Effects of spectacles on recognition memory for faces: Evidence from a distractor-free test

STUART J. McKELVIE
Bishop's University, Lennoxville, Quebec, Canada

Twenty subjects were administered a distractor-free recognition memory test for 20 photographs of faces (10 with and 10 without spectacles) that they had seen 1 week earlier. Hit rates were less than 1 and were similar for the two kinds of faces. Since the same results were obtained for hits on a conventional test, it was argued that the distractor-free test is a viable method for measuring recognition memory. Implications for the practice of routinely correcting recognition memory hits were also discussed.

When comparing visual recognition memory for different kinds of material, researchers usually employ a test in which the old (seen) stimuli are mixed with new (unseen) ones. This procedure has been adopted to keep subjects honest—that is, to prevent them from simply claiming that all stimuli are familiar (Wallace, 1980). Unfortunately, results cannot be clearly interpreted, since performance is affected by distractor similarity, which may vary across materials in unknown ways because it is difficult to quantify (Faw, 1992; Wallace, 1980).

However, Wallace (1980) points out that distractors change the task from "re-cognition" (experiencing familiarity) to recognition discrimination, and he challenges the assumption that subjects would invariably respond positively if they were tested only with old stimuli. In fact, three studies have shown that hit rates (saying "yes" to an old stimulus) for distractor-free tests were lower than 1 and did not differ from those with conventional tests (Faw, 1990; Ley & Long, 1987; Wallace, 1982). Furthermore, in two cases, independent variables (word frequency, Ley & Long, 1987; list structure, Wallace, 1982) affected both sets of hit rates in the same manner.

In Faw's (1990) experiment, recognition memory performance (total number correct) was generally higher for faces than for names on the conventional test but was lower for faces on the distractor-free test. Faw argues that the name decrement may be due to greater interitem similarity among names than among faces and that the distractor-free test is more equitable. However, it is also possible that subjects may have adopted a more lenient response criterion with names than with faces. On the conventional test, this would produce a higher number of false alarms (saying "yes" to a new stimulus) and a lower number of total correct responses for names than for faces,

whereas on the distractor-free test, it would yield a higher number of hits for names.

Faw did not report false alarms, but Ley and Long (1987) found that they were higher for medium- and low-frequency words than for high-frequency words. Similarly, Wallace (1982) reported that false alarms, but not hits, were higher for lists with high structure than with low structure. Under such circumstances, investigators usually assume that some of the obtained hits are contaminated by response bias, and they correct the hit rate by subtracting the false-alarm rate (Woodworth, 1938) or converting it to a measure of sensitivity (d', Green & Swets, 1966; or A', Pollack & Norman, 1964). In contrast, since the effect of list structure on hits was also not significant on a distractor-free test, Wallace states that its effect on false alarms has no implications for hits on the conventional test. This argument assumes that distractors *induce* response bias, but only on new stimuli. However, it is also possible that both sets of hits were contaminated in the same manner by a positive response bias. This implies that the lack of effect of list structure on hits in both cases conceals a poorer memory performance for high than for low list structure. By the same token, reliance on Ley and Long's distractor-free hit data may conceal a positive effect of word frequency on recognition memory.

Although Wallace (1980) acknowledges that distractors are important when signal detection theory is applied to recognition memory, neither he nor the other authors emphasize that the distractors' function in this context is to detect response bias under conditions of uncertainty (Green & Swets, 1966). According to this approach, test responses reflect both memory strength and a response criterion. When subjects feel unsure, they may respond positively on a strategic basis. Using this technique, I found that hit rates were similar for faces with than without spectacles but that, since false alarms were higher with spectacles, immediate recognition memory strength (d', corrected hits) was poorer for bespectacled faces (McKelvie, 1987, 1988, 1991). However, after a 1-week delay, results were less consistent. In one case, hits, false alarms,

Table 1
Mean Accuracy and Confidence Scores in Each Condition

Recognition Test	n	Spectacles		No Spectacles	
		M	SD	M	SD
Hits					
Conventional	24	7.92	0.97	7.42	1.24
Distractor-free	20	6.60	1.98	6.45	1.59
Confidence					
Conventional	24	3.05	0.46	3.15	0.50
Distractor-free	20	2.99	0.55	3.06	0.68
Joint Accuracy/Confidence					
Conventional	24	6.31	0.55	6.06	0.74
Distractor-free	20	5.54	1.10	5.48	0.93

Note—Maximum scores were 10 (hits), 4 (confidence), and 8 (joint). Data for the conventional test were obtained from McKelvie's (1993a) Experiment 2.

and memory strength did not differ on the two kinds of faces (McKelvie, 1988). In another case, false alarms were higher for faces with than without spectacles, but hits and corrected hits did not differ (McKelvie, 1993a).

