11-6-1992

The influence of processing instructions at encoding and retrieval on face recognition accuracy

Garrett L. Berman

Florida International University

ABSTRACT OF THE THESIS

The Influence of Processing Instructions at Encoding

and Retrieval On Face Recognition Accuracy

by

Garrett L. Berman

Florida International University, 1992

Miami, Florida

Professor Brian L. Cutler, Major Professor

   Whereas previous research has demonstrated that trait
ratings of faces at encoding leads to enhanced recognition
accuracy as compared to feature ratings, this set of
experiments examines whether ratings given after encoding and
just prior to recognition influence face recognition accuracy.
In Experiment 1 subjects who made feature ratings just prior
to recognition were significantly less accurate than subjects
who made no ratings or trait ratings. In Experiment 2 ratings
were manipulated at both encoding and retrieval. The
retrieval effect was smaller and nonsignificant, but a
combined probability analysis showed that it was significant
when results from both experiments are considered jointly.
In a third experiment exposure duration at retrieval, a
potentially confounding factor in Experiments 1 and 2, had a
nonsignificant effect on recognition accuracy, suggesting that
it probably does not explain the results from Experiments 1
and 2. These experiments demonstrate that face recognition

accuracy can be influenced by processing instructions at
retrieval.

To Professors: Brian Cutler
                 Ronald Fisher
                 Paul Foos

This thesis having been approved in respect to form and
mechanical execution, is referred to you for judgement
upon its substantial merit.

_____
Arthur Herriott, Dean
College of Arts and Sciences

The thesis of Garrett Lee Berman is approved.

_____
Paul W. Foos

_____
Ronald P. Fisher

_____
Brian L. Cutler, Major Professor

Date of Examination: November 6, 1992

_____
Richard Campbell, Dean
Division of Graduate Studies

FLORIDA INTERNATIONAL UNIVERSITY, 1992

ii

ACKNOWLEDGEMENTS

I would like to thank the following people for their time, patience, and encouragement: Dr. Brian L. Cutler, Dr. Ronald Fisher, Dr. Paul Foos. Without their support, this thesis would not have been possible.

I am also indebted to Janie Wagner who assisted with data collection. I would also like to thank Rod Lindsay, Mike Toglia, and David Ross for their valuable suggestions for improving this manuscript.

TABLE OF CONTENTS

List of Tables

List of Figures

The Influence of Processing Instructions at Encoding

and Retrieval On Face Recognition Accuracy

Attempts to understand the psychological processes underlying face recognition focus primarily on the encoding stage of memory (Light, Kayra-Stuart, & Hollander, 1979; Patterson & Baddeley, 1977; Strnad & Mueller, 1977; Warrington & Ackroyd, 1975; Wells & Hryciw, 1984; Winograd, 1976, 1981; Sporer, 1991). Authors of these studies assume that processing strategies can be manipulated via instructions given to subjects prior to encoding to-be-recognized faces. Specific processing hypotheses are tested by manipulating the instructions and examining their effects on recognition accuracy. In general this research has successfully demonstrated that instructions can affect processing strategies and hence recognition accuracy.

Although this research provides insights into the processes underlying face recognition, it provides little practical information about how face recognition accuracy might be improved in an ecologically valid setting, such as eyewitness identification of a crime perpetrator. Police investigators have little control over the conditions surrounding a crime (Wells, 1978), such as the processes used by a witness to encode a perpetrator's facial characteristics. In contrast, police have much more control over the conditions surrounding the retrieval of crime information (Wells, 1978), such as instructions given to witnesses prior to an identification

1

test. Thus, researchers with practical interests in improving recognition accuracy typically manipulate factors at the retrieval, rather than the encoding stage of memory (e.g., Cutler & Penrod, 1988; Fisher, Geiselman & Amador, 1989; Linsday, Lea, & Fulford, 1991; Luus & Wells, 1991). The goal of the current research is to test whether processing instructions, which have been shown to affect face recognition accuracy when manipulated at the encoding stage, can be used to improve face recognition accuracy when manipulated at the retrieval stage.

Processing Instructions at Encoding

Craik and Lockhart's (1972) levels of processing framework provided the conceptual foundation for an early test of processing strategies. Bower and Karlin (1974) found that subjects instructed to make honesty or likableness judgments about to-be-recognized faces, as compared with subjects who made gender judgments, performed better on a subsequent face recognition test. Bower and Karlin (1974) reasoned that judging a face for its honesty and/or likableness leads to more elaborate processing than does rendering gender judgments, which is thought to produce a shallower memory trace. Winograd's (1976) subjects rated faces for one of the following during the encoding phase: personality traits, occupations, hair types, or nose size. Subjects who made judgments about traits or occupations performed more accurately on the subsequent face recognition test than did

2

subjects who made more superficial judgments about hair type and nose size. Additional tests continued to confirm the superiority of trait judgments over feature judgments for face recognition accuracy (Baddeley, 1979; Patterson & Baddeley, 1977; Smith & Winograd, 1978; Strnad & Mueller, 1977; Warrington & Ackroyd, 1975; Wells & Hryciw, 1984; Winograd, 1981).

