

Verbal Overshadowing: A Sound Theory in Voice Recognition?

THEA VANAGS¹, MARIE CARROLL^{2*} and TIMOTHY J. PERFECT³

¹*University of Canberra, Australia*

²*Australian National University, Canberra, Australia*

³*University of Plymouth, Plymouth, UK*

SUMMARY

Verbal overshadowing is the impairment of a person's recognition ability as a result of generating a verbal description. Two experiments involving 169 participants examined the effects of verbal overshadowing, race of voice (own/other) and cognitive style (holistic/analytic) on voice recognition. In Experiment 1, participants heard a recorded voice (own- or other-race) saying a short phrase. After completing a cognitive style analysis and 15-minute filler task, the verbalisation group gave a written description of the voice while the control group did a filler task. Participants then attempted to identify the voice from a 6-voice lineup tape. Experiment 2 manipulated the similarity of the own-race voices by using telephone recordings of the voices, and the encoding-test similarities of the stimuli by using different phrases at encoding and test. Results showed a strong own-race bias with superior own-race voice recognition and no verbal overshadowing in Experiment 1, and a strong verbal overshadowing effect in Experiment 2. Cognitive style was predictive of voice identification in both experiments. Results are discussed with reference to the own-race bias, cognitive style and encoding-test similarity of the stimulus in verbal overshadowing. Copyright © 2005 John Wiley & Sons, Ltd.

While much research into human memory has suggested that verbal rehearsal and elaboration of information improves memory performance, a series of experiments by Schooler and Engstler-Schooler in 1990 raised doubt that this is always the case. In their experiments, Schooler and Engstler-Schooler asked participants to watch a 30-second videotape of a staged bank robbery. After a 20-minute delay, during which participants did an unrelated filler task, they were asked either to describe each feature of the robber's face in detail, or to continue with the filler task. Finally, participants were asked to identify the robber from a line-up of eight verbally similar faces. The results showed that participants who had described the target face were significantly less likely to identify the face.

As Schooler and Engstler-Schooler (1990) observed, much of the previous research into memory performance used verbal stimuli, and such stimuli were well-suited to verbal rehearsal and elaboration. In contrast, the modalities of stimulus and task in their own experiments were disparate (visual and verbal), with faces not lending themselves easily to verbal description. They hypothesised that verbal rehearsal or elaboration would be

*Correspondence to: Marie Carroll, QESS, Australian National University, ACT 0200, Australia.
E-mail: Marie.Carroll@anu.edu.au

effective in strengthening recognition ability only when a stimulus translates easily into words, and not when it is difficult to generate an accurate or complete description of the stimulus, particularly when the purpose of that description is to discriminate a target from distractors. Subsequent studies into verbal overshadowing have found this to be the case with continued evidence for verbal overshadowing in faces (Brown & Lloyd-Jones, 2003; Dodson, Johnson, & Schooler, 1997; Fallshore & Schooler, 1995; Finger, 2002; Finger & Pezdek, 1999; Schooler, Ryan, & Reder, 1996; Westerman & Larsen, 1997), as well as other difficult-to-describe stimuli such as colours (Schooler & Engstler-Schooler, 1990), wine (Melcher & Schooler, 1996), deep-structure memory (Lane & Schooler, 2004), and voices (Perfect, Hunt, & Harris, 2002).

Schooler and Engstler-Schooler (1990) initially suggested the interference between competing verbal and nonverbal memory representations of the stimulus at the time of retrieval as an explanation of verbal overshadowing. They believed this interference to be disruptive because the difficulty of a face description task results in a verbal encoding which is a distortion or misrepresentation of the original (visual) memory. Participants' descriptions focused on the most easily verbalised characteristics of the face, and tended to ignore the difficult-to-describe configurable information (Meissner, Brigham, & Kelley, 2001; Melcher & Schooler, 2004; Schooler & Engstler-Schooler, 1990). However, it is the configurable information which is most helpful in discriminating between target and distractor faces (Searcy & Bartlett, 1996; Tanaka & Farah, 1993, 2003; Tanaka & Sengco, 1997), and the absence of this important information results in a verbal encoding that is an inaccurate representation of the visual stimulus.

Indeed, it is this hypothesis that Meissner et al. (2001) have continued to favour as an explanation of verbal overshadowing, in particular, that verbal overshadowing is related to participants' response criterion during retrieval. Meissner and colleagues' findings have shown that when participants lower their response criterion (i.e. they are 'forced' to generate a description with more information about the target than they would naturally recall), they appear more susceptible to any self-generated misinformation (verbal overshadowing). In contrast, participants who are left to set their own response criterion (i.e. having 'free' recall and not being required to keep recalling additional information) are less likely to generate misinformation, and so less likely to experience interference from that (verbally encoded) misinformation during recognition.

