Hypermnesia in the eyewitness to a crime

PAUL EUGENIO, ROBERT BUCKHOUT, and STEPHEN KOSTES
Brooklyn College, City University of New York, Brooklyn, New York 11210

and

KATHERINE W. ELLISON
Montclair State College, Upper Montclair, New Jersey 07043

Forty students served as subjects in two groups of witnesses to which a surprise filmed crime was shown. All subjects gave free recall statements about the crime. Following either a directed thinking interval or a diversionary task to block rehearsal, witnesses gave two additional statements. Hypermnesia for hits and memory intrusions was observed, but d' scores did not show hypermnesia. The witnesses fell into two groups: "good" witnesses who had ascertained the correct schema and "bad" witnesses who were obviously on the wrong track. While most witnesses showed short-term hypermnesia for correct responses (hits), the good witnesses showed a significant increase in memory intrusions (false alarms) during the third recall trial (p < .01). Only with correct-schema witnesses did d' scores above the guessing level occur. Witnesses who had ascertained the correct schema but who were interrogated too often tended to make things up to suit the demands for detail, with a resultant increase in memory intrusions and lower d'.

In the real world of crime, witnesses and victims talk—and talk and talk and talk. In a short period of time, they may tell what happened to several police officers, a doctor, an assistant prosecutor, and their friends. Later, they may talk at a preliminary hearing and a grand jury hearing, to the next prosecutor who takes over the case, to the judge, and possibly to the jury. The stories told by witnesses tend to get longer over repeated sessions (Buckhout, 1974; E. F. Loftus, 1979), but are they any more accurate? Do witnesses remember more as they are cued? Or, like most people, do they just forget specific details and lengthen the narrative at the expense of accuracy?

The cognitive psychologist's view of memory as a three-stage process (Crowder, 1976; G. R. Loftus & E. F. Loftus, 1976) is not widely accepted by those whose responsibility it is to interview eyewitnesses. Acquisition and retention are taken for granted by many authorities in law enforcement and the courts to be akin to a photographic process characterized by permanence, with memory failure attributed to defective retrieval techniques by the witness or the investigator. In this framework, "more" is better, even though it may not be possible to independently check the accuracy of the more lengthy report. Controlled studies

of eyewitnesses to a simulated crime show that poor recall and recognition are the norm in those cases in which accuracy can be checked (Buckhout, 1974; Munsterberg, 1908). This poor performance can be attributed to factors affecting any or all of the stages of the memory process. Perhaps the main problem is that the critical events may not have been perceived at all. Eyewitnesses to a crime see the event only once. It is a "one-shot" perception that is affected by many factors, such as speed and movement, stress, stimulus overload, the fact that the perpetrator is a stranger to the witness, diversion of attention, excessive arousal, surprise, and limitations on the opportunity to observe the face (Ellison & Buckhout, 1981).

The eyewitness's acquired information may be interfered with during the retention stage, or it may be forgotten over the course of time. The classic study of Ebbinghaus (1885/1964) showed that the amount of retrievable verbal information will generally decline with the passage of time. Yet these classic findings have been questioned. A recent concentration of research effort has been directed toward the phenomenon of hypermnesia, the increase of recall over repeated recall trials (Erdelyi & Becker, 1974; Erdelyi, Finkelstein, Herrell, Miller, & Thomas, 1976; Erdelyi & Kleinbard, 1978; Madigan, 1976). The results of these studies raise questions about the inevitability of the Ebbinghaus forgetting curve. The work of Erdelyi and Becker (1974), which has been found to be reliable and replicable, indicates that memory for certain types of images may increase rather than decrease over time. They found a significant improvement in recall of pictures but not words over multiple recall trials. They also reported that the hypermnesic effect was somewhat greater when the subjects engaged in a "think" interval as a means of

improving memory. By a think interval, we mean that a subject is simply told to review the stimulus materials in his mind for a fixed period of time. Cady (1924) and Lipton (1977) found that when a narrative form of recall was requested (free recall), memory for details of the event was less complete but more accurate than the memory elicited by an interrogatory form of probing.

