Massey University Library. Thesis Copyright Form

Title of thesis: The Effect of Retention Interval and Target-Decoy Similarity on Facial Recognition

(1) (a) I give permission for my thesis to be made available to readers in the Massey University Library under conditions determined by the Librarian.

~~(b) I do not wish my thesis to be made available to readers without my written consent for ______ months.~~

(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

~~(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for ______ months.~~

(3) (a) I agree that my thesis may be copied for Library use.

~~(b) I do not wish my thesis to be copied for Library use for ______ months.~~

Signed W. Rockel / Rockel

Date 1/3/91

NAME AND ADDRESS

DATE

THE EFFECT OF RETENTION INTERVAL AND TARGET - DECOY SIMILARITY ON FACIAL RECOGNITION

A thesis presented in partial fulfilment
of the requirements for the degree
of Master of Arts in Psychology
at Massey University.

Warren Rockel

1991

ACKNOWLEDGEMENT

I would very much like to thank my thesis supervisor, John Podd, for his expert guidance in helping me through all the stages involved in the preparation of this thesis.

ABSTRACT

This research was an attempt to resolve the inconsistent results for the effect of delay interval on facial recognition. The theory tested was that the degree of target / decoy similarity may act either to enhance or to diminish the effect of delay primarily by influencing false alarm rates. The first experiment used a novel method to scale the 80 faces along the dimension of similarity. The results showed that the method used was reasonably successful in ordering the faces along the similarity dimension. It enabled the use of four sets of 20 faces as either low or high similarity decoy and target sets in a second experiment aimed at testing the proposed theory. It was predicted that high target / decoy similarity would result in a greater effect of delay than low target / decoy similarity. Six groups of 15 subjects completed a standard face recognition experiment which crossed 0, 1 and 21 days delay with high and low similarity target / decoy sets. The results showed a main effect for similarity, but, surprisingly, no main effect for delay. Nor was there the predicted interaction between similarity and delay for false alarms. The failure of the second experiment to test adequately the theory, and reasons for failure are discussed, along with the importance of the link between similarity and delay.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Overview

As humans, the faces of others of our species are for most of us the primary means of distinguishing one person from another. Ellis (1981) wrote that "No other object in the visual world is quite so important to us as the human face."(p.1). It is by their facial characteristics that we learn to recognize friends, relatives, and acquaintances. As social animals, this ability is extremely important if we are to interact and communicate effectively with others of our kind. So our ability to remember and recognize faces is of prime importance to our adequate functioning in everyday life.

The field of facial recognition research is one aspect of the cognitive approach to face processing, which studies how it is that we perceive and remember faces. In the real world outside the laboratory we recognize familiar faces daily. For instance, as we pass someone in the street, by a glance at their facial features we are able to identify that person as someone that we know. We make a judgement as to whether or not the face that we are looking at is that of so-and-so.

For controlled research, it is necessary for experimenters to define what exactly they are investigating. Thomson (1986) distinguishes four different definitions of face recognition:

a) The observer knows that a particular shape or form is that of a face;

b) The observer knows that a particular face has been seen before;

c) The same as b), but the observer knows that the face was seen before at a particular time or place; and

d) The observer knows the name or identity of the face.

For the purposes of the present study, face recognition is considered to be whether the observer knows that a particular face has been seen before. The observer is not required to identify a face by name, nor to state where or when the face was seen before. Recall of faces is a different process, which requires that the details of a particular face be retrieved from memory when the face to be recalled is not actually present. This phenomenon is not a common occurrence in our everyday lives, so is accordingly little studied.

An obvious practical use of face recognition research is for investigating the accuracy of eyewitness evidence. In a criminal trial it is crucial that if an identification on the basis of a suspect's face is to be made, that it can be done accurately. The injustices that can occur after the misidentification of a suspect have been well documented in the literature (e.g., see Yarmey, 1979b; and Shepherd, Ellis, & Davies, 1982). In a police lineup or a criminal trial there is great motivation on the part of a victim of a crime or accident to remember the face of their assailant. And here, too, the complexity and individuality of a face is critical.

It is important to note that eyewitnesses to a crime or accident are most often asked to try to recognize an offender when some time has elapsed since the incident occurred. Just how long after the initial viewing of a face is an observer able to state accurately whether or not that face has been seen before? So it is vital that we know not only the circumstances under which eyewitness testimony may be deemed reliable, but more specifically, for exactly what period of time after the event can a face be accurately retrieved from memory. Hence, the retention interval between study and test phases is one of the most important variables in facial recognition research.

Although common sense might imply that one would expect subjects' memory for previously seen faces to deteriorate over time, research to date portrays a confused picture. Deffenbacher (1986) states that of the 33 studies of forgetting that he examined, roughly half yielded no statistically reliable effect of retention interval. Goldstein and Chance (1981) remark that although the psychology literature is replete with studies demonstrating that forgetting occurs over time, there has been a lack of systematic laboratory research on the effect of delay on facial recognition. Even in the nine years since Goldstein et al. made that comment, very little in the way of systematic research has been carried out. So it seems a fair supposition that there may be other variables interacting with delay to produce these confusing results, and a systematic investigation into what exactly these variables might be is long overdue.