Several researchers (Winograd, 1981; Wells & Hryciw, 1984), however, proposed alternative explanations for the superior effects of trait judgments over feature judgments. Winograd's (1981) "elaboration hypothesis" proposed a quantitative relation between encoding processes and recognition accuracy. Trait judgments result in superior recognition accuracy because they lead to a higher number of encoded features than do judgments that require the subject to isolate individual facial features. If so, argued Wells and Turtle (1988), then subjects who made trait judgments should produce more accurate descriptions of faces than subjects who made facial feature judgments. In contrast, Wells and Turtle demonstrated that subjects who made featural judgments about to-be-recognized faces were better describers than were subjects who made trait judgments. This finding contradicts Winograd's elaboration hypothesis. Sporer (1991) also examined Winograd's (1981) elaboration hypothesis and found contradictory evidence. Sporer (1991) manipulated the number of physical features subjects were instructed to scan during encoding operations.

Number of scanned features did not significantly influence recognition accuracy.

Wells and Hryciw (1984) argued that the superior recognition accuracy produced by trait judgments over featural judgments is not due to depth of processing nor to number of features encoded. They argued that the two types of instructions produced qualitatively different memory traces. Due to these qualitatively diverse traces, they argued that retrieval accuracy is driven by the match between encoding strategy and retrieval strategy. Wells and Hryciw (1984) theorized that trait judgments lead to holistic processing, which they specifically define as more attention to interfeatural information (e.g., distances between features) and less attention to intrafeatural information (e.g., size of the nose). In contrast, featural judgments lead to featural processing, defined as more attention to intrafeatural information and less attention to interfeatural information. Wells and Hryciw (1984) further argued that face recognition follows a holistic process, requiring more attention to interfeatural than intrafeatural information. Thus, the "superiority" of trait judgments over featural judgments in previous research is not due to better quality encoding; rather, it is attributable to the match between encoding and retrieval processes: Both trait judgments and recognition accuracy involve holistic processing, or greater attention to interfeatural information.

It follows, according to Wells and Hryciw (1984), that if a retrieval task involving featural processing were used, then feature judgments at encoding would produce more accurate retrieval than trait judgments at encoding. In order to test their matching hypothesis, Wells and Hryciw (1984) had subjects make trait (presumably requiring holistic processing) or featural judgments (presumably requiring featural processing) about to-be-recognized faces. Half of the subjects in each encoding condition then participated in a recognition test (presumably requiring holistic processing) and the other half participated in a facial composite production task (presumably requiring featural processing). The composite production task involved use of the Identi-Kit; facial composites were constructed by overlaying acetate transparencies containing individual facial features. Consistent with the matching hypothesis, subjects who made trait judgments at encoding performed significantly more accurately on the recognition test but significantly less accurately on the composite production test, in comparison to subjects who made featural judgments at encoding. Results of Wells and Turtle's (1988) study, mentioned above, also support the matching hypothesis. Subjects who made trait judgments at encoding were significantly more accurate at recognizing the faces but significantly less accurate at verbally describing the faces than subjects who made featural judgments at encoding. Support for Wells and Hryciw's (1984) matching

hypothesis is consistent with Tulving and Thomson's (1973) encoding-specificity hypothesis: Successful retrieval is dependent upon the similarity between cognitive processes utilized during encoding operations and those involved in subsequent retrieval stages.

The Current Research

As defined above, holistic processing refers to attention to interfeatural information, whereas featural processing refers to attention to intrafeatural information. It is important to note that these two types of processing are not mutually exclusive; rather, one can engage in either process to different degrees. Moreover, the processes are not, in theory, inversely related. If a subject is shown a face for 10 msec, he or she is not likely to process much intrafeatural or interfeatural information. If given unlimited time, a subject may process substantial amounts of both types of information. Different encoding instructions and retrieval tasks may pose constraints on how much of each type of processing a subject engages in. Given a fixed period of time for encoding, holistic instructions reduce the amount of attention paid to intrafeatural information and increase the amount of attention paid to interfeatural information. Retrieval tasks can affect attention as in Wells and Hryciw (1984). The Identi-Kit is based on recognition of individual facial features. In selecting features, interfeatural information is absent. Facial description allows somewhat

more opportunity for interfeatural processing than does the Identi-Kit, but not as much as face recognition.