Although the recoding interference hypothesis has proved a viable explanation for some verbal overshadowing findings (e.g. Finger & Pezdek, 1999; Meissner et al., 2001; Schooler & Engstler-Schooler, 1990), Schooler (2002) and his colleagues (Schooler, Fiore, & Brandimonte, 1997) have suggested that some findings challenge this explanation: (1) the lack of a relationship between the accuracy of a participant's verbal description and his or her ability to recognise the perpetrator (Brown & Lloyd-Jones, 2003; Fallshore & Schooler, 1995; Kitagami, Sato, & Yoshikawa, 2002; Schooler & Engstler-Schooler, 1990); (2) verbal overshadowing induced by verbalisation of a non-target face or a car (Dodson et al., 1997; Westerman & Larsen, 1997); (3) inducing a verbal overshadowing-type effect by replacing the verbalisation task with a non-verbal cognitive task (Navon letter task) (Macrae & Lewis, 2002; Perfect, 2003); and (4) the release of verbal overshadowing by engaging in a non-verbal task such as listening to instrumental music (Finger, 2002). In these cases, Schooler (2002) argues, the concept of interference between the verbal and visual encodings at retrieval cannot account for the existence (or lack of) verbal overshadowing.

In 1997, Schooler et al. put forward the transfer-inappropriate retrieval hypothesis (TIR) of verbal overshadowing. This hypothesis suggests that if the cognitive processes used during retrieval are not the same as those used during encoding of the memory, recognition will be impaired. For example, faces are encoded using primarily holistic, configural or non-reportable processes (Bartlett, Searcy, & Abdi, 2003), so recognition of those faces would be most effective when the same processes are engaged during retrieval. However, the task of generating a verbal description of a face activates a processing mode appropriate for verbalisation (likely to be featural, verbal or reportable in nature), and subsequent face recognition is hampered because the active (featural) processing mode is not helpful, and is even *inappropriate*, for the holistic task of face recognition (Melcher & Schooler, 2004; Schooler, 2002; Schooler et al., 1997).

Schooler (2002) has further refined this hypothesis to encompass the broader concept of a transfer-inappropriate processing shift (TIPS). The TIPS hypothesis reduces emphasis on the retrieval aspect, and takes into account Macrae and Lewis's (2002) findings that a cognitive task *not* involving retrieval of the original memory can invoke a similar outcome to the standard verbal overshadowing paradigm. Because voices, like faces, meet the criterion for the modality mismatch assumption fundamental to verbal overshadowing (Perfect et al., 2002), an investigation into own- and other-race voice recognition may provide further insight into the TIPS hypothesis while adding to the research on verbal overshadowing in voice recognition.

The limited research that has been done on eyewitness testimony to date shows that it can be subject to the same problems as eyewitness testimony (Bull & Clifford, 1999; Clifford, 1980), yet only one study has reported a significant effect of verbalisation on voice recognition— that of Perfect and his colleagues (2002). Furthermore, the limited research done into the effect of race on voice recognition has produced ambiguous results. The findings of McGehee, and Goldstein, Knight, Bailis and Conover (cited in Yarmey, 1995) suggest no difference between identification of own- and other-race voices, yet other research (Doty, 1998; and Thompson, 1987; Goggin, Thompson, Strube, & Simental, 1991; Hollien, Majewski, & Doherty, 1982; cited in Yarmey, 1995) suggests that there is a race effect with accented speakers being more poorly recognised than unaccented speakers (cf. the own-race bias in face recognition).

If an own-race bias does exist for voices, it is possible that a verbal overshadowing outcome similar to that of own- and other-race faces (see Fallshore & Schooler, 1995) could be expected. As with faces, individuals have greater perceptual expertise with own-race voices; expertise has been associated with an enhanced ability to process the configural rather than the featural aspects of the stimulus (Fallshore & Schooler, 1995; Melcher & Schooler, 1996, 2004), so individuals may use implicit, nonverbal, non-analytic processes for encoding own-race voices. In contrast, individuals have little experience or skill in describing voices (Huss & Weaver, 1996; Ormerod, 2001; Yarmey, 2001b), and could be assumed to rely on explicit, verbal, analytic processing for this task (Melcher & Schooler, 2004). If this were the case, a verbal overshadowing outcome would be expected for own-race voices, as the verbalisation task would result in an analytic (and inappropriate) strategy for the recognition task. Conversely, verbal overshadowing would not be expected for other-race voices, as individuals have little perceptual expertise with the voices themselves and little expertise in describing the voice characteristics.

Such discrepancies in the levels of verbal and perceptual expertise may also be attributable to individual differences. Ryan and Schooler (1998) found that for individuals

whose verbal expertise or ability is below the median level, verbalisation *itself* can cause a disruption to the individual's ability to rely on perceptual expertise, particularly when his or her perceptual ability is above the median. They further suggested that the susceptibility of individuals to verbal overshadowing might even be predictable from measures of individuals' perceptual and verbal ability.

Such differences in the levels of perceptual and verbal ability can be due to learning and the associated experience gained with a stimulus, or they may be due to innate differences between individuals. Some individuals have a pre-existing tendency to process information in a configural or holistic manner, while others favour a featural or analytic manner (Littlemore, 2001; Riding, 2000; Riding & Cheema, 1991; Zhang, 2002). The way in which a person habitually retrieves, perceives and processes information is often referred to as 'cognitive style' (Littlemore, 2001; Riding & Cheema, 1991), and according to Zhang (2002), an analytic cognitive style is good for processing verbal or featural information, whereas a holistic style happens to be good for processing spatial or configural information.