As a model for this task, assume that subjects encode each original face according to its distinctive feature or features (Winograd, 1981). For faces with spectacles, it is likely that this feature will be the eyes, including the spectacles (McKelvie, 1993b). If interitem similarity is higher for these faces than for those without eyeglasses, true recognition accuracy and subjective confidence may be lower for faces with spectacles. However, when subjects are unsure, memory for the distinctive feature of eyeglasses may induce them to respond positively on both new and old bespectacled faces, leading to more false alarms and to a similar number of obtained hits on faces with than without spectacles.

The purpose of the present experiment was to investigate delayed recognition memory for faces with and without spectacles using the distractor-free test, to compare the results with previous performance on the conventional test, and, as assumed by the model, to ascertain if response bias operates, particularly on old faces. The final goal was accomplished by analyzing confidence ratings. Although obtained hits may not differ for faces with and without spectacles, subjects' confidence might be lower if some of the hits represent a positive response bias.

Finally, if the presence of spectacles increases task difficulty, it can also be predicted that the relationship between confidence and accuracy will be lower for faces with than without spectacles (which is known as the optimality hypothesis; see Deffenbacher, 1980). This result has been found with faces (McKelvie, 1990), particularly

when task difficulty was manipulated with eyeglasses (McKelvie, 1993a). However, accuracy in these cases was measured by corrected hits. For present purposes, the confidence/accuracy relationship was computed from hits on the previous conventional test and compared with results obtained from the distractor-free test administered here.

METHOD

Subjects
Participating were 20 (6 male, 14 female) high school students from 14 to 15 years of age. All volunteered as part of a classroom demonstration.

Materials and Procedure
The subjects viewed 20 black-and-white photographic slides of students (5 men and 5 women with and without eyeglasses; see McKelvie, 1987, 1988, 1991, 1993a, 1993b) for 3 sec each. They inspected the faces carefully, anticipating a memory test after 7 days. Just before this test, they were told that they would be shown faces, some of which they had previously seen, and that they should classify each one as either old or new and then rate their confidence in this decision on a 4-point scale (4 = certain, 3 = reasonably confident, 2 = somewhat confident, and 1 = guessing). In fact, the original 20 photographs were shown in the same order as before.

RESULTS

Table 1 shows the mean scores on accuracy and confidence for the present experiment and for McKelvie's (1993a) Experiment 2, which was conducted with a conventional recognition memory test. To measure Wallace's (1980) feeling of familiarity, a joint accuracy/confidence score was also constructed using a scale ranging from 8 for a certain hit to 1 for a certain miss (saying new to an old face).

Each score was treated with a 2 × 2 (test × spectacles) mixed-model analysis of variance (ANOVA). For both hits and joint scores, performance was higher on the conventional than on the distractor-free test [Fs (1,42) = 12.55, 10.67, p < .01, respectively]. However, neither the effect of spectacles nor the test × spectacles interactions were significant (Fs < 1.20). There was a weak trend toward lower confidence on faces with than without spectacles [$F(1,42) = 2.64, p = .11$], but the other confidence effects did not approach significance (Fs < 1.10).

To investigate the relationship between accuracy and confidence, mean confidence was compared for correct (hits) and incorrect (misses) responses (see McKelvie, 1990, 1993a). Here, a 2 × 2 × 2 (test × spectacles × accuracy) mixed-model ANOVA showed that the subjects were more confident when correct than when incorrect [$F(1,39) = 41.62, p$ < .01; see Table 2]. Again, the ef-

Table 2
Mean Confidence When Correct and Incorrect

Recognition Test	n	Spectacles				No Spectacles			
		Correct		Incorrect		Correct		Incorrect	
		M	SD	M	SD	M	SD	M	SD
Conventional	22	3.28	0.45	2.39	0.87	3.34	0.47	2.51	0.88
Distractor-free	19	3.14	0.64	2.65	0.68	3.23	0.69	2.79	0.73

Note—Sample sizes were slightly smaller than in Table 1, because some subjects scored perfectly in some conditions.

fect of spectacles was marginally significant [$F(1,39)$ = $2.72, p$ = .10]. The test × accuracy interaction also approached significance [$F(1,39)$ = $3.62, p$ < .07], indicating that the difference in confidence on hits and misses was greater for the conventional than for the distractor-free test. Neither the spectacles × accuracy nor the test × spectacles × accuracy interactions were significant (Fs < 1).

DISCUSSION

On the distractor-free test, hit rates were less than 1, confirming previous demonstrations that subjects did not give uniformly positive responses when they viewed only old stimuli (Faw, 1990; Ley & Long, 1987; Wallace, 1980, 1982). Indeed, hits were lower on the distractor-free test than on the previous conventional test. This effect may have occurred because the present subjects were approximately 4 years younger than those in McKelvie's (1993a) Experiment 2, which would indicate that the difference in memory is real and would provide validating evidence for the distractor-free test.

