "descriptive" and "nondescriptive" subsets, thus adjusting for the differential number of "yes" and "no" responses.

The subjects were recruited from introductory psychology courses; there were 15 subjects in each group, with men and women randomly distributed.

Materials. The 80 black-and-white slides were made from pictures in college yearbooks. Half were of men and half of women; all were Caucasians. Portraits with distinctive features (e.g., glasses) were not used. These were divided into two sets of 20 males and 20 females each, with sets counterbalanced as study slides and distractors.

Procedure. The 40 study slides were presented one at a time at a 5-sec rate, under the guise of a rating study. There were two versions of the abstract trait judgment, friendliness and intelligence, and two versions of the facial feature task, lip thickness and distance between the eyes. Each version was used equally often; no differences were observed for this factor, so it will not be considered further. Subjects with the nonself incidental learning tasks marked their answer sheets (yes, no) according to whether or not the person shown was "above average" on the feature in question. These groups will be referred to as the abstract-nonself and face-nonself conditions, because no self-comparison was involved. In the self-comparison tasks, subjects decided whether the person shown had more or less of the feature in question "than you." These groups will be referred to as the abstract-self and face-self conditions. The subjects in the molar self-reference group responded in terms of whether the person "looks like you." Each subject made only one decision throughout.

The unannounced test followed immediately. Forty new slides were randomly mixed with the 40 study slides, and the test sequence was shown at a 5-sec rate. A single-item test format was used; subjects marked an answer sheet in terms of whether each slide had been shown in the study phase. There were no constraints on guessing or the number of "old" responses allowed.

The procedure for the intentional learning group was the same, except they were told before study that a test would follow and they performed no orienting task during study. This group was included to determine how beneficial self-processing is compared with the strategies subjects spontaneously employ.

Experiment 2

The main procedural difference was that faces were rated on a 5-point scale during study; for analysis, these were converted by using faces rated 1-3 as nondescriptive and 4-5 as descriptive. (This pooling was done to obtain sufficient nonzero entries.) There were five groups: abstract-self, facial-self, molar-self, abstract-nonself, and facial-nonself. The first three were identical to those in Experiment 1. The last two required the subject to make judgments of abstract traits or facial features relative to their favorite television or movie star, instead of "average." The design thus was a 2 by 2 factorial, for task (abstract, physical) by self-reference (self, other), with the molar-self group outside the factorial. Decision outcome was added to the design for all analyses except false alarms. There were 16 subjects per group.[1]

RESULTS

Experiment 1

When analyzed as a one-way (five groups) design, the results showed a group main effect for false alarm rate [$F(4,70) = 5.28$, MSe = .012].[2] The molar self-reference group had a mean false alarm rate of .18, compared with .17 for abstract-self, .19 for abstract-nonself, .28 for face-nonself, and .30 for the face-self groups. For comparison, the intentional learning group had a mean false alarm rate of .22. Thus, relative to the subject-generated processing evoked by "learn" instructions, abstract-nonself and self-oriented processing led to slightly better rejection of new items, and physical feature processing led to somewhat more confusion of new items (with or without self-reference).

The hit rate also revealed a significant group main effect [$F(4,70) = 3.81$, MSE = .047], with mean rates (pooling over choice) as follows: abstract-self = .79, abstract-nonself = .75, molar-self = .71, face-nonself = .65, and face-self = .59. (For comparison, the intentional group had a mean hit rate of .73, thus ranking the same for old- as for new-item recognition.) Analyzed as a 2 by 2 factorial (dropping the molar-self group), there was no choice main effect or Task by Choice interaction for hit rate ($Fs < 1.91$).