Impressive as the hypermnesia research may be, it has, for the most part, been limited to studying memory for fixed lists of words or pictures. Our focus on dynamic, changing events takes us into an area in which the unsuspecting observer has a difficult acquisition task, a retention interval punctuated by interrogation and variety of retrieval techniques that can influence the process of reconstructing the memory of the event. To determine whether hypermnesia could be observed in such a setting, we created a staged mock crime, varied the type and character of the retention interval, employed repeated open-ended free recall tests, and measured both the number of correct items recalled and the number of memory intrusions. It was initially predicted that a "think" interval would facilitate hypermnesia in eye-witnesses.

METHOD

Subjects

Forty Brooklyn College undergraduates served as subjects, receiving experimental credit in their introductory psychology course for their participation. The sample group's average age was 22.4 years and consisted of 70% males and 30% females.

Apparatus

The subjects were shown 38 2 x 2 slides of a staged assault in a simulated prison setting. The filming was done at a distance of 5 m. The experiment was conducted in a room set up with a stereo tape recorder synchronized to present the slides on a slide projector. The film was displayed on a screen 3.1 m in front of the subject.

In the "play" condition, a Drueke "Shoot the Moon" perception toy was used as a means of blocking rehearsal by the subject.

The film depicts three prison inmates playing cards in what appears to be a prison recreation room. Two of them accuse the winner of cheating, and, after arguing, they get up and leave. Another inmate then enters the scene, holding a glass bottle at his side, and argues with the winner about his having cheated him that morning in a prior game. He then hits the winner (who is still seated) in the face with the bottle and leaves. As the victim is screaming, a guard enters to assist the victim and then leaves to get help.

All of the prison inmates were dressed in white T-shirts and blue jeans. Two of the inmates, the perpetrator and one of the card players, clearly had numbers visible on their backs. The guard was wearing a striped shirt and blue jeans. The table also had many details: cards, cigarettes, black radio, eyeglasses, and a red card box. The sound track was rich in detail, consisting of rock music, arguing, names being mentioned, the sound of a bottle breaking, screaming, and moaning. The total length of the slide presentation was 64 sec. The presentation rate of each slide was 1-1.5 sec, timed to match events on the sound track, which had been recorded simultaneously with the film. The perpetrator's face was clearly visible for 2.4 sec, and the number on his back was visible for 9 sec.

Procedure

Subjects were randomly assigned to either "think" or "play" conditions. Subjects were told that they were to be shown a short film, to put on eyeglasses if required, and to make themselves comfortable.

The experimenter interviewed each subject for 5 min and recorded the conversation on tape. The experimenter was trained to use neutral, nonleading probing questions in asking for more details. An example of some of the questions was "Can you please be specific?", "Please go on," "Do you remember anything else?", "Can you tell me everything that happened?"

Subjects in the "think" condition were then instructed in the following manner: "Thank you very much. Now we are going to ask you to sit and think over what you just saw for a few moments. The purpose will be to see if you can remember any better after thinking for a while. Just imagine that you are looking at a movie screen where the whole scene is playing back. Try to think about specific details. After a few moments, I will return and ask you to make another statement on what you remember. This is just like real police investigations where you have to tell the story over and over again. Do you have any questions?"

Subjects in the "play" condition were told, "We now have a perception game for you. The purpose will be to see if you can improve your performance with practice."

Following each "think" or "play" period, which was timed for 5 min, the experimenter conducted another 5-min interview, which was recorded on tape. The instructions they received were similar to those given initially, depending on the condition to which they belonged. Subjects were to retell what they had witnessed, as well as to add any new details they now remembered. This procedure was repeated in exactly the same way for the third recall session. The subjects were then asked to fill out a simulated police form, which again asked for data about the "crime." The form consisted of questions concerning (1) the subject's demographics, (2) the perpetrator's description, (3) objects they had seen, (4) anything they might have heard, and (5) a specific probe to determine if they could remember any part of the prison number from the back of the perpetrator's T-shirt.

