1993

When Eyewitnesses Are Also Earwitnesses: Effects on Visual and Voice Identifications

Hunter A. McAllister
Southeastern Louisiana University

Robert H.I. Dale
Butler University, rdale@butler.edu

Norman J. Bregman
Virginia Commonwealth University

Allyssa McCabe
University of Massachusetts - Lowell

C. Randy Cotton
University of Southern Mississippi

When Eyewitnesses Are Also Earwitnesses: Effects on Visual and Voice Identifications

Hunter A. McAllister, Robert H. I. Dale, Norman J. Bregman, Allyssa McCabe, C. Randy Cotton

Abstract

In Experiment 1, subjects witnessed a mock crime either visually or both auditorily and visually. A visual lineup was conducted with either a guilty or an innocent suspect present. Identification accuracy of visual-only versus auditory-visual witnessed did not differ, although the diagnosticity ratio for the visual-only condition was more than twice as large. Thus, there was only limited support for auditory information interfering with encoding visual information. In Experiment 2, subjects witnessed a mock crime either auditorily or both auditorily and visually. A voice lineup was conducted with either a guilty or an innocent suspect present. Consistent with Yarmey's (1986) prediction that visual information can interfere with encoding auditory information, guilty-suspect identification was significantly higher in the auditory-only condition.

In observing a crime, witnesses often have both visual and auditory information. Although there has been much research on both eyewitnesses (see Loftus, 1979; Yarmey, 1979) and earwitnesses (see Bull & Clifford, 1984), surprisingly little research deals with both the visual and auditory modalities in the same design.

In one of the few studies concerned with both witnessing modalities, Yarmey (1986) addressed the encoding question of whether auditory information might be attended to more carefully when visual information was absent or weak. To test the influence of visual information on the processing of auditory information, Yarmey (1986) had subjects witness a simulated rape; subjects saw the rapist and heard his voice. The main experimental manipulation was the degree of illumination under which the rape was observed. Yarmey argued that if witnessed behaved logically, as the visual impairment increased, subjects would give greater attention to the auditory information. Contrary to this prediction, Yarmey found that subjects did not allocate greater attention to or have better memory for voice as the visual conditions became worse. From this research, then, it appears that the two sensory channels operate independently and that little has been lost in studying them separately. Before coming to such a conclusion, however, it would be useful to consider some of the possible limitations of Yarmey's study.

Yarmey (1986) suggested that the visual conditions may not have influenced the voice identification accuracy because of the generally poor performance on both visual and voice identification. Yarmey pointed out that the overall poor performance in both the visual and voice lineups might have created floor effects that minimized the influence of the illumination manipulation. Before ruling out the influence of visual information on the accuracy of voice identification, research should be conducted with visual and voice information that has greater impact than that used by Yarmey.

Although not mentioned by Yarmey, a further limitation concerns the absence of an auditory-only (no visual information) witnessing condition. The failure to find differences between daylight and night conditions does not necessarily mean that visual information does not interfere with the processing of auditory information. It might mean that any visual information, no matter how weak, disrupts the processing of auditory information. Comparing the voice identification performance of an auditory-visual witnessing condition with an auditory-only control condition would provide the strongest test of visual interference with the processing of auditory information.

One of the main purposes of the current research was to test Yarmey's original hypothesis through the use of stronger auditory and visual stimuli and the inclusion of an auditory-only control condition. An additional purpose was to extend Yarmey's predictions concerning voice lineups with visual lineups. Just as visual information was hypothesized to interfere with the processing of auditory information, it might also be that auditory information could interfere with the processing of visual information. In a recent study, Yarmey (1992) showed that voice provides very little information about an individual's physical characteristics. Because there is little crossover of information from one modality to the other, to the extent that voice was focused on, it could distract from the processing of visual cues. Thus, witnesses who both saw and heard a criminal would be predicted to be less accurate in visual lineups than those who had only seen the criminal. Yarmey (1986) was unable to test such as question because all his witnessing conditions involved auditory stimuli being present.