My principle concern is with face recognition and with manipulations at the retrieval stage. When subjects are not given explicit processing instructions at the retrieval phase of face recognition, they apparently process holistically, at least more so than when they describe faces (Wells & Turtle, 1988) or construct facial composites (Wells & Hryciw, 1984). But because attention to intrafeatural and interfeatural information are theoretically independent, and because both types of cues are present in face recognition tasks, I hypothesized that it is possible to influence attention to intrafeatural and interfeatural information within a face recognition task. I attempted this in Experiments 1 and 2 by having subjects make personality inferences (which presumably leads processing of interfeatural information, or holistic processing) and featural judgments (which presumably leads to processing of featural information, or featural processing) after encoding and just prior to making a recognition judgment. My rationale for focussing on manipulations at the retrieval stage is due to my applied concerns: In cases involving eyewitness identification, police investigators have little control over the conditions surrounding the encoding of information and more control over retrieval tests. Thus, in comparison to the manipulation by Wells and colleagues (Wells

& Hryciw, 1984; Wells & Turtle, 1988), these manipulations are much more practical.

<center>Experiment 1</center>

<u>Overview and Hypotheses</u>

The primary question addressed in Experiment 1 is, can recognition accuracy be influenced by directing subjects' attention to intrafeatural information versus interfeatural information at retrieval? On the one hand, earlier research (Wells & Hryciw, 1984; Wells & Turtle, 1988) provides evidence that subjects attend more to interfeatural information in recognition tasks than they do in description and facial composite tasks. What researchers do not know is the extent to which they attend to interfeatural versus intrafeatural information during recognition. If subjects attend only to interfeatural information at recognition, then further attempts to direct their attention to interfeatural information should not improve recognition accuracy. In contrast, if subjects primarily, though not exclusively, attend to interfeatural information during recognition, then further attempts to direct their attention to interfeatural information might improve recognition. In any case, attempts to direct attention to intrafeatural information at recognition, because it is at the expense of interfeatural information, is likely to decrease recognition accuracy. By including a no-instruction control group, and comparing its performance with the instructed groups, I have a baseline

against which to examine whether performance is improved or decreased by the instructions.

Subjects viewed a series of 60 faces at encoding and attempted a recognition task from another 60 faces. The design of the experiment was a 2 (holistic versus featural processing at retrieval) X 2 (use of one versus three questions to induce type of processing) factorial where both variables were manipulated between subjects. An appended control group used no processing instructions at retrieval.

I hypothesized that instructions to process faces featurally at recognition would lead to inferior recognition performance as compared to instructions to process faces holistically. I have no basis for predicting the performance of the control group relative to the other conditions. The control group is included in the design for exploratory purposes. The manipulation of number of questions used to induce processing at retrieval was also included for exploratory purposes: Does asking more questions affect extent of processing (of either type) and hence recognition accuracy?

Method

Subjects. Subjects were 72 (49 female, 23 male) students from introductory psychology classes at Florida International University. Subjects were randomly assigned to one of the five experimental conditions (12 subjects in each).

Materials. The materials consisted of 90 photographs of faces obtained from a local high school yearbook. All faces were head and shoulder photographs of young Caucasian men and women. In addition, all faces were photographed in a straight-on pose. The faces were cut out of the yearbook and centered on 13.3 X 21.3 cm white index cards. No faces had distinguishing characteristics such as beards, mustaches, or glasses. The photos were divided into three sets of 30, and sets were counterbalanced as targets and distracters at both encoding and retrieval.

Procedure. All subjects were tested individually. Upon arrival, the subject signed an informed consent form and sat directly in front of a tachistoscope. The subjects was informed that (s)he is participating in a study designed to find out how people recognize faces. The subjects was then instructed that (s)he would be shown a series of faces, each of which would be presented one at a time for a few seconds.

The subject was then shown 60 faces (30 targets 30 distracters) for 3 s. each using a tachistoscope. In between each presentation, the subject viewed a blank white card for 3 s. After viewing all 60 faces, the subject completed a self-report measure of anxiety as a filler task. The filler task took approximately 10 minutes to complete, after which the subject proceeded to the retrieval phase of the experiment.

At retrieval, the subject was asked to view another set of pictures. This set consisted of 30 previously seen targets and 30 new distracters. Depending on condition, the subject was given an instruction sheet stating either 1 or 3 trait or physical feature questions regarding each face. Subjects in the three- holistic condition rated each face for intelligence, attractiveness, and height (all of which are believed to induce holistic judgments). Subjects responded to these questions in a dichotomous fashion stating whether the target face was above or below average in each of the three traits. In contrast, subjects in the three-featural condition rated each face's nose (long or short), eyes (close together-far apart), and lips (thin, full). Subjects in the one-question condition were asked one of three questions mentioned above (with question counterbalanced across subjects). After rating each face for either holistic or featural questions, the subject stated whether the present stimulus was an old (previously seen) or a new face.