The results of the signal detection analyses are shown in the top of Figure 1-A. The d' analysis revealed a significant group main effect [$F(4,70) = 8.39$, MSe = 1.321], with mean values as follows (pooling over study choice): abstract-self = 1.84, abstract-nonself = 1.81, molar-self = 1.58, face-nonself = 1.01, and face-self = .78. As for hits and false alarms, the dominant factor in the rank

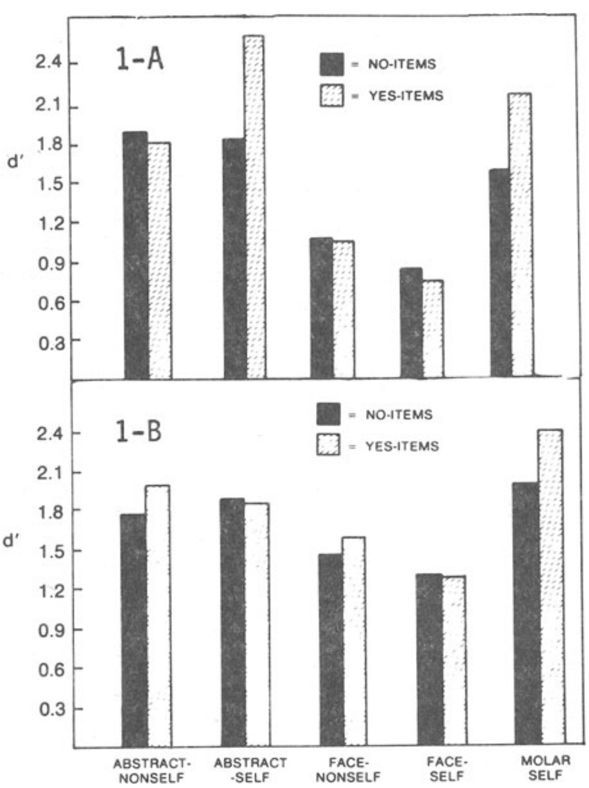

Figure 1. Recognition performance for each condition in Experiment 1 (A) and Experiment 2 (B), by judged similarity (yes, no) to the self or other criterion.

ordering was feature "depth," not self-referencing. (The intentional group produced an overall mean d′ of 1.60; in this case, of course, the study-choice breakdown is not feasible.) In a 2 by 2 analysis, the choice main effect [F(1,70) = 3.06, MSe = .727, p < .09] and the Task by Choice interaction [F(4,70) = 2.04, MSe = .727, p < .10] were only marginally significant.

Experiment 2

The average false alarm rates were .24 for the facial-self condition, .18 for the facial-nonself condition, .17 for the abstract-nonself condition, .11 for the abstract-self condition, and .14 for the molar-self condition. When analyzed in a one-way (five groups) design, the groups main effect was significant [F(4,75) = 3.57, MSe = .777]. When analyzed in the 2 by 2 design, dropping the molar-self group, the task main effect was significant [F(1,60) = 8.06, MSe = .589], as was the Task by Self-Reference interaction [F(1,60) = 5.45]; the self-reference main effect was not significant (F < 1).

Clearly, self-reference reduced false alarms when an abstract trait was involved in the decision, but self-reference actually increased false alarms when a physical feature had been examined. This pattern of results was in the same direction in Experiment 1, but well short of significance. This effect of self-processing on false alarms should not be confused with the "false alarms effect" reported by Rogers, Rogers, and Kuiper (1979). They found more false alarms as the item was more similar to the self, as expected given a prototype manipulation. The new items here were not structured for similarity to the self. Such inadvertent similarity may be one reason a subject would misidentify a distractor as "old," at least following self-reference in study, but the group differences here seem dependent on the type of feature processed, implying that a different interpretation may be required, at least for abstract features.

The hit rates were higher for the abstract-self, abstract-nonself, and molar-self groups (means = .72, .74, and .78, respectively) than for the facial-self and facial-nonself groups (means = .70 and .66), but no effects were significant (Fs < 2.41).

The analysis of d′ revealed a significant main effect for feature level [F(1,60) = 10.95, MSe = .652]. As can be seen in the bottom of Figure 1-B, the abstract conditions performed better than did the physical feature groups (means = 1.88 and 1.41, respectively), but there was no interaction with self-reference (F < 1) or decision outcome (F < 1). Self-reference was not significant as a main effect (F < 1), nor were there any interactions involving self-reference in the factorial groups. The molar-self condition performed nonsignificantly better (mean = 2.01) than did the abstract groups, with a slightly more pronounced study choice difference.