Experiment 1

Experiment 1 was conducted to (a) demonstrate the strength of the visual witnessing stimuli and (b) test the extension of Yarmey's interference hypothesis. This was accomplished by manipulating the witnessing situation so that subjects either received both auditory and visual information or received only visual information; subjects were then tested in a visual lineup.

Method

Subjects. A total of 144 introductory psychology students (72 women, 72 men) volunteered as subjects. Subjects received course credit for participation.

Design. The design was a 2 (Suspect in Lineup: Guilty vs. Innocent) × 2 (Witnessing Modality: Auditory-Visual vs. Visual Only) factorial. Subjects were randomly assigned to one of the four experimental conditions.

Apparatus and materials. The visual and auditory stimuli used in the original witnessing conditions and in the lineups were constructed from a pool of 48 White male college students who had volunteered to be photographed and (audio) tape-recorded. Each subject was photographed twice: (a) outdoors with a brick-wall background, and (b) inside a large room with

a blank-beige-wall background. The outside photographs were full body with the men wearing typical casual clothing. The inside photographs were of head and shoulders with all subjects wearing the same white laboratory coat so that no other clothes were visible in the photograph. All photographs were 9 mm × 13 mm and in color. The man selected as the perpetrator for the simulated crime: (a) was clean shaven, (b) had no unusual physical characteristics, (c) had no unusual voice characteristics, and (d) had given a convincing crime monologue reading. In a pilot study, six photographs of other subjects were selected based on three criteria: (a) the similarity of their inside photographs to the inside photograph of the target subject, (b) their having no unusual voice characteristics, and (c) their having given a convincing crime monologue reading. The individual with the greatest visual similarity to the target was assigned the role of *lure* to be used in target-absent lineups in place of the target. The five next highest ranking individuals were chosen as distractors for the six-person lineups.

The voice stimulus used in the experiment was a cassette tape-recording of a 60-sec, 233-word monologue describing a conspiracy to commit murder. The monologue was by a husband hiring someone to kill his wife. The husband went into various details concerning the conspiracy such as when his wife would be alone at home, how to get into the house, and so forth. The hired killer was never heard.

Procedure. Subjects were led to an experimental cubicle by the experimenter. Instructions were delivered via audiotape recordings. Subjects were informed that they would be exposed to information concerning a mock crime and then asked questions concerning the crime. Subjects in the *auditory-visual* witnessing condition were given the outside picture of the perpetrator and told to examine it while listening to his voice. Subjects in the *visual-only* condition were simply given the picture to examine. After a 1-min inspection period, subjects performed a 5-min filler task.

Following the filler task, subjects were given a notebook containing 12 head-and-shoulder pictures, each on a separate page. Subjects were told that they would be looking at a 12-person lineup, 1 person at a time. Subjects were told that, just as in the real world, the perpetrator might or might not be in the lineup. When the experimenter rang a bell, subjects turned to the first picture. After 10 sec, the bell rang again; subjects recorded their judgment as to whether the person was the perpetrator or not and how confident they were of their answer. This procedure was continued until 6 lineup members had been seen. Following the 6th stimulus presentation, the experimenter explained that there would be no further stimuli to evaluate. This procedure of letting the subjects believe that the lineup would involve 12 people while actually stopping at 6 was developed by Lindsay and Wells (1985) to reduce any tendency to increase the probability of making a *yes* response as the end of the lineup approached.

Half the lineups contained the face of the target (perpetrator) whereas the other half contained the face of the lure (innocent suspect) in place of the target. The distractor stimulus material was the same for target and lure lineups. Order was counterbalanced such that the target (and lure) occurred equally often in each of the first six lineup positions.