One of the experimenters scored the tape-recorded interviews in a random order of tape segments. He had not participated in the interrogation of the subjects. Each of the taped interviews was scored separately. The experimenter did not know the conditions to which subjects had been assigned, thus minimizing bias on his part. The scoring of recall accuracy was accomplished by means of a standardized checklist composed of facts gleaned from repeated observations of the slides by the experimenters and a large group of pilot subjects. Correct responses (hits) and incorrect responses (intrusions) were counted. The frequency of hits and intrusions constituted the main dependent variables. Based on the observed maximum range of hits and intrusions, a specially derived d' score was computed, for which the a priori probability of a hit was equal to the number observed out of 37, a value determined by counting the maximum total number of unique correct items recalled by all subjects. The maximum total number of unique memory intrusions was 11, so that the a priori probability of an intrusion was the number observed out of 11.

RESULTS AND DISCUSSION

A groups by trials ANOVA (Winer, 1971) of the hit data revealed no significant main effect for type of rehearsal [$F(1,38) = .13$], but we found a significant increase in hits over trials [$F(2,76) = 8.91$, $p < .001$] and no significant interaction. An ANOVA on the intrusion data showed no significant main effect [$F(1,38) = .24$], but recall trial effects [$F(1,38) = 3.75$, $p < .03$] and the interaction [$F(2,76) = 3.79$, $p < .03$] were significant. Interpretation of the interaction of the memory intrusion data with orthogonal polynomials

showed that intrusions increased monotonically over trials in the "think" condition [F(1,38) = 5.72, p < .01], while showing a flat, nonsignificant trend in the "play" condition [F(1,38) = 1.76].

The analysis of the d' data showed that the main effect for type of rehearsal was not significant [F(1,38) = .02], but there was a significant Rehearsal by Trial interaction [F(2,76) = 4.77, p < .01] and a nonsignificant trial effect [F(2,76) = .28]. Reflecting the trends in the hit and intrusion data, witnesses in the "think" condition showed a significant decrease in d' over trials [F(1,38) = 5.57, p < .02], whereas witnesses in the "play" rehearsal condition showed a significant increase in d' over trials [F(1,38) = 4.77, p < .01].

The hypothesis that the "think" rehearsal condition (in which witnesses were asked to review the crime events as if viewing a film of the crime) would facilitate hypermnesia was rejected. In fact, there were no overall differences between "think" witnesses and the "play" group (control), except for the tendency of witnesses in the "think" group to increase memory intrusions and, hence, show a slight decrease in d' (forgetting) over recall trials. The modest hypermnesia shown by witnesses in the "play" condition (blocked rehearsal) must be interpreted cautiously since the mean d' scores were very low, did not increase monotonically, and, in fact, showed a decline on Recall Trial 3 (see Table 1). While hypermnesia for hits was shown for both groups, intrusions increased fast enough to produce a flat d' function, which indicates that overall, hypermnesia did not occur.

Because there were no overall differences between the two rehearsal groups, the data were combined and then divided into two new groups on the basis of an analysis of answers from the witness's interview protocol from Recall Trial 1, which were coded according to whether or not the witness had ascertained the correct schema. The correct schema was that the crime was committed by a new person entering the scene, rather than one of the earlier card players who left. As shown in Table 2, 16 witnesses had the correct schema and 24 blamed the crime on (and tried to describe) one of the earlier, innocent, card players. A groups by trials ANOVA showed that witnesses with the correct schema had a significantly greater number of hits than those with an incorrect schema [F(1,38) = 22.59, p < .0001] and that both groups showed significant increases in hits over trials [F(2,76) = 8.85, p < .0004]. An ANOVA on the

Table 1
Summary of Recall Data Over Rehearsal Conditions

Measure	"Think" Rehearsal Condition			"Play" Blocked Rehearsal Condition		
	1	2	3	1	2	3
Hits	15.00	16.70	19.00	15.80	19.10	19.10
Intrusions	3.50	4.60	5.20	4.20	3.70	4.20
d'	.33	.09	.08	.01	.36	.22

