

# Eyewitness identification: Accuracy of individual vs. composite recollections of a crime

ANDREA ALPER, ROBERT BUCKHOUT, SUSAN CHERN, RICHARD HARWOOD  
and MIRIAM SLOMOVITS

*Brooklyn College of City University of New York, Brooklyn, New York 11210*

Individual eyewitness recall reports were gathered from witnesses who were later put into groups to discuss and reach a consensus on a description of a simulated crime they had witnessed. Groups gave more complete reports but at the price of a significant increase in errors of commission (the fabrication of details under group pressure).

The investigation of a crime after the fact, where there were multiple witnesses, poses a dilemma which underlies the research. Any criminal investigator will be searching for a coherent set of facts to explain *all* of the events in order to proceed in searching for a suspect or in presenting a case for prosecution. Any one witness may not have seen enough or may be unable to supply the completeness of detail which the investigator needs to make his report. He may naturally turn to a group and attempt to get a composite description, or he may be rushed to write down the inputs of several witnesses at the scene of the crime. The advantages are many—more details, consensus, screening out irrelevancies, etc. Add these to the tendency to give less credence to the perceptions of a single observer, and we describe the practice of many a police officer who must himself become a spokesman for what happened. Common sense coincides with empirical research (for a change) in demonstrating, as did Hall, Mouton, and Blake (1963), that decisions made by groups after interaction are superior to decisions based on the statistical pooling of individual judgments.

In the case of research on eyewitnesses, the percentage recall accuracy of individual witnesses to simulated crimes has consistently been found to be low (Buckhout, 1975; Marshall, 1966; Muensterberg, 1908), due to variations in attention, set, stress, and environmental factors. In the practice of criminal investigation, great emphasis is placed on composite descriptions and the accumulation of individual reports (e.g., *Warren Commission Report*, 1964). As far as we know, there has been no research on the comparative accuracy of group vs. individual reports of events which have the complexity of most crimes. Yet the faith of criminal investigators in consensual validation of uncheckable prior perceptions may be poorly founded if we heed the lessons of research on the effects of group dynamics on relatively simple perceptual events.

The essence of the classic research of Asch (1956) was to demonstrate that a group that puts forth a false consensus could effectively sway a correct perceiver to adopt the majority's position. A minority group can successfully influence a majority's decision by the degree of persistence with which the minority maintains its position (Moscovici & Neve, 1973). Thus, conformity may be shown to be an effect of consistency of position, not necessarily the more commonly assumed notion of strength of social pressure.

The purpose of the present study was to determine whether a group of witnesses to a simulated crime would be more accurate in their collective recall as compared with their pooled individual data. The standard of accuracy used a percentage of the total number of facts recalled, permitting the measurement of errors of omission and errors of commission (the number of fabricated facts).

## METHOD

### Subjects

The subjects were 29 undergraduate students at Brooklyn College who were enrolled in a course in social psychology. Participation in the experiment was required.

### Materials

Two questionnaires were employed. The witnesses (subjects) were required to complete a form that consisted of questions relating to the incident itself, a description of the perpetrator, and their own relationship to the incident. The leaders (experimenters) were asked to complete a form that included questions relating to their observations of the processes employed in their respective groups.

A sweep-second stopwatch was used to time the duration of the incident.

### Design

The general research design was to stage an incident in front of a group of students. The incident was such that it appeared to be genuine at the time it took place. Following the incident, the students were then required, individually, to answer questions on what had occurred. Upon completion

**Table 1**  
**Mean Errors in Recall by Individuals and Groups**

	Group	Averaged Individual Data	Group	Averaged Group Data
Omission Errors	1	10.13	1	9.0
	2	14.29	2	3.0
	3	16.17	3	7.0
	4	9.25	4	3.0
	Total	10.96	Total	5.5
Commission Errors	1	6.63	1	5.0
	2	3.71	2	11.0
	3	5.66	3	5.0
	4	6.38	4	11.0
	Total	5.60	Total	8.0
Total Errors	1	16.75	1	14.0
	2	18.0	2	14.0
	3	15.83	3	12.0
	4	15.63	4	14.0
	Total	16.55	Total	13.5

of the individual questionnaires, the students were divided into four groups and required to complete the same questionnaire (as a group). Each of the four experimenters took the role of leader/observer of one of the four groups. The leader/observer asked the questions and recorded the responses.

