

Retrospective duration estimates of an abductor's speech

A. DANIEL YARMEY and EVA MATTHYS
University of Guelph, Guelph, Ontario, Canada

Two hundred ten male and female introductory psychology students served as subjects in a simulated abduction study purportedly investigating voice identification. The subjects listened to a 72-sec tape recording of either a male or a female abductor's voice. The tape was played for one massed trial or was given in two distributed exposure trials of 36 sec each or three distributed exposure trials of 24 sec each. Each of the distributed exposure trials was separated by a 5-min intertrial interval. The subjects were asked to estimate the total amount of time the abductor's voice was heard and the duration of the intertrial interval immediately, 1 day, or 1 week after presentation. The results indicated that total duration was significantly overestimated, especially by the subjects given repeated exposures to the target voice. Intertrial intervals were either accurately estimated or slightly underestimated. No significant main effects were found for delay of testing, sex of subjects, or sex of abductor. The results are discussed in terms of their forensic relevance.

Information regarding the estimation of time duration can have considerable forensic significance. The U.S. Supreme Court has taken the position that the greater the opportunity a witness has to become familiar with a suspect, the less the chance of misidentification (*Neil v. Biggers*, 1972). Witness credibility may be enhanced when witnesses report having observed a suspect at the time of the crime for a relatively long period of time. A critical issue of importance, of course, is the general accuracy of witnesses' judgments of duration.

Loftus, Schooler, Boone, and Kline (1987) described a 1974 homicide case in which accuracy of duration estimation was critical in the defense of a woman charged with first-degree murder, but who claimed self-defense. Following a bitter argument between the defendant and her boyfriend, the woman ran to her bedroom, grabbed a gun, and shot the victim six times. The accused and her sister stated that 2 sec elapsed between the time the accused grabbed the gun and fired the first bullet. In contrast, another witness testified that 5 min had elapsed. Another "real-world" example of the court's interest in duration judgments involved the first author of this paper, who served as an expert witness for the prosecution on issues related to identification evidence. The police and the crown attorney were interested, among other issues, in the accuracy of duration judgments of two male witnesses who were held hostage by a criminal who wore a ski mask. Throughout the 2-h incident, the criminal repeatedly talked to the victims, searched the house, and

returned to talk to the victims. During his absence, the two victims reminded each other about the importance of remembering his voice. The law was interested in the probability of accuracy judgments about the total time the victims had an opportunity to listen to the voice of the perpetrator, as well as the amount of time that passed while he left to search the house.

The present study is relevant to the two above anecdotes. Subjects participated in a simulated abduction designed to test the accuracy of voice identification. As part of this research, we examined the effects of sex differences on the accuracy of time estimates as a function of massed versus distributed exposures to the target's voice, delay in testing, and sex of the abductor. Subjects were tested for their accuracy in estimating the total duration of hearing the target voice (72 sec) as well as the duration of the intertrial interval (300 sec) when the abductor was not present.

Several studies examining retrospective time estimation have shown that judgments are dependent upon many variables, such as the complexity of the event and interest value of the information contained in the event (Brown, 1985; Fraise, 1984; Ornstein, 1969; Schiffman & Bobko, 1974). The "storage-size hypothesis" provides one explanation for this phenomenon. According to Ornstein (1969), memory of an event's duration depends upon the amount of information stored about that event. If more stimuli occurred during a time period, or if the stimuli were coded in a more complex way and greater storage space utilized, the judged experience of time would be longer. Therefore, complex or highly motivating circumstances, such as attending to a perpetrator's voice, should be estimated as being longer than less complicated or less motivating circumstances, such as a casual conversation with an experimenter during an intertrial interval.

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. We gratefully acknowledge the assistance of Claudia Turowski for testing subjects. Correspondence may be addressed to A. D. Yarney, Department of Psychology, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

The evidence to date for sex differences in time-estimation studies has been equivocal. Some researchers using traditional laboratory procedures have reported finding sex differences (Fraisse, 1963; Geer, Platt, & Singer, 1964; Rammsayer & Lustnauer, 1989), whereas Roeklein (1972) and Smith (1969) reported no statistically significant differences. In an applied experiment in which subjects saw a short videotape of a bank robbery, Loftus et al. (1987) found that duration was consistently overestimated. Overestimations by females were significantly higher than by males, and a more stressful version of the robbery produced higher estimates than did a low-stress version.