Results

For each subject the number of hits and false alarms were calculated. The mean number of hits and false alarms were, respectively, 17.69 (sd = 5.07) and 8.78 (sd = 4.44), each out of a possible 30. In addition, d', a signal detection measure of sensitivity or accuracy which takes into account both hit-rate and false alarm-rate, was calculated for each

11

subject. These scores ranged from -.17 to 2.41 and averaged .86 (sd = .50).

The first analysis compared d' as a function of question type (holistic versus featural) and number of questions (one versus three) in a 2 X 2 analysis of variance. This analysis yielded a significant main effect for question type, wherein holistic questions produced greater recognition accuracy than featural questions, F (1, 44) = 5.96, p < .05, eta-squared = .12, as shown in Figure 1.

Number of questions produced a nonsignificant main effect, F (1, 44) = 1.12, p > .05, eta-squared = .02. The interaction was also nonsignificant, F (1, 44) = .02, p > .05, eta-squared < .01. Thus, the data was collapsed across number of questions for further analyses. The next analysis compared d' scores as a function of holistic, featural, or no question at retrieval by computing a one-way, three level analysis of variance. The mean d' for the featural, holistic, and control conditions are shown in Figure 2.

This analysis produced a significant main effect for question type, F (2, 69) = 6.04, p < .01, eta-squared = .15. Pairwise comparisons revealed that subjects who responded to either no questions or to holistic questions were significantly more accurate than subjects who responded to featural questions (p < .05 for each comparison). The control and holistic groups did not differ significantly (p > .05).

Discussion

Results suggest that processing instructions can influence recognition accuracy. That performance was comparable for the holistic and no instruction groups. That both differed significantly from the featurally instructed group, suggests that subjects, when left alone, devote maximum attention to interfeatural information, at least within the time constraints posed by the experimental procedure. Alternatively, it may be that three questions is not enough to enhance holistic processing above baseline. One featural question is apparently sufficient to reduce recognition accuracy. I refer to this as the "featural inferiority effect." Like Sporer (1991), I found that increasing the number of featural questions at encoding had trivial effects on subsequent recognition performance.

Experiment 2

Overview and Hypotheses

Experiment 2 has two purposes. First, I attempted to replicate the findings of Experiment 1. Second, I examined whether encoding strategy qualifies the effect of processing instructions at retrieval. The design of the Experiment 2 is a 3 (holistic versus featural versus no questions at encoding) X 3 (holistic versus featural versus no questions at retrieval) + 2 (same holistic and featural instructions at encoding and retrieval) factorial in which all variables were manipulated between subjects. Subjects in the current

13

experiment viewed 60 faces at encoding and received either holistic, featural, or no questions during encoding. At retrieval, subjects were asked to view another set of pictures. This set consisted of 30 previously seen targets and 30 new distracters. During retrieval operations subjects also received either holistic, featural, or no questions regarding each individual face prior to giving a recognition judgment.

Results from Experiment 1 indicate that featural processing at retrieval reduced recognition accuracy. I entertain two hypotheses regarding the effects of the encoding conditions.

Matching superiority. The matching hypothesis states that the effects of featural judgments on processing at retrieval is comparable to its effect at encoding, thus producing a match between encoding and retrieval processes. As in Wells and Turtle (1988) and Wells and Hryciw (1984), this hypothesis predicts that recognition accuracy would be significantly higher when the encoding and retrieval processes match (holistic/holistic, featural/featural) than when there is a mismatch (holistic/featural, featural/holistic). Evidence for the matching superiority hypothesis would be a significant interaction showing the pattern just described.

Featural inferiority. The featural inferiority hypothesis holds that, although featural processing at

14

recognition reduces attention to interfeatural information, its effect at retrieval is small in comparison to its effect at encoding. Because holistic processing dominates the recognition task regardless of the processing instructions, the effect of featural processing at either encoding or retrieval reduces recognition accuracy. Evidence for the featural inferiority hypothesis would be main effects for processing conditions at encoding, retrieval or both, with featural conditions showing inferior performance.