#### Procedure

The "incident" that took place during the class was as follows: Shortly after the class began, a male student knocked at the door and stated that he thought he had left a book in the room during the preceding class and that he would like to look for it. The professor (who was aware that this was "staged") consented to let him look but asked him to make it brief since he was in the middle of a lecture. The student walked over to the chairs where two of the experimenters were sitting. He bent over between the chairs as if to look for a book, then grabbed the pocketbook of one of the experimenters and fled out the door. The experimenter whose pocketbook was stolen, screamed and then ran out the door after him. The instructor cautioned the class to remain seated, and then he, too, ran out of the classroom. Several minutes later, both the instructor and the experimenter returned to the classroom and announced that the pocketbook had been dropped in the hallway. The experimenter, very tense and nervous, checked the contents of the pocketbook and said that everything was there. The excitement level in the room was rather intense, and there was little doubt that the class believed that an actual purse snatching had occurred.

After several minutes, one of the experimenters announced that the incident had been staged and that this was part of an experiment. The questionnaires were distributed to the members of the class and instructions were given for them to answer the questions to the best of their ability and not consult with each other with regard to the answers.

Upon the completion and collection of the individual questionnaires, the class was divided into four groups of equivalent size. Each of the four experimenters served as leader/observer of one of the groups. Each group was asked to complete the same questionnaire again, but this time they were instructed to arrive at a "group answer" for each of the questions. They could use any method they found suitable, but only one answer for each question would be accepted. This procedure was followed until all questions on the form were completed. The

experimenter's function was to ask the questions, keep the group at the task until a conclusion was reached, and record all final answers.

## RESULTS

A comparison of the mean errors in recall by individuals and groups is shown in Table 1. Group consensus descriptions were more complete than averaged individual descriptions, with groups making an average of 5.5 errors of omission as compared with an average of 10.96 errors for individuals. However, completeness was obtained at the cost of increased inaccuracy, with groups obtaining 40% more errors of commission (including incorrect details in their descriptions) than individuals.

The mean confidence levels for groups and individuals are shown in Table 2. All subjects expressed a reasonably high degree of confidence in their answers, reporting an average of 7.1 [on a scale of 1 (least) to 10 (most)] for both individual and consensus decisions. However, this confidence was not reflected in the accuracy of their responses with an averaged individual accuracy score of 32% and an averaged group accuracy score of 44%.

Group reports on the duration of the incident were much closer to the truth than averaged individual estimates (Table 2). The actual incident lasted 14 sec, with groups estimating an average of 19.38 sec and individuals estimating 37.93 sec ( $p < .05$ ).

Figure 1 presents a graphic presentation of the differences in the main descriptive data, averaging the groups and the individuals composites on which further testing would be based. The groups under-

**Table 2**  
**Mean Confidence Levels and Time Estimates by Individuals and Groups**

	Group	Averaged Individual Data	Group	Averaged Group Data
Confidence Level*	1	8.14	1	6.0
	2	6.14	2	6.0
	3	6.25	3	
	4	7.38	4	10.0
	Total	6.97	Total	7.3
Time Estimate**	1	57.14	1	12.5
	2	40.50	2	15.0
	3	14.40	3	20.0
	4	39.69	4	30.0
	Total	37.93	Total	19.38

\*Confidence level is reported on a scale of 1 (least) to 10 (most).

\*\*Time estimates are reported in seconds.

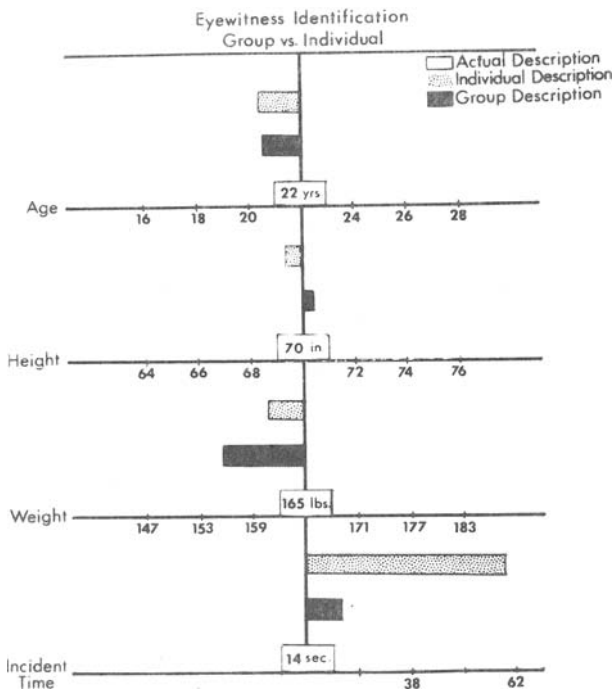


Figure 1. Comparison of mean individual and group estimates with actual physical characteristics of perpetrator.

estimated the weight, while individuals overestimated time.