The gender of the speaker has not been investigated in retrospective time-estimation studies, particularly in an applied context of a forensic nature. No specific predictions were made regarding the influence of the sex of the speaker on time estimates.

The interval between a crime or accident and police questioning of witnesses can range from several minutes to a week or more. Ornstein (1969) found that a 2-week delay in testing produced significantly shorter estimates of duration than did an immediate test. In the present study, we examined three delay periods: an immediate test, a 24-h delayed test, and a 1-week delayed test. It was expected that judgment of duration would be most accurate on an immediate test, and accuracy of estimates would decline over time.

Finally, the effects on time estimates of a single massed exposure to a perpetrator's voice compared to distributed exposures were examined. Verbal learning research indicates that distributed practice with words as stimuli produces higher recall and superior recognition memory than does massed practice. This difference is attributed to differential rehearsal effects with distributed practice (Underwood, 1969). Similarly, Goldstein and Chance (1985), using a 2-week retention interval, found that speaker identification is enhanced if subjects listen to short samples of a target's voice distributed over 3 consecutive days in contrast to listening to the total speech sample in one observation trial. It was hypothesized in the present study that more information would be coded about a speaker's voice and more storage space would be utilized under distributed exposure than under massed exposure. Consequently, it was predicted that subjects would estimate longer experiences of total time after hearing the abductor's voice with distributed exposures of two periods of 36 sec each, or three periods of 24 sec each, as opposed to a single period of 72 sec.

METHOD

Subjects and Design

A total of 210 male and female undergraduates participated as subjects for extra course credit. The project had been advertised as involving a voice-identification task. The experiment was a $3 \times 3 \times 2 \times 2$ between-subjects groups factorial. The factors were exposure (massed exposure of 72 sec, distributed exposure of two trials of 36 sec each, and distributed exposure of three trials of 24 sec each), delay of test (immediate test, 24-h delayed test, and 1-week delayed test), sex of abductor's voice (male, female), and sex of subjects (male, female).

Procedure

The subjects were asked to assume that they were the daughter/son of a wealthy businessman and that they had just been kidnapped by an abductor who wore a mask throughout the incident. They were told that they would hear a tape recording of the voice of the abductor and that they were to be tested later on their ability to identify the speaker in a voice lineup. They were not forewarned to expect a test for time estimates. After any questions were answered, the tape-recorded monologue of the abductor giving instructions and voicing complaints was played for the appropriate duration. During the 5-min intertrial intervals in the distributed exposure conditions, the subjects were told to assume that the abductor had left the room and would return later. The subjects and the experimenter engaged in informal conversation during the intertrial period. The first question given to the subjects in the test stage was to estimate the total amount of time they had heard the abductor's voice. The subjects in the distributed exposure conditions were then asked to estimate the duration spent in conversation with the experimenter during the intertrial interval(s).

In order to minimize stimulus sampling errors, 2 male volunteers aged 25 and 26 years and 2 female volunteers aged 24 and 25 years participated as abductors. The abductors spoke in monotone without any noticeable accents or other distinctive voice characteristics. Their voices were recorded on a Sony tape recorder with an external microphone with high-quality Maxell tapes.

RESULTS

The subjects generally answered the estimation questions with rounded approximations in minutes and/or seconds. All estimates were converted into seconds for statistical analysis. Estimates of the total speech sample of 72 sec ranged from 30 to 960 sec, with a mean estimate of 312 sec. Ninety-eight percent of all subjects overestimated the actual duration. Table 1 shows the means and standard deviations of estimates across treatment conditions.