Table 1. Proportion of visual identifications of the suspect as a function of Suspect in Lineup and Witnessing Condition

Group	Suspect in Lineup	
	Guilty	Innocent
Witnessing condition		
Auditory-visual	.64	.11
Visual	.75	.06

Results

Table 1 shows the mean proportions of visual identifications of the criminal and innocent suspects as a function of witnessing conditions. Chi-square tests revealed no significant differences in suspect identifications between the visual-only and the auditory-visual conditions for either the guilty-suspect lineups or the innocent-suspect lineups. When the suspect manipulation was tested, it was found that significantly more suspect identifications were made in the guilty-suspect lineups than in the innocent-suspect lineups, $\chi^2(1, N = 144) = 56.57$, $p < .05$.

Another method of comparing the usefulness of identifications from the two witnessing conditions is to compute their *diagnosticity* (Wells & Lindsay, 1980). Diagnosticity is the ratio of correct identifications of the suspect in the guilty-suspect lineups to false identifications of the suspect in the innocent-suspect lineups. Visual-only witnessing produced a diagnosticity ratio of 12.50 as compared with the auditory-visual witnessing ratio of 5.82. Thus, the identifications of visual-only witnesses were more than twice as diagnostic as identifications by auditory-visual witnesses.

Discussion

The first goal of Experiment 1 was to demonstrate the strength of the visual stimulus material. The identification rate for the guilty suspect was significantly higher than for the innocent suspect. Further, as can be seen in Table 1, the recognition rates for the criminal suspect averaged .69—more than twice the .33 rate obtained in Yarmey's (1986) daylight condition. Thus, the first goal, demonstrating the strength of visual stimuli, was achieved.

The second purpose of Experiment 1 was to test the extension of Yarmey's interference hypothesis to auditory information interfering with the processing of visual information. Contrary to the hypothesis, there were no significant differences in visual-lineup identification rates as a function of witnessing modality. Consistent with the hypothesis, however, the diagnosticity index of an identification from visual-only witnesses was more than twice that of the daignosticity of an identification by an auditory-visual witness. Thus, the findings for whether voice interferes with visual processing were not as straightforward as the findings for the first issue of the effectiveness of the visual stimuli.

Experiment 2

With Experiment 1 demonstrating the effectiveness of the visual stimuli, it was now possible to conduct a second experiment to provide a strong test of Yarmey's original hypothesis. In Experiment 2, the witnessing situation was manipulated so that subjects received either auditory-visual information or auditory information alone; subjects were then tested in an auditory lineup.

Method

Subjects. A total of 144 introductory psychology students (72 women, 72 men) volunteered as subjects. They received course credit for participation.

Design. The design was a 2 (Suspect in Lineup: Guilty vs. Innocent) × 2 (Witnessing Modality: Auditory-Visual vs. Auditory-Only) factorial. Subjects were randomly assigned to one of the four experimental conditions.

Apparatus and materials. The auditory and visual stimuli used in the original witnessing conditions were identical to those used in Experiment 1. The same guilty suspect, innocent suspect, and foils used in Experiment 1 were also used; however, their voices rather than their faces were used in the lineup. The voice stimuli used in the lineups were 10-sec recordings in which some of the key lines from the original monologue were repeated. Every lineup member (foils, lure, and target) made 1 of the 10-sec tapes for use in the voice lineups. The 10-sec monologue was 33 words in length.

Procedure. The procedure for Experiment 2 was identical to that of Experiment 1 with two major exceptions. The first difference concerned the original witnessing conditions. The auditory-visual witnessing condition (as compared with the visual-only condition of Experiment 1) subjects simply listened to the target's 60-sec monologue. The other difference in procedure was that a voice lineup, rather than the visual lineup of Experiment 1, was used. The basic lineup procedure and instructions were identical to Experiment 1 with the exception that following the bell indicating the next lineup member, subjects heard the 10-sec tape for a lineup member.

Results

Table 2 shows the mean proportions of voice identifications of the guilty and innocent suspects as a function of witnessing conditions. Identification of the guilty suspect was significantly more likely in the auditory-only condition than in the auditory-visual condition, $\chi^2(1, N = 72) = 4.51$, $p < .05$. There was no significant difference between the auditory-only and auditory-visual conditions in (mis)identifications of the innocent suspect.