## DISCUSSION

The experiment is very straightforward: A group of witnesses pooling their descriptions of a crime came up with a more complete description of what happened than any single witness was able to provide. This performance is very pleasing to any investigator, who is, himself, under pressure to completely describe the crime. But, of course, this completeness is uncheckable in a real crime, and the "complete" description may be more a function of conformity to group pressure than a product of memory enhancement. We are concerned especially with the possibility that group interaction may pressure witnesses into offering inferences rather than perceptions, or even into offering outright fabrications.

For example, in staging our simulated crime before a group of casually dressed college students, the tacit assumption of most groups was that the "student criminal" would be dressed as they were. Thus, virtually all but one student (who had correctly perceived and remembered that the "perpetrator" wore brown corduroy pants) were in strict accord that he was dressed in blue dungarees—the stereotypic college "uniform."

However, the group process provided information superior in its accuracy in the case of estimation of the duration of the simulated crime. Although both individual and composite recollection of the duration produced overestimations (in

agreement with previous findings, e.g., Hogan, 1975, that "excessively filled" time intervals are experienced as longer than "moderately filled" intervals of objectively equal duration). The average group estimate of 19.38 sec was significantly closer to the truth (14 sec) than the averaged pooled individual estimate of 37.93 sec. In the context of group discussion, "way-out" individual estimates were brought in check through the efforts of group members who questioned the use of 32-min and 1-min "ball-park" estimates of time.

Time limitations prevented us from giving each witness in this study a lineup test of his ability to recognize the culprit. In a follow-up study, Rupp, Warmbrand, Karash, and Buckhout (1976) tested independent groups of witnesses with and without group discussion of the crime. While groups with interaction showed a mean increase in the completeness of their descriptions, the discussions had *no* effect on lineup (recognition) performance. In some cases, the discussions led some witnesses to change their descriptions to fit the available, but innocent, participants in the lineup. Since U.S. Courts require that witnesses make an identification which has an "independent basis" in perception (Levine & Tapp, 1973), group discussions would appear, on the surface, to invalidate that independent basis. We conclude that group discussions with witnesses is an ill-advised procedure that may double the effort of the investigator and compound the error from both a legal and perceptual point of view.

## REFERENCE NOTE

1. Rupp, A., Warmbrand, A., Karash, A., & Buckhout, R. *Effects of group interaction on eyewitness reports*. Paper presented at the meetings of the Eastern Psychological Association, New York, 1976.

## REFERENCES

- ASCH, S. Studies of Independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 1956, **70** (9).
- BALLOT, P. (Ed.). *Warren commission report*. Island Park, N. Y.: Scantronics Corporation, 1964.
- BUCKHOUT, R. The psychology of eyewitness testimony. *Law and Psychology Review*, 1975, **1**, in press.
- HALL, E. J., MOUTON, J. S., & BLAKE, R. R. Group problem solving effectiveness under conditions of pooling vs. interaction. *Journal of Social Psychology*, 1963, **59**, 147-157.
- HOGAN, H. W. Time perception and stimulus preference as a function of stimulus complexity. *Journal of Personality and Social Psychology*, 1975, **31**, 32-35.
- LEVINE, R. J., & TAPP, J. L. The psychology of criminal identification: The gap from *Wade to Kirby*. *University of Pennsylvania Law Review*, 1973, **212**, 1079.
- MARSHALL, J. *Law and psychology in conflict*. New York: Anchor, 1966.
- MOSCOVICI, S., & NEVE, P. Studies on polarization of judgements: III. Majorities, minorities and social judgements. *European Journal of Social Psychology*, 1973, **3**, 479-484.
- MUENSTERBERG, H. *On the witness stand*. New York: Doubleday Page & Co., 1908.

(Received for publication March 9, 1976.)