Table 2
Hits, Intrusions, and d' Scores for Witnesses With Correct and Incorrect Schemas

Measure	Correct Schema (n = 16)			Incorrect Schema (n = 24)		
	1	2	3	1	2	3
Hits	19.375	21.563	23.438	12.625	14.875	15.667
Intrusions	3.563	4.250	5.250	4.042	4.208	4.417
d'	.650	.600	.410	−.150	−.030	−.020

intrusion data showed a nonsignificant main effect for schema [F(1,38) = .06] and an overall significant increase in intrusions over trials [F(2,76) = 4.50, p < .01]. There were no significant interactions. The d' data subjected to a group by trials ANOVA showed a highly significant superiority in d' for the correct schema group [F(1,38) = 11.62, p < .002], but the overall trial effect and interaction were not significant. A closer look at the correct schema group showed an almost significant quadratic trend downward in d' [F(1,38) = 3.19, p < .08].

The people we call "good witnesses," with the correct schema or theory of the crime, began with significantly more hits than the "bad witnesses" on the first recall trial and continued to improve, whereas those witnesses on the wrong track began with fewer hits and improved on the second recall trial but failed to improve significantly on the third recall trial. The two groups did not differ in the number of memory intrusions on the first two trials, but the correct-schema witnesses showed a significant increase in intrusions on Recall Trial 3 [t(15) = −2.60, p < .02]. These data suggest that when otherwise good witnesses are pushed too hard by repeated interrogation, having already recalled a large number of correct details, they may start constructing details to fit the social demands placed on them to talk. Presumably, the good witnesses had reached some kind of ceiling effect but felt obliged to keep talking, no matter what they said. The bad witnesses showed an increased number of hits as well, but they never even reached the starting level of hits achieved by the good witnesses. These increases in hits could suggest an illusory showing of hypermnesia to a criminal investigator, who might not be able to verify a statement as either a hit or an intrusion. Clearly, our witnesses provided more information over recall trials, but it was a mixture of fact and fiction. Few subjects were able to recall the "prison" numbers on the back of the actor's T-shirt, a single item of memory that would have provided the best possible identification of a perpetrator inside a prison population.

We note that, consistent with much of the current research on the eyewitnesses to crime by Buckhout (1974), E. F. Loftus (1979), and others, a close look at the d' scores on Recall 1 (immediately following the crime) reveals that most of the eyewitnesses were not showing what most memory researchers would regard as mastery of the material. Either the witnesses had not

encoded enough visual or verbal material to satisfy the recall demands placed on them by the investigator or they were affected in some unknown way by being asked questions. We tend to support the notion that the encoding was not very complete to begin with, because of the complexity of the dynamic stimulus. Further, we believe that the demonstration that many of the witnesses got off on the wrong track is evidence that the encoding process can produce distortions of crime events right from the start.

There is a parallel in our research to a study by Owens, Bower, and Black (1979). In an ingenious series of experiments, these researchers tested the recall of subjects exposed to simple (dull) narratives that differed in whether an implied motive was provided for the action's main character. The motive (the woman was described as pregnant) was hypothesized to be a schema for deciding the meaning of actions described in the narratives. The motive and control groups did not differ in the number of text propositions correctly recalled, but the group with a motive produced a significantly greater number of memory intrusions. In free recall, subjects who had a motive for the character were able to remember more of the episodes correctly. However, when the "motive" subjects were pressed to recognize correct statements from a list of propositions about the story that included both true and false but plausible statements, the subjects with a motive produced more false alarms or intrusions. Owens et al. (1979, p. 191) concluded that "recall will tend to reconstruct the motives, thoughts and attributions the reader imagined that the characters were experiencing." Better still is their title: "The Soap Opera Effect in Story Recall."