It seems possible that an analogy can be drawn between holistic and analytic cognitive styles and the TIPS hypothesis of verbal overshadowing. As Macrae and Lewis (2002) found, a verbal overshadowing-type effect can result from inducing a featural or analytic approach when the stimulus is more readily recognised using holistic processing. In contrast, inducing a global or holistic mode of processing for the same stimulus did not result in verbal overshadowing, but rather improved recognition performance. Hence, an investigation that incorporates such a measure of participants' cognitive style, as well as using a paradigm that attends to differences in processing mode due to the manipulation of expertise, may assist in investigating the role of TIPS in comparison to the role of recoding interference in verbal overshadowing.

EXPERIMENT 1

The first experiment in this study proposes to investigate verbal overshadowing, the own-race bias and cognitive style in voice recognition. Firstly, the intention is to replicate Perfect et al.'s (2002) findings that own-race voice recognition is impaired by verbalisation. Secondly, we expect that the well-established own-race bias that exists for face recognition will exist for voice recognition, and we also investigate the hitherto untested assumption that this bias, and the associated expertise, will result in a verbal overshadowing effect for own-race voices but not for other-race voices. Finally, this experiment will investigate the effect of cognitive style on voice recognition, with the anticipation that if cognitive style is a predictor of voice identification ability, then individuals with a holistic cognitive style will be better at identifying own-race voices, which are likely to be processed holistically, than those with an analytic cognitive style.

Method

Design

This experiment was a 2 (verbalisation or no-verbalisation) \times 2 (own-race or other-race) between-subjects factorial design. In addition, a continuous variable (holistic/analytic ratio) was recorded for each participant.

Participants

There were 133 native Australian-speaking participants in this experiment (113 female, 20 male). Participants were mostly University of Canberra first-year undergraduate psychology students who received course credit for their participation. Ages ranged from 17 to 55 years ($M = 25.9$, $SD = 9.39$).

Materials

A cassette tape recorder, 14 × 60-minute audiotapes, personal computers, Riding's (2001) Cognitive Style Analysis (CSA) program, cognitive style analysis interpretation sheets and participant booklets were used in this experiment.

Audiotapes

There were 14 audiotapes used. The voices of 12 different males (six Australian and six English) were recorded on 12 of these tapes. The phrase used for these recordings was the same phrase used by Perfect et al.'s (2002) study 'Just follow the instructions, don't press the alarm and no one will get hurt'. The remaining two audiotapes were fresh recordings of all six same-race (one tape Australian, one tape English) voices repeating the same phrase. These two voice lineup tapes were not copies of the 12 single audiotapes, but a second recording of each individual saying the same phrase a second time.

The audiotapes were made in a quiet room with only the experimenter and owner of the voice present. The tape recordings were made using a Sony cassette recorder (TCM-465V) using a uni-directional popup microphone and ACME C60 extra high-grade cassette tapes.

Australian and English voices

The Australian voices used in this experiment belonged to six professional men working at the University of Canberra. They were native Australian speakers ranging in age from approximately 30 to 50 years. The English voices belonged to six English men, of approximately the same age range, all of whom had been in Australia less than three years. The English voices were similar in terms of being relatively accent-free (they were from England, not Scotland, Wales or Northern Ireland), and not readily identifiable as from a particular region by native Australian speakers. Neither the Australian nor the English voices were selected from a larger pool of voices. All six Australian and English voices were randomly assigned the role of target voice during the experiment.

Cognitive style analysis (CSA) computer program

The CSA program (Riding, 2000) ran on personal computers. It generated a cognitive style for the user on two orthogonal axes: holistic/analytic and verbaliser/imager. Each of these two axes has a third value that falls between the two extremes: intermediate for the holistic/analytic axis and bimodal for the verbaliser/imager axis. The result was a cognitive style of holistic/intermediate/analytic and verbaliser/bimodal/imager.

An individual's position on these dimensions was determined by computing a ratio of reaction times to true/false questions (Riding, 2001). The program also provided speed and accuracy indicators to allow the exclusion of outlier data. Overall, indications are that this program is a valid predictor of people's performance on tasks that are hypothesised to involve either holistic or analytic processing (Peterson, Deary, & Austin, 2003; Riding,

2001, 2003). However, Riding (2001) concedes that there is a need for longer-term test-retest reliability studies. In a short-term, test-retest (8.5 days) and parallel forms and split-half analysis of the CSA program, Peterson et al. suggested that the holistic/analytic scale was stable (Mean $r = 0.69$) but the verbaliser/imager scale was not (Mean $r = 0.36$). For this reason the verbaliser/imager scale will not be used in these experiments.