Two concerns necessitated inclusion of appended cells. On the one hand, because I desire to make claims about processing instructions in general (i.e., holistic versus featural) and not about specific questions used to induce processing, it seemed reasonable to use different questions at encoding and retrieval within a processing condition. On the other hand, it could be argued that by using different questions at encoding and retrieval (i.e., within a processing condition), the processing strategies invoked by the different questions might create less of a match. In order to address this concern empirically, the present factorial design consists of different questions at encoding and retrieval, and appended cells contain identical questions at encoding and retrieval: one appended condition for identical featural questions and another for identical holistic questions. If the specific questions influence the match between processes, then the appended cells should

15

produce superior recognition performance than the cells containing different questions at encoding and retrieval (within processing condition).

Method

Subjects. Participants were 220 students (69 female, 21 male) from introductory psychology classes at Florida International University. Subjects were assigned randomly to one of the eleven conditions (20 per condition).

Materials. Photographs and counterbalancing orders were the same as in Experiment 1.

Procedure. All participants were tested individually. Upon arrival, the subject signed an informed consent form and sat directly in front of the tachistoscope. The Subject was then informed that (s)he is participating in a study designed to find out how people recognize faces. Depending on condition, the subject was then given an instruction sheet stating two trait or physical feature questions regarding each face. Subjects in the holistic condition rated each face on two of the following personal inferences: attractiveness, intelligence, friendliness and height. In contrast, subjects in the featural encoding condition rated each face on two of the following physical features: nose (long v. short), lips (thin v. thick), jawline (angular v. rounded) and eyes (close together v. far apart). The subject responded to these questions by stating whether the target face was above or below average on each of the two

traits. After a practice trial, the subject was instructed
that (s)he would be shown a series of faces, each of which
would be presented one at a time for a few seconds, and that
he/she should answer both (holistic or featural) questions
for each face. Subjects in the no-instruction condition
were instructed that they would be shown a series of faces
each presented individually for a few seconds.

The Subject was then shown 60 faces (30 targets, 30
distracters) for 3 s each, using the tachistoscope. In
between each presentation, the subject viewed a blank white
card for 3 s. After viewing all 60 faces, the subject
completed a self-report measure of anxiety as a filler task.
The filler task took approximately 10 minutes to complete,
after which the subject proceeded to the retrieval phase of
the experiment.

At retrieval, the subject was asked to view another set
of photos. This set consisted of 30 previously seen targets
and 30 new distracters. Depending on condition, the subject
was given another instruction sheet stating either the
identical encoding questions or the remaining 2 of the 4
trait or physical feature questions mentioned above. For
example, subjects receiving dissimilar holistic questions at
both encoding and retrieval (holistic/holistic group) rated
each face for friendliness and attractiveness at encoding
and intelligence and height at retrieval. In contrast,
subjects in the mixed questions conditions

(holistic/featural or featural/holistic) received two of the
four questions from each group. Questions were
counterbalanced across subjects. Subjects in the two
appended cells (holistic/same and featural/same) received
identical trait or featural questions at both encoding and
retrieval. After rating each face for either holistic or
featural questions, the subject stated whether each picture
was an old (previously seen) or a new face. Subjects in the
condition who receive no instructions at retrieval proceeded
directly to the recognition judgments.

Results

The mean number of hits and false alarms were,
respectively, 18.04 ($\underline{sd}$ = 4.60) and 9.63 ($\underline{sd}$ = 5.19), each
out of a maximum of 30. The mean $\underline{d'}$ was .78 ($\underline{sd}$ = .55), and
the range was -.15 to 2.23.

The first analysis compared $\underline{d'}$ as a function of encoding
instructions (holistic versus featural versus none) and
retrieval instructions (holistic versus featural versus
none) in a 3 X 3 analysis of variance. The cell means are
displayed in Table 1. This analysis yielded a main effect
for the encoding instructions, wherein the holistic group
was more accurate in recognizing faces compared to both the
control and featural conditions $\underline{F}$ (2, 171) = 8.83, $\underline{p}$ < .01,
eta-squared = .09 (see Figure 3).

Tukey-Kramer pairwise comparisons revealed that subjects
who responded to holistic questions were significantly more

18

accurate on face recognition than were subjects in the featural and control groups (p < .05). The featural and control groups did not differ significantly.

Retrieval instructions produced a nonsignificant main effect, F (2, 171) = .52, p > .05, eta-squared < .01 (see Figure 4). Although this main effect was nonsignificant, the pattern of means is comparable to that obtained in Experiment 1. I therefore conducted a combined probability analysis to assess whether the holistic and featural retrieval conditions differ significantly when the results of Experiments 1 and 2 are considered together. The standardized differences (d = mean difference in d' divided by the standard deviation) between the holistic and featural retrieval conditions were .64 for Experiment 1 and .30 for Experiment 2. The z test for combined probabilities indicated that the retrieval effect was significant across the two experiments (p < .05). See figure 4.