As with the visual lineup, the diagnosticity ratios were computed for both witnessing conditions. The diagnosticity ratio for the auditory-only condition was 10.9 whereas the ratio for the auditory-visual condition was 5.9.

Table 2. Proportion of voice identifications of the suspect as a function of Suspect in Lineup and Witnessing Condition

Group	Suspect in Lineup	
	Guilty	Innocent
Witnessing condition		
Auditory-visual	.33	.06
Auditory	.61	.06

Discussion

One of the major purposes of this experiment was to test a prediction from Yarmey (1986) that auditory information would be attended to more carefully when visual information was absent during the witnessing of the crime. This was tested by comparing the voice-lineup accuracy of the condition in which the subjects had both seen and heard the criminal with the condition where they had only heard him. As predicted, subjects' ability to identify the criminal suspect in a voice lineup was better when subjects had only auditory information available during witnessing; no differences occurred in (mis)identifications of the innocent suspect. Additional support came from the diagnosticity ratios, which revealed that auditory-only witnessing resulted in identifications that were more than 1.8 times as diagnostic as auditory-visual witnessing identifications. Thus, the methodological differences from Yarmey (1986) did produce a situation in which the presence of visual information during the witnessing of a crime interfered with the processing of auditory information.

In addition to the usefulness of the diagnosticity ratios in testing the experimental hypotheses, the overall levels of the ratios are worthy of note. Even in the auditory-visual witnessing condition, which had the lower diagnosticity ratio, that ratio was still quite high suggesting that a voice identification is a valuable piece of information in determining suspect guilt. Part of the reason for the high diagnosticity ratios was the small number of false-positive identifications of the innocent suspects. This finding of few false positives is at odds with past research which has found that voice identifications are unusually high in false-positive identifications (e.g., Hollien, Bennett, & Gelfer, 1983; Yarmey & Matthys, 1992).

One possible reason for the relatively small number of false positives in our study concerns the way in which the innocent suspect and foils were selected. Although the voices of the foils and lure were generally similar to the target in that they had no unusual voice characteristics and had given a convincing reading, they were not selected for their similarity to the target. The foils and lure were selected because of their *visual* resemblance to the target. This is in contrast to the general practice for research on voice recognition which is to select lineup members based on voice similarity to the target. The result of matching on voice is of course to increase the similarity of the innocent suspect and foils to the target. And, as discussed by Navon (1990) and Wells and Luus (1990), the greater the resemblance of the innocent suspect to the guilty suspect, the lower the diagnosticity ratio is likely to be due to the increases in false positives. Thus, the

difference in false positives between our research and past research may be a function of the method of selecting lineup members.

Which technique for selecting lineup members is most appropriate? It depends on the situation to which the researcher wishes to generalize. According to Wells (1988), there are two different situations in which voice identifications are made: (a) situations in which the witness both visually viewed the perpetrator and heard the perpetrators voice, and (b) situations in which the witness was exposed to the perpetrator's voice, but had little or no visual contact with the perpetrator. When the perpetrator was both seen and heard, then the voice identification will usually be conducted as part of a live lineup. Wells's (1988) recommendations concerning how to select foils for a live lineup focus almost exclusively on matching for visual characteristics. Voice is relegated to a brief discussion of "compelled actions"; there is no real concern about matching for voice. However, when the witness only heard the perpetrator, then a voice lineup is used. For these voice lineups, Wells (1988) suggested that foil voices should be matched for the general characteristics of the perpetrator's voice. Thus, the issue of whether to match on voices or not depends on whether the original witnessing involved just voice alone or involved both voice and visual cues.