The holistic/analytic dimension of cognitive style is assessed by the CSA through two subtests (40 items in total) (Riding, 2001). The first subtest presents items containing pairs of complex geometric figures. The participant is required to determine if the two figures are identical. Riding assumes a faster response time for such items indicates a holistic processing style. The second subtest displays a simple geometric shape, for example a square and a complex geometric shape. The participant has to determine if the simple figure is contained within the complex figure. Here it is assumed that individuals with an analytic processing style will respond more quickly than individuals with a holistic cognitive style.

Participation booklet

The participation booklets consisted of a cover page collecting demographic information (age and sex); five mazes which constituted the filler task; instructions for the verbalisation condition or control condition (a word search task); and a lineup response sheet asking participants to identify which voice was the original voice they heard, and to rate their confidence in their choice.

Cognitive style analysis interpretation sheet

The cognitive style analysis interpretation sheet (Riding, 2001) provided a brief description of the different cognitive styles and how they relate to the way in which people process information.

Procedure

Participants were tested in groups of between 1 and 10 people, and each participant was randomly allocated to the verbalisation or control condition, and the own- or other-race condition.

Participants were told that they were going to hear a tape recording and were asked to listen to it. At this point, they heard one of the six Australian or one of the six English voices saying the phrase 'Just follow the instructions, don't press the alarm and no one will get hurt'.

Participants were asked to then work through the CSA program. There was no time limit for this task, however no participant spent more than 10 minutes on this task. After completing the CSA program, participants worked on the maze task for 10 minutes. The maze task consisted of a maximum of five paper and pencil mazes. No one completed all mazes in the allotted time.

In the next phase, participants in the verbalisation condition were given the following instructions:

Please provide below a written description of the recorded voice you just heard, as if you were trying to describe the voice to the police. We are not interested in what was said, but the characteristics of the voice saying the phrase. Please provide as much information as you can, in as much detail as you can.

These participants were given five minutes for the description task while the control condition participants spent that time generating words from a set of nine letters. There was no verbal encouragement given to either the verbalisation or the control group to continue writing for the full five minutes.

For the final phase of the experiment, all participants were given the following instructions:

Please turn over the page to the next task. I am about to play you a voice lineup tape. This is a tape of six voices all saying the same thing as the original voice you heard. The original voice you heard may be one of the six or it may not. You will be asked to choose which voice was the original voice, or to indicate the original voice is not in the lineup. Once you have done that please complete the remaining questions on the page.

At this point, the appropriate voice lineup audiotape (Australian or English) was played. This marked the end of the experiment.

Results

Analysis of voice identification

The dependent variable, voice identification, was categorical, with two possible outcomes: correct identification or incorrect identification (as the target was always present). Analyses were undertaken to determine whether voice identification was comparable for each of the six voices used in the experiment.

For own-race voices, an exact significance test revealed that voice number and voice identification were *not* independent, $\chi^2(5) = 10.5$, $p = 0.05$. The majority of participants correctly identified voices 1 and 2 (89% and 80% respectively), whereas only about half the participants correctly identified voices 3, 4, 5 and 6. The proportions of correct and incorrect identifications for each voice are given in Table 1. For other-race voices, voice number and voice identification were independent, Pearson's $\chi^2(5) = 5.1$, $p = 0.44$, and can be seen in Table 1, the majority of identifications for each voice were incorrect.

Voice similarity ratings

For the six own-race voices, 12 native Australian speakers rated the similarity of each voice to each other voice on a five-point scale. The five-point similarity scale was 1 = 'very different', 2 = 'somewhat similar', 3 = 'reasonably similar', 4 = 'very similar' and 5 = 'identical', with a higher rating indicating a voice had greater similarity to the other voices. Ratings for all six voices ranged from 1 (very different) to 4 (very similar). All voices, except voice 3 ($M = 2.2$, $SD = 1.03$), had a mean rating close to 'somewhat similar'

Table 1. Proportion correct and incorrect by voice number

		Voice number (%)						Total
		1	2	3	4	5	6	
Own race	Incorrect	11	20	50	55	55	50	38
	Correct	89	80	50	45	45	50	62
Other race	Incorrect	70	60	70	100	78	67	74
	Correct	30	40	30	0	22	33	26

(voice 1 ($M = 1.7$, $SD = 0.99$); voice 2 ($M = 1.8$, $SD = 0.92$); voice 4 ($M = 1.8$, $SD = 0.96$); voice 5 ($M = 1.9$, $SD = 0.95$); voice 6 ($M = 1.7$, $SD = 0.96$)). No one voice was rated as significantly more or less similar than the other voices, $F_{(5, 354)} = 1.99$, $p = 0.08$.

Race of voice

A binary logistic regression analysis was performed for voice identification (correct, incorrect) with race of voice (own, other) as the predictor variable. The model was reliable with race of voice accounting for a change in deviance of $\chi^2(1) = 18.0$, $p < 0.001$. Participants were significantly better at identifying own-race voices (62% correct) than other-race voices (26% correct). The odds of a participant correctly identifying a voice increased by a factor of 4.7 when that voice was an own-race voice rather than an other-race voice. This race bias accounted for between 12.7% (Cox and Snell) and 16.9% (Nagelkerke) of the variance in voice identification. Table 2 shows the number of correct and incorrect identifications by race.