On both tests, hits did not differ significantly for faces with and without spectacles, showing that another independent variable had a similar effect on conventional and distractor-free hits (see Wallace, 1982; Ley & Long, 1987). However, contrary to the hypothesis that some hits on spectacles may have been biased positive responses generated under uncertainty, confidence was not clearly lower on faces with than without spectacles. There was a trend in this direction, but it did not translate into a lower joint (feeling of familiarity) score for bespectacled faces. Furthermore, although the optimality hypothesis received some support from the greater confidence difference between correct and incorrect responses when the hit rate was higher (conventional test) than when it was lower (distractor-free test), there was no evidence that it applied to faces with and without spectacles. This indicates that the latter faces were not perceived as more difficult to recognize and implies that the similar hit rates reflect similar memory accuracy. Future research might extend the present application of the optimality hypothesis by examining whether it holds for both conventional and distractor-free hits when they are clearly reduced by an independent variable (e.g., exposure time and inversion; see McKelvie, 1990). If it does, it would provide further evidence that the distractor-free test is valid.

Together, the results of this experiment add further support to the proposal that a distractor-free recognition test is a viable alternative to the conventional one. However, they also raise some questions about the assumption that recognition memory hit rates should be modified on the basis of false-alarm rates. Although the application of the signal detection measure of sensitivity (d') to memory has been questioned (e.g., see Faw, 1992; Kolers, Duchnicky, & Sundstroem, 1985; Pollack & Norman, 1964), the usual nonparametric substitute (A'; Pollack & Norman, 1964) also assumes that responses to old and new items have a common basis, an assumption that has not received empirical support (Kolers et al., 1985). The present analysis also challenges this assumption, since there was no evidence that some of the hits for bespectacled faces on either the conventional or distractor-free tests were biased positive responses. Since false-alarm rates have been consistently higher for faces with than without spectacles (McKelvie, 1987, 1988, 1991, 1993a), it is likely that response bias operated only on new faces (see also Wallace, 1992), which implies that hits should not routinely be altered to allow for false positives. Notably, Shapiro and Penrod (1986) recommend that hits and false alarms be treated separately, and some investigators have done so (e.g., Klatzky, Martin, & Kane, 1982). Applied to the effect of spectacles on facial memory, this argument

challenges the conclusion (McKelvie, 1987, 1988, 1990) that immediate recognition memory is poorer for faces with than without spectacles. For true recognition, or the feeling that a stimulus is familiar, the absence of a general effect of spectacles on hits on both conventional (McKelvie, 1987, 1988, 1990, 1993a) and distractor-free (the present experiment) tests indicates that memory is similar for both kinds of faces.

Of course, there may also be cases in which response bias contaminates both hits and false alarms. To investigate this possibility, conventional and distractor-free tests should be compared when the response criterion is strongly manipulated. It is predicted that increased hits would be accompanied by decreased confidence and that the optimality hypothesis would apply.

REFERENCES

DEFFENBACHER, K. A. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law & Human Behavior*, **4**, 243-260.

FAW, H. F. (1990). Memory for names and faces: A fair comparison. *American Journal of Psychology*, **103**, 317-326.

FAW, H. F. (1992). Recognition of unfamiliar faces: Procedural and methodological considerations. *British Journal of Psychology*, **83**, 25-37.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

KLATZKY, R. L., MARTIN, G. L., & KANE, R. A. (1982). Semantic interpretation effects on memory for faces. *Memory & Cognition*, **10**, 195-206.

KOLERS, P. A., DUCHNICKY, R. L., & SUNDSTROEM, G. (1985). Size in the visual processing of faces and words. *Journal of Experimental Psychology*, **11**, 726-751.

LEY, R., & LONG, K. (1987). A distractor-free test of recognition and false recognition. *Bulletin of the Psychonomic Society*, **25**, 411-414.

McKELVIE, S. J. (1987). Recognition memory for faces with and without spectacles. *Perceptual & Motor Skills*, **65**, 705-706.

McKELVIE, S. J. (1988). The role of spectacles in facial memory: A replication and extension. *Perceptual & Motor Skills*, **66**, 651-658.

McKELVIE, S. J. (1990). Effects of exposure time and inversion on the confidence-accuracy relationship in facial memory: A test of the optimality hypothesis. *Perceptual & Motor Skills*, **71**, 32-34.

McKELVIE, S. J. (1991). Effects of processing and transformation on recognition memory for photographs of faces. *Bulletin of the Psychonomic Society*, **29**, 98-100.

McKELVIE, S. J. (1993a). Confidence and accuracy in facial memory: Further evidence for the optimality hypothesis. *Perceptual & Motor Skills*, **76**, 1257-1258.

McKELVIE, S. J. (1993b). *Perceived attributes of faces with and without spectacles*. Unpublished manuscript.

POLLACK, I., & NORMAN, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, **1**, 125-126.

SHAPIRO, P. N., & PENROD, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, **100**, 139-156.

WALLACE, W. P. (1980). On the use of distractors for testing recognition memory. *Psychological Bulletin*, **88**, 696-704.

WALLACE, W. P. (1982). Distractor-free tests of memory. *American Journal of Psychology*, **95**, 421-440.

WINOGRAD, E. (1981). Elaborativeness and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning & Memory*, **7**, 181-190.

WOODWORTH, R. S. (1938). *Experimental psychology*. New York: Holt.