To summarize, our witnesses fell naturally into two groups: those with the correct "schema" and those on the wrong track. We might say that our good witnesses picked up the soap opera story line quickly whereas our bad witnesses behaved as if the soap opera lacked any integrating focus. Neither group showed hypermnesia; in fact, the better witnesses tended to forget a little. The particular form of rehearsal called "thinking" by Erdelyi and Becker (1974) showed some signs of being detrimental to memory, even though we are exploring this paradigm in connection with a study we are doing on the use of hypnosis as a memory enhancement technique (Buckhout, Eugenio, & Lecitra, Note 1).

Hypermnesia in eyewitnesses may be merely an artifact resulting from the impoverished amount of information originally stored in memory. What hypermnesia did occur in our study was evident only in hits. The more sensitive d' measure indicated that the witness's memory did not change over time. The free recall of a real episode is very different from the finite pool of pictures and words used by most hypermnesia researchers, in which, in effect, the schema is well understood by the subjects from the instructions. An eyewitness is expected not only to learn and remember facts and features, but also to grasp the schema of what happened after one brief exposure to the "stimulus." In our

research, witnesses who fail to grasp the schema produced d' scores at around the guessing level. Most witnesses gave more information over time, but the good witnesses who had the right schema gave better information to start with, making intrusion errors only when they apparently ran out of accurate things to say. The bad witnesses, already on the wrong track, showed a small increase in memory for correct details but started out so poorly that they had no place to go but up. We presume that some of this improvement may be the product of inference or suggestions inherent in almost any interrogation. While they may sound better over time, bad witnesses remain wrong witnesses.

REFERENCE NOTE

1. Buckhout, R., Eugenio, P., & Lecitra, T. *Is there life after hypnosis? Attempts to revive eyewitness' memory.* Paper presented to the annual meeting of the American Psychological Association, Los Angeles, August 25, 1981.

REFERENCES

BUCKHOUT, R. Eyewitness testimony. *Scientific American*, 1974, 231, 23-31.

CADY, H. M. On the psychology of testimony. *American Journal of Psychology*, 1924, 35, 110-112.

CROWDER, R. F. *Principles of learning and memory.* Hillsdale, N.J: Erlbaum, 1976.

EBBINGHAUS, H. *Memory.* New York: Dover, 1964. (Originally published, 1885.)

ELLISON, K. W., & BUCKHOUT, R. *Psychology and criminal justice.* New York: Harper & Row, 1981.

ERDELYI, M. H., & BECKER, J. Hypermnesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology*, 1974, 6, 159-171.

ERDELYI, M. H., FINKELSTEIN, S., HERRELL, N., MILLER, B., & THOMAS, J. Coding modality vs. input modality in hypermnesia: Is a rose a rose a rose? *Cognition*, 1976, 4, 311-319.

ERDELYI, M. H., & KLEINBARD, J. Has Ebbinghaus decayed with time? The growth of recall (hypermnesia) over days. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 275-289.

HARRIS, R. J. Answering questions, containing marked and unmarked adjectives and adverbs. *Journal of Experimental Psychology*, 1973, 97, 399-401.

LIPTON, J. P. On the psychology of eyewitness testimony. *Journal of Applied Psychology*, 1977, 62, 90-93.

LOFTUS, E. F. *Eyewitness testimony.* Cambridge, Mass: Harvard University Press, 1979.

LOFTUS, G. R., & LOFTUS, E. F. *Human memory: The processing of information.* Hillsdale, N.J: Erlbaum, 1976.

MADIGAN, S. Reminiscence and item recovery in free recall. *Memory & Cognition*, 1976, 4, 233-236.

MUNSTERBERG, H. *On the witness stand.* New York: Doubleday, Page, 1908.

OWENS, J., BOWER, G. H., & BLACK, G. H. The soap opera effect in story recall. *Memory & Cognition*, 1979, 7, 185-191.

TULVING, E. The effects of presentation and recall in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 175-184.

WINER, B. J. *Statistical principles in experimental design,* New York: McGraw-Hill, 1971.