A $3 \times 3 \times 2 \times 2$ analysis of variance (ANOVA) showed a significant main effect for exposures [$F(2,174) = 8.85, p < .0002$]. Reliably longer estimates were given to the two distributed exposure conditions than to the massed exposure condition ($p < .01$, Newman-Keuls), but there was no significant difference between the two distributed conditions. The only other significant

Table 1
Means and Standard Deviations of Time Estimates (in sec) for the Total 72-sec Duration

	Exposure			Delay			Sex of Voice		Sex of Subject	
	Massed (72 sec)	Distributed (2 × 36 sec)	Distributed (3 × 24 sec)	Immediate	24 h	1 week	Male	Female	Male	Female
	<i>M</i>	235	362	341	278	346	315	296	328	331
<i>SD</i>	129	184	225	173	197	197	166	212	204	173

effect was the interaction between delay of testing and sex of the perpetrator's voice [$F(2,174) = 3.00, p < .05$]. Estimates of duration for the male voices increased after a 1-day delay in testing [$F(2,103) = 9.88, p < .01$], whereas the subjects tested after a 1-week delay with male voices did not differ from those tested immediately. Female abductor voices elicited significantly longer duration estimates than did male abductor voices over the three test conditions [$F(1,69) = 6.97, p < .01$]; however, estimates of duration of the speech samples of the female abductors' voice did not increase over time.

Table 2 shows the estimates of duration for the single intertrial interval in the 2×36 sec distributed exposure condition and the first 300-sec intertrial interval in the 3×24 sec distributed condition. The average estimate for all subjects in the distributed exposure conditions was 277 sec, with 51% of the subjects underestimating the duration. Thirty-two percent were correct in their estimation of 5 min. An ANOVA of these estimates revealed a significant interaction between exposure and sex of the perpetrator's voice [$F(1,114) = 4.20, p < .05$]. The subjects who heard the male abductors two or three times did not differ in their underestimations of the intertrial duration, but the subjects who heard the female abductors two times overestimated, and the subjects who heard the female abductors' voice three times underestimated, the intertrial interval [$F(1,67) = 7.42, p < .01$]. An ANOVA of the estimates of duration for the first versus the second intertrial interval for the subjects who heard the abductor's voice for three sample periods of 24 sec each revealed no significant differences.

DISCUSSION

The present experiment was designed to simulate a complex forensic situation involving an abductor and his or her victim earwitness. Retrospective time estimates of the total duration of a speaker's voice were greatly overestimated. Exposure to an abductor's voice for a total duration of 72 sec was estimated to last an average of 312 sec, which is a ratio approximating a factor of overestimation of 4 to 1. As predicted, repeated short exposures to a voice produced significantly longer estimates of total duration than did a single long exposure. However, estimates of a 5-min intertrial interval, which was spent in casual conversation with the experimenter, were either accurately estimated or slightly underestimated in duration. No statistically significant differences in time estimates were found between male and female subjects or between voices of male and female abductors. A significant interaction between delay of testing and sex of the abductor's voice was found, but no interpretation for this effect can be offered. Similarly, no meaningful interpretation can be made of the significant interaction between exposures to the

target voice and the sex of the abductors' voice on estimations of the intertrial interval.

It is probable that the subjects carefully followed the experimenter's instructions to attend to the voice of the abductor since they knew that they were participating in a voice-identification study. As a consequence of these instructions and the complexity of the task, selective attention would have been high while processing voice characteristics and low while processing temporal cues. We found that retrospective time was substantially overestimated, with 98% of all subjects overestimating the total duration—particularly in the distributed exposure conditions. The subjects in these conditions may have made longer estimates because of differential rehearsal of voice characteristics each time they heard the abductor's voice (see Elmes, Greener, & Wilkinson, 1972). Each rehearsal may have taken more time (see Shaughnessy, Zimmerman, & Underwood, 1972) and utilized more storage space (Ornstein, 1969) in the distributed conditions than in the massed exposure condition, which may have led to the greater retrospective estimations of total time duration. In contrast, the relatively accurate estimates of the 5-min intertrial interval (32% were accurate and 51% slightly underestimated the interval) may have occurred because the subjects gave a "ballpark" estimate of 5-min duration. The use of a 5-min intertrial interval may not have been a wise decision by the experimenters and is not recommended for any subsequent investigations. Another major difference in the two types of time estimates was the fact that in attending to the abductor's voice, the subject was an active listener (deep levels of processing for identification purposes), whereas during the intertrial interval, the subject was engaged in a conversation with the experimenter. These procedural and processing differences may account for the respective over- and underestimations.