The interaction between encoding and retrieval conditions was nonsignificant, F (4, 171) = .77, p > .05, eta-squared = .02. Comparisons between the appended cells, wherein subjects received identical processing question at encoding and retrieval, and the cells in which the processing category was the same but the questions were different, revealed no significant differences in recognition performance (ps > .05).

Discussion

Results of Experiment 2 support the featural inferiority hypothesis. First, the main effect of retrieval condition was nonsignificant and was substantially smaller in magnitude than the effect of encoding condition. Surprisingly, the effect of retrieval condition in Experiment 2 was considerably smaller than the effect in Experiment 1, although the combined probability analysis suggests that the effect is significant when the results of both experiments are combined. Given the similarity in the subjects, procedures and materials, I am at a loss to explain the difference in the magnitudes of the effect. Second, the lack of evidence of an interaction between encoding and retrieval conditions further argues in favor of the featural inferiority hypothesis for face recognition. On the other hand, the lack of evidence for the matching hypothesis could be due to the general weakness of the retrieval effect in Experiment 2.

As in the previous research described above, holistic processing instructions at encoding produced superior recognition accuracy in comparison to featural processing instructions. Unlike previous research, a no instruction group was included to assess baseline performance. The no instruction group performed comparably to the featural processing group. Thus, while the results support featural inferiority at the retrieval stage, the encoding effects are better characterized as holistic superiority.

20

A potential confound in Experiments 1 and 2 may be that subjects in the holistic condition spent more time examining the faces at retrieval than did subjects in the featural processing condition. Enhanced exposure time at retrieval could conceivably influence recognition accuracy. I know of no empirical test of exposure duration at retrieval, and I did not assess processing time in my experiments. Thus, this confound is addressed this in Experiment 3.

Experiment 3

Subjects examined 60 faces for 3 seconds each at encoding and then for either 3, 10, or 30 seconds (manipulated between subjects) at retrieval. No processing manipulation was utilized at either encoding or retrieval. It was expected that viewing time would have minimal effects on subjects' recognition accuracy.

Method

Subjects. Participants were 30 students (17 females, 13 males) from a introductory psychology classes at Florida International University. Subjects were assigned randomly to one of the three conditions (10 subjects in each).

Materials. Materials and counterbalancing conditions were the same as those used in Experiments 1 and 2.

Procedure. All subjects were tested individually. Upon arrival, the subject sat directly in front of the tachistoscope and was informed that (s)he is participating in a study designed to find out how people recognize faces.

The subjects was then instructed that (s)he would be shown a series of faces, each of which would be presented one at a time for a few seconds. At encoding, the subject was shown 60 faces (30 targets, 30 distracters) for 3 s each using a tachistoscope. In between each face, the subject viewed a blank card for 3 s. After viewing all 60 faces, the subject was taken into another room to read an unrelated manuscript as a filler task. The filler task lasted approximately 10 min., after which the subject participated in the retrieval phase of the experiment. At retrieval, the subject was asked to view another set of photos. The retrieval stimuli consisted of 30 previously seen targets and 30 new distracters. The subjects was instructed to study each face for the full duration as it appeared on the screen. Depending on condition, subjects viewed each face for either 3 s, 10 s, or 30 s. After viewing an individual face for the specified time, the subject stated whether the face was old (previously seen) or new (not previously seen). Thus, the subject gave recognition judgments after each face was removed from the screen (and a blank card was in view) and did not view another face until the recognition assessment was disclosed and recorded.

Results and Discussion

The mean number of hits and false alarms were respectively, 19.07 (sd = 4.31) and 5.53 (sd = 3.64), (each out of a possible 30). The Mean d' was 1.20

($\underline{sd}$ = 5.08). The mean d' for the 3 s, 10 s, and 30 s conditions were 1.24, 1.04, and 1.33 respectively. A one-way three level analysis of variance yielded a nonsignificant main effect for viewing time $\underline{F}$ (2, 29) = .931, $\underline{p}$ > .05, eta-squared = .07. Thus, exposure duration at retrieval is unlikely to explain the findings of Experiments 1 and 2.

General Discussion

In conclusion, results of this set of experiments support the notion that face recognition performance can be influenced by processing instructions at retrieval, but its effect is considerably smaller than the effects of processing instructions at encoding. The difference in magnitude of the retrieval effect from Experiment 1 to Experiment 2 is troublesome. There are several ways to interpret this. First, as demonstrated above, the combined probability analyses reveals a significant main effect when the two sets of results are considered jointly. Second, it could be that the significant retrieval effect of Experiment 1 is due to Type I error or that the nonsignificant retrieval effect of Experiment 2 is due to Type II error. In any case, further attempts at replication and investigations of qualifying effects are called for. In addition, it would be useful to test the retrieval effect with a more ecologically valid methodology, such as a crime simulation. Future research should also vary the methods

23

used to induce holistic and featural processing at retrieval. Perhaps some methods would be more effective than others.