Podd (1990) has suggested that an obvious variable is the degree of similarity between targets and distractors. Davies, Shepherd, and Ellis (1979) have pointed out that in the bulk of face research, target and distractor faces have been selected at random, and systematic study of relative similarity between targets and decoys has been neglected. Davies et al. found that the degree of similarity had an effect on recognition performance in their study.

Shepherd et al. (1982) report the Revised Scottish Guidelines for the composition of identification parades (these may reasonably be taken as representative of practices in Police forces elsewhere) which stress the importance of placing the accused "beside persons of similar age, height, dress and general appearance." (p.133). These guidelines merely codify what has long been known regarding lineup composition: That putting a suspect or target individual in line with decoys who are physically dissimilar to that suspect results in a biased lineup. The suspect "sticks out like a sore thumb", because the witness is given no real choice.

So, there are some indications that the mixed results of previous delay studies could be due to a lack of control over the degree of similarity between targets and distractors. The major aim of this study was to investigate the relationship between delay and similarity.

Typical Recognition Study

In some studies, trying to present as realistic a situation as possible, live crime scenarios are played out in front of an unsuspecting audience. Subsequently they are asked to play the role of witnesses to the crime, in identifying suspects, rating their degree of confidence in their choice, and so on, as happens in actual police investigations (e.g., Buckhout, Alper, Chern, Silverberg, & Slomovits, 1974; Egan, Pittner, & Goldstein, 1977). But not all researchers go to such lengths in attempting to emulate every action that the witness to a crime or incident goes through. More often, the process is reduced to the fundamental act of subjects being shown a face or faces in a laboratory setting, then undergoing a standard recognition test, as described below.

A typical face recognition study involves presenting subjects with a number of photographs of faces, usually in the form of slides presented sequentially. The photographs usually have been black and white, and show only a full-frontal view of the face. This is the "study phase" (or "inspection phase"), and generally subjects are informed that they should pay close attention to the faces (known as targets), as they shall later be requested to attempt to recognize them. However, Courtois and Mueller (1981) found that it made little difference to the results whether or not subjects were told that a recognition test would follow.

In the "recognition phase" (or "test phase"), which can take place either immediately or after a delay, the subjects are shown the target faces again, this time randomly interspersed with other faces, called distractors or decoys, which they have not seen before. Subjects are asked to indicate which of the faces they think they have seen before, by rating them as either old (previously seen) or new (not seen before), often giving the level of confidence in their decision also. The ratio of targets to distractors used in different studies varies a great deal. Laughery, Fessler, Lenorovitz, and Yoblick (1974) used just one target to 149 distractors in their recognition test. However, more commonly a ratio of between 1 : 2 and 1 : 4 is used. Shapiro and Penrod (1986), in their meta-analysis of facial identification studies, found a mean of 22 targets shown at study and recognition phases, with a mean of 40 decoys in the recognition test.

The faces used are either of males only, or of males and females; seldom are only female faces used. Most studies use only white (Caucasian) faces. The length of delay between study and recognition phases varies greatly, with many studies using several different retention intervals for comparison, as well as an immediate test as a control. Deffenbacher (1986) reported a "vast range" of retention intervals tested in the literature on laboratory studies of face recognition, from "one minute to 350 days" (p.63). Shapiro and Penrod (1986) found a mean delay of 4.5 days, with a standard deviation of 21 days.

The present study is limited in that it is a laboratory study of face recognition. The faces used as stimuli are still photographs, and show only a full frontal view. It is hoped that in spite of this simplification in the present study and others in the literature, the subject's task remains an adequate representation of what takes place in the real world, and is generalizable to it.

Signal Detection Theory Measures

Most relatively recent studies make use of Signal Detection Theory (SDT) measures to determine performance in face recognition. In the present study, four of the most common SDT measures were used: hits, false alarms, d', and A_g. The following discussion of these measures is largely drawn from Banks (1970) and McNicol (1972).

In applying SDT to facial recognition, the memory trace is considered as a signal which the subject must detect. SDT is used to separate the truly retention-based aspects of memory performance from the decision aspects (for instance, subjects may appear to be insensitive because they are extremely cautious and only report signals they are certain of). The subject is required to make one of two possible responses to each stimulus, according to whether he or she can detect a memory trace for it: "yes - this is an old item," or "no - this is a new item." Thus, hits occur when the subject gives an "old" response, given that the stimulus was seen before.

A false alarm is when the subject states that a stimulus is old, when in fact it is new.

Hits and false alarms, collected under varying degrees of decision bias, can be plotted against each other to yield a receiver-operating-characteristic (ROC) curve. A_g is the area underneath this curve, being a measure of observer sensitivity, independent of the decision criterion. Like A_g, d′ is also a criterion-free index of recognizability. It is defined as the z-score of the false alarm rate minus the z-score of the hit rate. But unlike A_g, d′ assumes underlying normal-normal equal variance distributions.

According to SDT, all points on an ROC curve represent equivalent retention. They differ only in the degree of caution shown by the subject. A cautious subject may score fewer hits, but also gets fewer false alarms, and, likewise, a lax subject produces more hits but also more false alarms. Some researchers have reported the results of their studies in terms of hits or false alarms alone. This practice of reporting one or other in isolation may be misleading, because either can vary as a result of changes in response bias, and may not in fact indicate that there has been a change in recognition accuracy.