Verbalisation and control groups

Logistic regression analysis for voice identification (correct, incorrect) was performed with group (control, verbalisation) as the predictor variable for both own-race and other-race voices. There was no verbal overshadowing effect for own-race voices, change in deviance of $\chi^2(1) = 0.03$, $p = 0.86$ and no verbal overshadowing for other-race voices, change in deviance of $\chi^2(1) = 0.006$, $p = 0.94$. Furthermore, the group accounted for no variance (0%) in voice identification for either own- or other-race voices.

Cognitive style

Cognitive styles ranged from 0.5 to 3.6 ($M = 1.5$, $SD = 0.56$), and fell within the typical range of 0.4 to 4.0 (Riding, 2000). The distribution of cognitive styles was unimodal, leptokurtic (3.3) and positively skewed (1.5), possibly due to ratios being derived from response times (Riding). According to Riding's guidelines, 14.3% were holistic, 32.3% were intermediate and 53.4% were analytic.

To evaluate the effect of cognitive style on voice identification, a further logistic regression was performed with cognitive style as the predictor variable. The full model was considered an acceptable fit for the data, $-2 \text{ Log Likelihood} = 178.99$, Hosmer and Lemeshow $\chi^2(8) = 11.32$, $p = 0.18$, with a change in deviance of $\chi^2(1) = 4.48$, $p = 0.034$. However, cognitive style accounted for only 3.3% (Cox and Snell) to 4.4% (Nagelkerke) of the variance in voice identification.

The odds ratio indicated that an *increase* in the value of the holistic/analytic ratio of one (towards a more *analytic* style) was associated with an *increase* in the odds of correctly identifying a voice by a factor of 2. In contrast, with each decrease in the value of the

Table 2. Voice identifications by group

	Group		
	Control	Verbalisation	Total
Incorrect	38 (54%)	34 (54%)	72 (54%)
Correct	32 (46%)	29 (46%)	61 (46%)
Total	70 (100%)	63 (100%)	133 (100%)
	Own Race	Other Race	Total
Incorrect	27 (38%)	45 (74%)	72 (54%)
Correct	45 (62%)	16 (26%)	61 (46%)
Total	72 (100%)	61 (100%)	133 (100%)

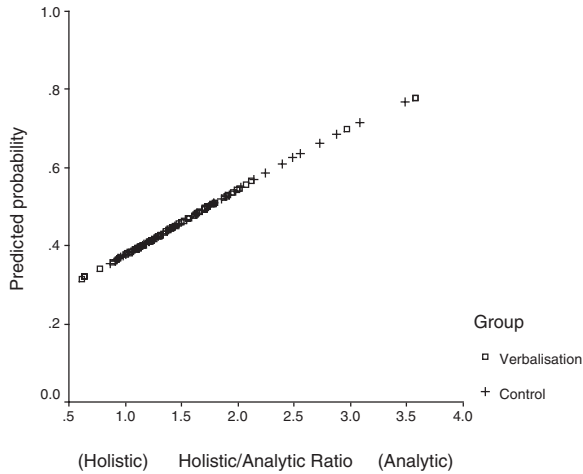


Figure 1. Predicted probability of identification as a function of cognitive style in Experiment 1

holistic/analytic ratio towards a more holistic style, the odds of correctly identifying a voice halved.

The relationship between cognitive style and voice identification can be seen in Figure 1 which shows the predicted probability of identifying a voice against the holistic/analytic ratio. It can be seen from Figure 1 that the more *analytic* the cognitive style of the participant, the *greater* the probability of identifying the voice from the lineup, and the more holistic the cognitive style of the participant, the lower the probability of identifying the voice.

Group and cognitive style

A final logistic regression for voice identification was performed to determine whether an interaction existed between group and cognitive style. The full model was a good fit for the data, $-2 \text{Log Likelihood} = 182.21$; Hosmer and Lemeshow $\chi^2(4) = 5.84$, $p = 0.21$, however there was no interaction between group and cognitive style with a change of deviance of $\chi^2(1) = 1.25$, $p = 0.26$.

Discussion

This experiment failed to find verbal overshadowing in own- or other-race voice recognition. The absence of verbal overshadowing made it impossible to identify whether the verbal overshadowing effect was stronger for own-race voices than other-race voices, as had been predicted. However, the findings clearly indicated that an own-race bias exists for voices, with a highly significant difference between participants' identification of own- and other-race voices. Race of voice accounted for approximately 13% to 17% of the variance in voice identification, and the odds of a participant identifying an own-race voice were 4.7 times greater than those of identifying an other-race voice. This provides support to earlier findings by Doty (1998) and Thompson, Goggin et al. and Hollien et al. (cited in Yarmey, 1995) that individuals are significantly worse at identifying voices of individuals of countries other than their own.