DISCUSSION

Both experiments replicated the basic abstract vs. facial feature task difference. However, these studies do not support a strong effect for self-processing, which would require that self-reference be superior to other types of processing, although it may be that self-comparisons are equivalent to other beneficial study activities. When induced for specific features, self-reference seems to produce no effect for either abstract or physical features. When self-comparisons are more open-ended, the result is good retention, but still no better than other tasks that involve processing several features (such as friendliness judgments or "intent" to learn). We have also observed (Mueller & Courtois, Note 1) that the molar-self case produces performance virtually identical to judgments about one's parents or best friends, other cases in which multiple features can be examined. For molar-self comparisons, faces judged to be similar to one's own are remembered better than are dissimilar faces, a result more in line with self-reference data on verbal memory. However, this result was not consistent when self-comparisons referred to a single feature (see Figure 1).

Some reasons for the absence of a stronger self-comparison effect can be considered briefly. First, it is possible that a functional ceiling effect prevents self-reference from yielding further gains. Face memory is generally good, which lends some credibility to this idea. The ceiling effect might be eliminated by using a longer retention interval. However, some features of the data argue against a ceiling effect; for example, the molar-self and intentional groups were comparable, but the abstract feature groups did better yet, indicating performance could be improved even here.

A second possibility is a recall-recognition difference: The verbal memory studies have generally used a recall test, whereas the face studies have used recognition. One way to interpret this is that the self-reference benefit is greater when retrieval is involved in the test.

It should also be acknowledged that the actual decisions are somewhat different for faces and words. The verbal memory studies compare the self task to a "semantic" task, such as synonymity, in which the latter involves a well-defined (single?) dimension. The face studies compare the self judgment to an "abstract" task, such as friendliness. Both the self and abstract tasks may require examination of numerous features, thus leading to equivalent performance. This discrepancy may be part of the problem, but it seems unlikely to be the only one.

Finally, it may be that the answer does not lie with a methodological difference between studies of face memory and verbal memory, but with some more fundamental problem. Lord (1980) has discussed a distinction between the self-schema, as a verbal feature list perhaps, and the self-image, noting that the latter may actually be less effective as a memory aid. It seems likely that the image of one's self is involved when the self-comparison involves a photograph; thus it may be that the self-reference manipulation cannot be expected to be as effective with face stimuli as it is for verbal stimuli.

REFERENCE NOTE

1. Mueller, J. H., & Courtois, M. R. *Self-other comparisons in face memory.* Unpublished manuscript, 1980.

REFERENCES

Bailis, K. L., & Mueller, J. H. Anxiety, feedback, and self-reference in face recognition. *Motivation and Emotion,* in press.

Bower, G. H., & Gilligan, S. G. Remembering information related to one's self. *Journal of Research in Personality,* 1979, **13,** 420-432.

Bower, G. H., & Karlin, M. B. Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology,* 1974, **103,** 751-757.

Courtois, M. R., & Mueller, J. H. Processing multiple physical features in facial recognition. *Bulletin of the Psychonomic Society,* 1979, **14,** 74-76.

Craik, F. I. M. Depth of processing in recall and recognition. In S. Dornic (Ed.), *Attention and performance* (Vol. 6). Hillsdale, N.J: Erlbaum, 1977.

Keenan, J. M., & Baillet, S. D. Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and performance* (Vol. 8). Hillsdale, N.J: Erlbaum, 1980.

Kuiper, N. A., & Rogers, T. B. Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology,* 1979, **37**, 499-514.

Lord, C. G. Schemas and images as memory aids: Two modes of processing social information. *Journal of Personality and Social Psychology,* 1980, **38**, 257-269.

Mueller, J. H., Bailis, K. L., & Goldstein, A. G. Depth of processing and anxiety in facial recognition. *British Journal of Psychology,* 1979, **70**, 511-515.

Mueller, J. H., Carlomusto, M., & Goldstein, A. G. Orienting task and study time in facial recognition. *Bulletin of the Psychonomic Society,* 1978, **11**, 313-316.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology,* 1977, **35**, 677-688.

Rogers, T. B., Rogers, P. J., & Kuiper, N. A. Evidence for the self as a cognitive prototype: The "false alarms effect." *Personality and Social Psychology Bulletin,* 1979, **5**, 53-56.

NOTES

1. Measures of imagery ability and situational anxiety were taken after both experiments, but because these revealed no significant correlations, they will not be discussed further.

2. Effects described as significant involve $p < .05$, or better. Effects not discussed failed to reach this level.

(Received for publication January 30, 1981.)