Our study employed both auditory-only and auditory-visual witnessing situations. To test the interference hypothesis, a voice lineup was required; however, the standard procedure for constructing a voice lineup was not really appropriate. Because the ultimate goal of our research program is to explore the relation between auditory and visual information cues when they occur together at both acquisition and retrieval, we followed the live lineup procedure of selecting foils based on their visual similarity to the target. This procedure makes our research compatible with a second line of research on the contribute that voice makes to the auditory-visual lineup. Thus, our lineup procedures facilitate generalizations to situations where both auditory and visual cues are present, but may limit their usefulness for pure voice lineups.

General Discussion

The general conclusion to be drawn from Experiments 1 and 2 is that auditory and visual information during witnessing can interfere with each other. There was clear evidence of visual information interfering with the processing of auditory information, and mixed support of auditory information interfering with the processing of visual information. It is not surprising that there was clearer evidence of visual interference. Past research has often demonstrated that visual cues predominate when visual and auditory stimuli are presented together (Howard & Templeton, 1966).

In spite of the reasonableness of the findings, they must be interpreted cautiously. The witnessing situation was more artificial than Yarmey's in that subjects simply looked at a picture of the criminal as opposed to seeing him in the act of committing the crime. Further, there was a briefer delay before lineup than in Yarmey's research. Finally, as neither we nor Yarmey attempted to sample different types of targets, it is possible that the differences in findings are a function of

the particular targets used in the two experiments. It may be that the current situation was an optimal situation for showing interference effects as contrasted with the Yarmey situation which may not have been suited to showing the effect.

To the extent that the current results do generalize, there are some practical implications for both the criminal justice system and for researchers of criminal justice issues. For the criminal justice system, the current findings suggest that police and ours should treat voice identifications made by auditory-visual witnesses with caution. For researchers, the fact that there *can* be interference effects suggests that eyewitness and earwitness research should not continue in isolation from each other. Research that is conducted on earwitnesses in the absence of visual information may not generalize to earwitness situations where visual cues are also available.

Acknowledgments

This research was supported by a grant from the Research and Grants Committee of Southeastern Louisiana University.

References

1. Bull R. & Clifford B. R. (1984). Earwitness voice recognition accuracy. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 92-123). Cambridge, England: Cambridge University Press.
2. Hollein H, Bennett G, & Gelfer M. P. (1983). Criminal identification comparison: Aural versus visual identifications resulting from a simulated crime. *Journal of Forensic Sciences, 28,* 208-221. PMID: 6680738
3. Howard I. P. & Templeton W. B. (1966). *Human spatial orientation*. New York: Wiley.
4. Lindsay R. C. L. & Wells G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70,* 556-564. http://dx.doi.org/10.1037/0021-9010.70.3.556
5. Loftus E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.
6. Navon D. (1990). How critical is the accuracy of an eyewitness memory? Another look at the issue of lineup diagnosticity. *Journal of Applied Psychology, 75,* 506-510. http://dx.doi.org/10.1037/0021-9010.75.5.506
7. Wells G. L. (1988). *Eyewitness identification: A system handbook.* Toronto: Carswell.
8. Wells G. L. & Lindsay R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88,* 776-784. http://dx.doi.org/10.1037/0033-2909.88.3.776
9. Wells G. L. & Luus C. A. E. (1990). The diagnosticity of a lineup should not be confused with the diagnostic value of non-lineup evidence. *Journal of Applied Psychology, 75,* 511-516. http://dx.doi.org/10.1037/0021-9010.75.5.511
10. Yarmey A. D. (1979). *The psychology of eyewitness testimony.* New York: Free Press.

11. Yarmey A. D. (1986). Verbal, visual, and voice identification of a rape suspect under different levels of illumination. *Journal of Applied Psychology, 71,* 363-370. http://dx.doi.org/10.1037/0021-9010.71.3.363
12. Yarmey A. D. (1992). The effects of dyadic discussion on earwitness recall. *Basic and Applied Social Psychology, 13,* 251-263. http://dx.doi.org/10.1207/s15324834basp1302_8
13. Yarmey A. D. & Matthys E. (1992). Voice identification of an abductor. *Applied Cognitive Psychology, 6,* 367-377. http://dx.doi.org/10.1002/acp.2350060502