The hypothesis that if cognitive style could be used as a predictor of voice recognition individuals with a holistic cognitive style would be better at voice identification than those with an analytic cognitive style was not supported. In fact the reverse was true. As participants moved from a holistic cognitive style towards an analytic cognitive style they were about twice as likely to correctly identify the voices in this study. The effect, however, was not as strong as that for race of voice, and accounted for only 3% to 4% of the variance in voice identification.

The lack of verbal overshadowing, and the tendency for analytic individuals to be better at voice identification than holistic individuals, could lead to a suggestion that unlike faces, voices may be featurally processed. However, analysis of own-race data showed that identification and voice number were not independent, and independent raters found the voices only 'somewhat similar' to one another. This suggests that participants may have found some own-race voices easier to identify than others, and would be consistent with Kitagami et al.'s (2002) findings that when test-set similarity of the target and distractors is low, verbal overshadowing is less likely to occur.

Nevertheless, there may be another explanation for the lack of verbal overshadowing. While many previous verbal overshadowing studies have used stimuli that are featurally quite different at encoding and test, for example video at encoding versus static photograph at test (Perfect, 2003; Schooler & Engstler-Schooler, 1990; Westerman & Larsen, 1997), or casual photograph at encoding versus formal photograph at test (Dodson et al., 1997; Fallshore & Schooler, 1995; Ryan & Schooler, 1998), the encoding and test stimuli in this experiment had a high level of similarity with the same phrase used for encoding and lineup recordings. This high encoding-test similarity of the target may prevent verbal overshadowing from occurring.

In fact, Read and Craik (1995) have suggested that high encoding-test similarity can account for greater accuracy in voice identification. They intimate that replaying the original voice recording or having the same voices say the same words at encoding and test results in better rates of identification because people have certain ways of uttering specific phonemes, and these are more easily detected when the content of the speech remains the same. So a high encoding-test similarity may encourage the use of featural processing, which could explain both better recognition of the voices by analytic participants over holistic participants as well as the lack of verbal overshadowing in this experiment.

Although it may appear that this experiment did not replicate Perfect et al.'s (2002) findings, a closer analysis of those findings shows the same absence of verbal overshadowing for 'normal' voices repeating the same phrase at encoding at test. Perfect et al.'s verbal overshadowing effect came from the distortion condition of voices recorded over the telephone. Furthermore, Bull and Clifford (1999) have proposed that telephones distort human voices, and this would reduce the likelihood of voice identification when compared to 'normal' voices; they also maintain that creating voice lineups by recording the target and distractor voices over the telephone (to match the modality of the original voice heard by the participants) does not compensate for the loss of information participants experience when hearing the original voice over the telephone.

EXPERIMENT 2

The aim of Experiment 2 is to reduce the opportunity for participants to rely on featural information, such as idiosyncrasies of speech, when identifying the voices, and to

decrease the similarity of the stimuli across encoding and test. It will employ the same six Australian voices used in Experiment 1, but all recordings of the voices will be made over the telephone as in Perfect et al.'s (2002) distortion condition. In addition, the phrase recorded for the lineup tape will be different from the initial phrase heard by participants.

For Experiment 2, only own-race (Australian) voices will be used. We expect that verbal overshadowing will occur, as the voices will not be easily identified by features such as peculiarities of speech, and recording the voices over the telephone will result in a loss of information, possibly featural information, as in the loss of featural information from visual blurring (Collishaw & Hole, 2000). We also predict that individuals with a holistic cognitive style will be better than those with an analytic style at identifying these voices, as there will be less featural information on which to rely.

Method

Participants

The participants in this experiment were 36 native Australian speaking University of Canberra first-year undergraduate psychology students (25 females, 11 males) who had not participated in Experiment 1. They ranged in age from 17 to 38 years ($M = 22.3$, $SD = 5.50$).

Materials and design

This experiment was a simple verbalisation or control between-subjects design. In addition, the continuous variable holistic/analytic ratio was recorded for each participant. The materials used were as for Experiment 1 in every respect except for two factors: (1) new recordings were made of the Australian voices over the telephone, and (2) the phrase used at lineup, 'For your own safety don't make a noise and don't move and nothing will happen to you,' was different to that used at encoding 'Just follow the instructions, don't press the alarm and no one will get hurt'.

Procedure

The procedure was the same as in Experiment 1; participants were tested in groups of between 1 and 10 people, and were randomly allocated to the verbalisation or control conditions.

Results

Analysis of telephone (other sentence) voice identification

Analysis showed that voice number and voice identification were independent, $\chi^2(5) = 8.1$, $p = 0.15$. The proportions of incorrect and correct responses for each voice are given in Table 3.