Although the magnitude of the retrieval effect remains in question, the pattern of means appears clear: featural processing at retrieval reduces recognition accuracy. These results are consistent with the findings of Woodhead, Baddeley, and Simmonds (1979). In their attempt to enhance face recognition accuracy, they trained subjects to recognize facial features. The training consisted of lectures, films, and panel discussions. Training did not significantly improve face recognition accuracy; indeed, in one study, trained subjects were significantly less accurate than untrained subjects on a face recognition test.

If the pattern of results in these experiments continues to hold in future research, then implications for police practices would be as follows. It would appear that instructions designed to facilitate holistic processing at retrieval do not improve identification accuracy (as compared to leaving witnesses to their own devices); rather, instructions to focus on individual facial features impairs recognition accuracy. Thus, police investigators should refrain from drawing the eyewitness' attention to individual facial features during mugshot searches and identification tests.

Research on retrieval factors associated with eyewitness memory has produced valuable techniques for enhancing eyewitness recall (Fisher et al., 1989) and identification accuracy (Cutler & Penrod, 1988; Lindsay et al., 1991; Luus & Wells, 1991). Future research should continue to investigate the effectiveness of processing instructions and other potentially useful retrieval factors.

References

Baddeley, A. D. (1979). Applied cognitive and cognitive applied psychology: The case of face recognition. In L. G. Nilsson (Ed.), <u>Perspectives on memory research</u>. Hillsdale, NJ: Erlbaum.

Bower, G. H., & Karlin, H. B. (1974). Depths of processing pictures of faces and recognition memory. <u>Journal of Experimental Psychology</u>, <u>103</u>, 751-757.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. <u>Journal of Verbal Learning and Verbal Behavior</u>, <u>11</u>, 671-684.

Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. <u>Journal of Applied Psychology</u>, <u>73</u>, 281-290.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. <u>Journal of Applied Psychology</u>, <u>74</u>, 722-727.

Light, R. C. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. <u>Journal of Experimental Psychology: Human Learning and Memory</u>, <u>5</u>, 212-228.

Lindsay, R. C. L., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. <u>Journal of Applied Psychology</u>, <u>76</u>, 741-745.

Luus, C. A. E., & Wells, G. L. (1991). Eyewitness
    identification and the selection of distracters for
    lineups. Law and Human Behavior, 15, 43-57.

Patterson, K. E., & Baddeley, A. D. (1977). When face
    recognition fails. Journal of Experimental Psychology:
    Human Learning and Memory, 3, 406-417.

Smith, A. D., & Winograd, E. (1978). Adult age differences
    in remembering faces. Developmental Psychology, 14,
    443-444.

Sporer, S. L., (1991). Deep--Deeper--Deepest? Encoding
    strategies and the recognition of human faces. Journal of
    Experimental Psychology: Learning, Memory, and Cognition,
    17, 323-333.

Strnad, B. N. & Mueller, J. H. (1977). Levels of processing
    in facial recognition memory. Bulletin of the Psychonomic
    Society, 9, 17-18.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity
    and retrieval processes in episodic memory. Psychological
    Review, 80, 352-373.

Warrington E. K., & Ackroyd, C. (1975). The effects of
    orienting tasks on recognition memory. Memory and
    Cognition, 3, 140-142.

Wells, G. L., (1978). Applied Eyewitness Testimony Research:
    System variables and Estimator variables. Journal of
    Personality Social Psychology, 36, 1546-1557.

Wells, G. L., & Hryciw, B. (1984). Memory for faces:
Encoding and retrieval operations. <u>Memory and Cognition,</u>
<u>12</u>, 338-344.

Wells, G. L., & Turtle, J. W. (1988). What is the best way
to encode faces. In M. M. Gruneberg, P. E. Morris, & R.
N. Sykes (Ed.), <u>Practical aspects of memory: Current</u>
<u>research and issues: Vol. 1. Memory in everyday life</u> (pp.
163-168). New York: John Wiley & Sons.

Winograd, E. (1976). Recognition memory for faces following
nine different judgments. <u>Bulletin of the Psychonomic</u>
<u>Society</u>, <u>8</u>, 419-421.

Winograd, E. (1981). Elaboration and distinctiveness in
memory for faces. <u>Journal of Experimental Psychology:</u>
<u>Human Learning and Memory</u>, <u>7</u>, 181-190.