Contrary to predictions and the findings of Ornstein (1969), delay in testing over a 1-week interval did not affect the retrospective estimates of duration. These results support the findings of Yarmey (in press), who found no significant differences in retrospective duration estimates of 20- and 40-min 2-person discussions on an immediate test and a 1-week delayed test.

It is concluded that witnesses are likely to make significant overestimations of time after listening to the voice of a perpetrator for a relatively short period. The results of this study, along with the findings of Loftus et al. (1987) and others (Buckhout, 1977; Johnson & Scott, 1976; Marshall, 1966), suggest that time estimates in forensic situations should be accepted with great caution.

REFERENCES

- BROWN, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, *38*, 115-124.
- BUCKHOUT, R. (1977). Eyewitness identification and psychology in the courtroom. *Criminal Defense*, *4*, 5-10.
- ELMES, D. G., GREENER, W. J., & WILKINSON, W. C. (1972). Free recall of items presented after massed- and distributed-practice items. *American Journal of Psychology*, *85*, 237-240.
- FRAISSE, P. (1963). *Psychology of time*. New York: Harper & Row.
- FRAISSE, P. (1984). Perception and estimation of time. In *Annual Review of Psychology* (pp. 1-36). Palo Alto, CA: Annual Reviews.
- GEER, J., PLATT, P., & SINGER, M. (1964). A sex difference in time estimation. *Perceptual & Motor Skills*, *19*, 42.

Table 2
Means and Standard Deviations of Duration Estimates (in sec)
for the First 5-min Intertrial Interval

	Exposure		Delay		Sex of Voice		Sex of Subject	
	Distributed (2 × 36 sec)	Distributed (3 × 24 sec)	Immediate	24 h	Male	Female	Male	Female
<i>M</i>	294	111	161	192	180	171	195	155
<i>SD</i>	131	227	184	209	219	215	210	233

Note—There was only one intertrial interval in the 2×36 sec condition. Data are presented for the first of two intertrial intervals in the 3×24 sec condition.

- GOLDSTEIN, A. G., & CHANCE, J. E. (1985). *Voice recognition: The effects of faces, temporal distribution of "practice," and social distance*. Paper presented at the Midwestern Psychology Association meeting, Chicago.
- JOHNSON, C., & SCOTT, B. (1976). *Eyewitness testimony and suspect identification as a function of arousal, sex of witness and scheduling of interrogation*. Paper presented at the American Psychological Association meeting, Washington, DC.
- LOFTUS, E. F., SCHOOLER, J. W., BOONE, S. M., & KLINE, D. (1987). Time went by so slowly: Overestimation of event duration by males and females. *Journal of Applied Cognitive Psychology*, *1*, 3-15.
- MARSHALL, J. (1966). *Law and psychology in conflict*. New York: Bobbs-Merrill.
- NEIL v. BIGGERS, 409 U.S. 188 (1972).
- ORNSTEIN, R. E. (1969). *On the perception of time*. New York: Penguin.
- RAMMSAYER, T., & LUSTNAUER, S. (1989). Sex differences in time perception. *Perceptual & Motor Skills*, *68*, 195-198.
- ROECKELEIN, J. E. (1972). Sex differences in time estimation. *Perceptual & Motor Skills*, *35*, 859-862.
- SCHIFFMAN, H. R., & BOBKO, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology*, *103*, 156-159.
- SHAUGHNESSY, J. J., ZIMMERMAN, J., & UNDERWOOD, B. J. (1972). Further evidence on the MP-DP effect in free-recall learning. *Journal of Verbal Learning & Verbal Behavior*, *11*, 1-12.
- SMITH, N. (1969). The effect on time estimation of increasing the complexity of a cognitive task. *Journal of General Psychology*, *81*, 231-235.
- UNDERWOOD, B. J. (1969). Some correlates of item repetition in free-recall learning. *Journal of Verbal Learning & Verbal Behavior*, *8*, 83-94.
- YARMEY, A. D. (in press). Accuracy and confidence of duration estimates following questions containing marked and unmarked modifiers. *Journal of Applied Social Psychology*.

(Manuscript received September 25, 1989.)