Table 3. Proportion correct and incorrect by voice number

	Voice number (%)						Total
	1	2	3	4	5	6	
Incorrect	50	50	86	38	100	60	64
Correct	50	50	14	62	0	40	36

Table 4. Voice identifications by group

	Group		
	Control	Verbalisation	Total
Incorrect	9 (47%)	14 (82%)	23 (64%)
Correct	10 (53%)	3 (18%)	13 (36%)
Total	19 (100%)	17 (100%)	36 (100%)

Verbalisation and control groups

In contrast to Experiment 1, binary logistic regression analysis for voice identification, with group as the predictor variable, did show a significant verbal overshadowing effect. The full model was reliable with group accounting for a change in deviance of $\chi^2(1) = 4.96$, $p = 0.026$. Participants in the control group were significantly better at identifying the voices (53% correct) than participants in the verbalisation condition (18% correct). The odds of a control group participant correctly identifying a voice were 5.2 times greater than those in the verbalisation condition. This model accounted for approximately 12.9% (Cox and Snell) to 17.6% (Nagelkerke) of the variance in voice identification. Table 4 shows the number of correct and incorrect identifications by group.

Cognitive style

Cognitive styles ranged from 0.8 to 3.1 ($M = 1.4$, $SD = 0.50$), and fell within the typical range of 0.4 to 4.0 (Riding, 2000). As with Experiment 1, the distribution of cognitive styles was unimodal, leptokurtic (4.3) and positively skewed (1.8). According to Riding's guidelines, 22.2% were holistic, 36.1% were intermediate and 41.7% were analytic. A Kolmogorov-Smirnov two-sample test indicated no significance difference between the distribution of cognitive styles in Experiments 1 and 2 ($Z = 1.21$, $p = 0.11$).

A logistic regression analysis was performed with cognitive style as the predictor variable. The full model was considered a good fit for the data, $-2 \text{ Log Likelihood} = 43.0$; Hosmer and Lemeshow $\chi^2(7) = 6.9$, $p = 0.44$, with a change in deviance of $\chi^2(1) = 4.12$, $p = 0.042$. Cognitive style accounted for 1.8% (Cox and Snell) to 14.8% (Nagelkerke) of the variance in voice identification.

The odds ratio indicated that a *decrease* in the value of holistic/analytic ratio of one (towards a more *holistic* style) was associated with an *increase* in the odds of identifying a voice by a factor of almost 7 (6.9). Conversely, with each increase in the value of the holistic/analytic ratio towards a more analytic style, the odds of correctly identifying a voice dropped by a factor of 0.14.

The predicted probability of identifying a voice was graphed against cognitive style; the trend in Figure 2 shows that the more *analytic* the cognitive style of the participant the *lower* the odds of identifying the voice from the lineup, and the more *holistic* the cognitive style of the participant the *greater* the odds of identifying the voice.

Group and cognitive style

A final logistic regression analysis for voice identification was performed to investigate an interaction between group and cognitive style. The model was not considered a good fit for the data, $-2 \text{ Log Likelihood} = 45.96$; Hosmer and Lemeshow $\chi^2(4) = 8.61$, $p = 0.072$, and

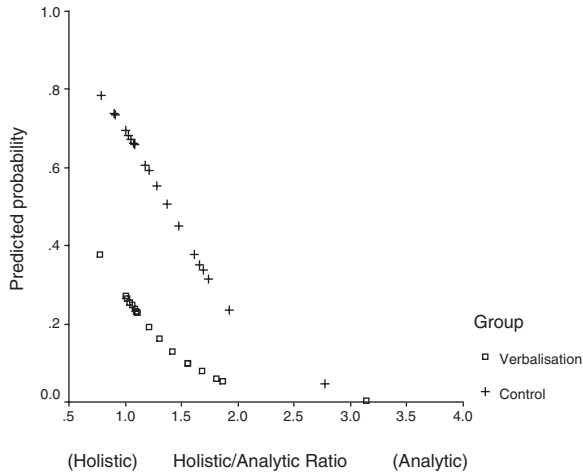


Figure 2. Predicted probability of identification as a function of cognitive style in Experiment 2

there was no interaction between group and cognitive style with a change in deviance of $\chi^2(1) = 1.13$, $p = 0.29$.

Discussion

The hypothesis that verbal overshadowing would occur as a result of generating a verbal description of the target voice was supported. There was a significant difference in identification by participants in the verbalisation and control groups, with group accounting for between 13% and 18% of the variance in voice identification. In line with Perfect et al.'s (2002) earlier findings, participants in the control group were significantly (5.2 times) more likely to correctly identify the voice than those in the verbalisation group. There was also support for the prediction that holistic individuals would better identify voices, with cognitive style being a significant predictor of voice identification and accounting for 11% to 15% of the variance in voice identification.

The relationship between cognitive style and recognition accuracy showed the opposite trend to that of Experiment 1. Although analytic individuals were more likely to correctly identify the voice in Experiment 1, holistic individuals were more likely to make a correct identification in Experiment 2. Despite no predictions having been made concerning an interaction between cognitive style and group, this trend reversal prompted investigation, as it could have been suggested that such an interaction would exist. Control participants with a holistic processing style could be predicted to be better at identifying a stimulus of that nature. Control participants with a featural style and holistic individuals in the verbalisation condition might be expected to suffer greater impairment in the recognition task than verbalisation condition participants' whose natural style was featural and similar to that induced by manipulation. However, investigations revealed no such interaction between cognitive style and group.