Woodhead, M. M., Baddeley, A. D., & Simmonds, D. C. V.
(1979). On training people to recognize faces.
<u>Ergonomics</u>, <u>22</u>, 333-343.

Table 1

Experiment 2 Cell Means ($\underline{d}'$)

| | Retrieval Condition | | |
	Featural	No Questions	Holistic
Encoding Condition			
Featural	.62	.80	.56
No Questions	.56	.70	.67
Holistic	.97	.97	1.11

Appended Conditions

Featural/Featural (Same Questions): .52

Holistic/Holistic (Same Questions) 1.08

Figure 1

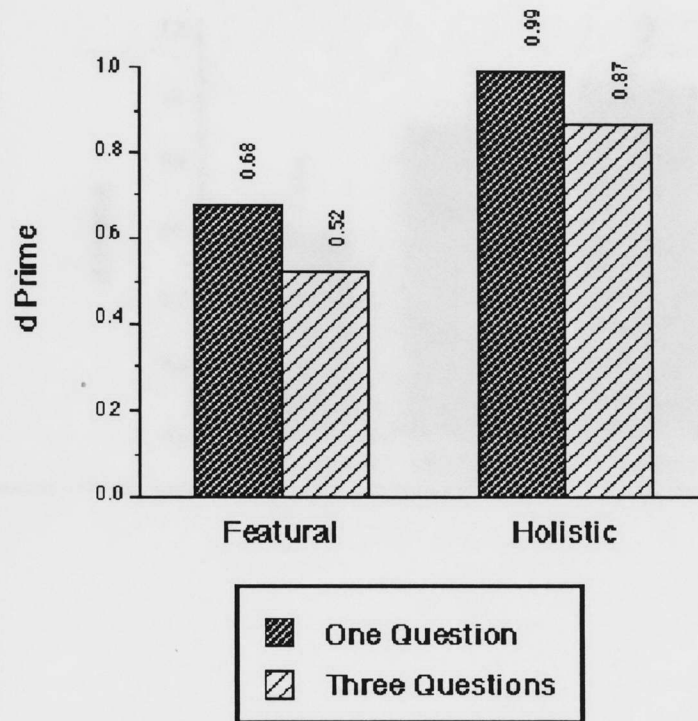## Effects of type and number of questions on face recognition accuracy.

Figure 2

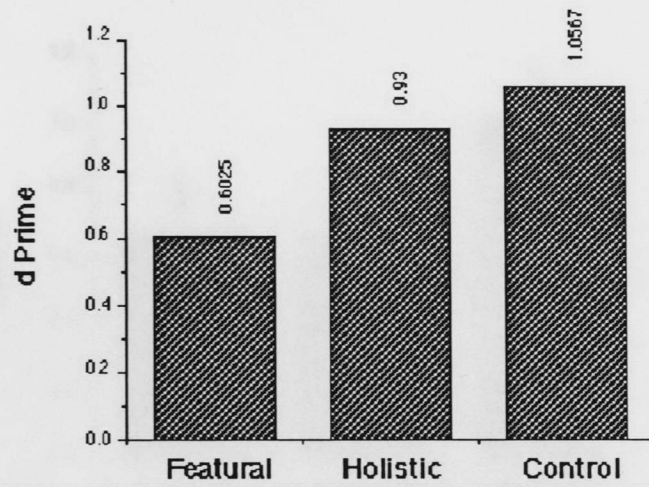## Effect of retrieval conditions upon recognition accuracy.

Figure 3
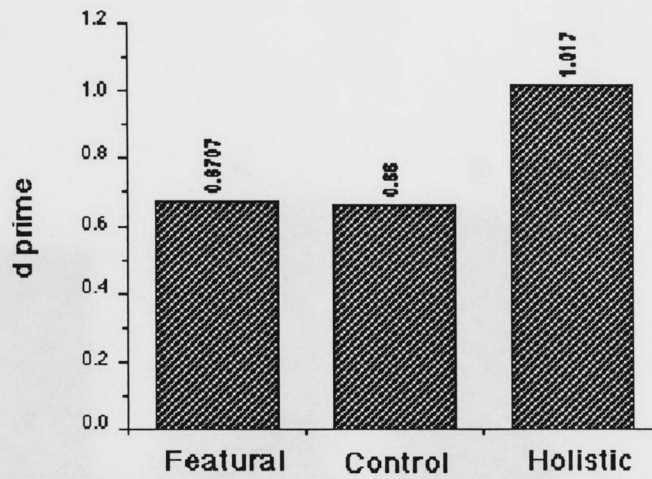
Effect of Encoding Questions on Face Recognition Accuracy.

Figure 4