GENERAL DISCUSSION

These experiments examined the effect of verbal overshadowing, own-race bias and cognitive style on voice recognition. While Experiment 1 failed to find verbal

overshadowing in normal voice recognition, we did, however, find strong support for an own-race bias in voices. Participants were 4.7 times more likely to identify correctly an own-race voice over an other-race voice, and race of voice accounted for up to 17% of the variance in voice identification. These findings add support to earlier findings (e.g. Doty, 1998; Yarmey, 1995), and to Yarmey's (1995) claim that listeners will find people speaking with an accent less distinguishable than speakers of the listener's own regional area.

The absence of verbal overshadowing in other-race voices could be explained by Melcher and Schooler's (1996) discrepancy of expertise account. Participants had a low level of familiarity (perceptual expertise) with other-race voices and a low level of verbal expertise in voice description (Perfect et al., 2002; Yarmey, 2001a); hence there was no discrepancy in verbal and perceptual expertise levels, and no resultant verbal overshadowing. However, this explanation alone cannot account for the absence of verbal overshadowing in own-race voices. Participants could be assumed to have a high level of perceptual expertise with own-race voices, and still a low level of verbal expertise in describing those voices. Yet this discrepancy did not result in verbal overshadowing, suggesting that the occurrence of verbal overshadowing with voices is reliant on more than just a disparity in these levels of expertise.

In their review of the verbal overshadowing literature, Meissner and Brigham (2001) identified that verbal overshadowing is a somewhat fragile phenomenon. Their analysis showed that two factors could significantly influence its occurrence: post-description delay and instructional bias. A short delay (less than 10 minutes) between the description and identification phases results in a stronger verbal overshadowing effect than a long delay. However, this factor is unlikely to have influenced our results, as there was a delay of less than five minutes between description and identification in both experiments. Although instructional bias—the requirement for participants to keep producing detailed verbal descriptions past the point of free recall (Meissner, 2002)—also has a significant influence on the occurrence of verbal overshadowing, it does not explain the differential outcomes here. Again, the same recall instructions were given in both experiments. By similar reasoning, lineup format, which is known to influence verbal overshadowing (Kneller, Memon, & Stevenage, 2001; Lindsay & Wells, 1985; Meissner, 2002; Steblay, Dysart, Fulero, & Lindsay, 2001; Wells, 1984), is not likely to have influenced the disparity in outcomes. In fact, the only difference between Experiments 1 and 2 was the voice recordings themselves (normal or over the telephone employing a different phrase at lineup). This suggests that the nature of the stimulus and the encoding-test similarity of the stimulus are factors that may contribute to verbal overshadowing.

In Experiment 1, the stimuli were more readily identified by analytic individuals, and this points to featural processing being demanded by the stimulus (Zhang, 2002). If this were the case, then in line with the TIPS hypothesis, no verbal overshadowing effect would be expected, as the processing mode for the stimuli and description task were congruent (Schooler, 2002; Schooler et al., 1997). In Experiment 2, when the voices were recorded over the telephone and a different phrase was used at encoding and test, the trend was clearly the opposite of Experiment 2; holistic individuals were more likely to correctly identify the voice than analytic individuals suggesting the voices in Experiment 2 were better suited to holistic processing (Zhang, 2002). The resultant verbal overshadowing was consistent once more with the TIPS hypothesis, and the induced featural processing being inappropriate for a stimulus requiring holistic processing.

The reduction of encoding-test similarity in Experiment 2, through the use of a different phrase at lineup, may have resulted in participants being less able to rely on particular featural aspects of the voices such as the way specific phonemes were uttered. Instead, participants may have been forced to rely on more perceptual dimensions such as the pitch and loudness of the voice (Yarmey, 1995). If such dimensions are more holistic, then generating a verbal description and invoking a featural processing strategy would result in a transfer-inappropriate processing strategy and the verbal overshadowing outcome that occurred.

While this data appears to fit the TIPS hypothesis quite well, the recoding interference hypothesis does not seem to accommodate these results to the same extent. Considered in the light of response criterion (Meissner et al., 2001), the results from Experiment 2 should have been similar to those of Experiment 1, as no change was made to the description instruction; response criterion, whether determined to be high or low, was unchanged. Meissner et al. (2001) have suggested that criterion shift influences description accuracy as a lower criterion results in more inaccurate information being generated. However, if that featural information is no longer applicable at test (as in Experiment 2 due to the different phrase removing participants' ability to rely on idiosyncrasies of speech), then that misinformation should not affect recognition and verbal overshadowing should not occur. Yet, verbal overshadowing did occur. Conversely, if that featural information is still applicable at test (as in Experiment 1), it might be expected that this would elicit the greatest interference and verbal overshadowing should occur, but no verbal overshadowing was seen.

In summary, this research adds to the body of knowledge on voice recognition and earwitness testimony in two areas: the own-race bias and verbal overshadowing. Firstly the results showed a strong own-race bias with individuals being significantly better at recognising voices of speakers of their own country than those of another country. In addition, the results relating to the own-race bias indicate that a discrepancy in perceptual and verbal expertise is insufficient in itself to produce verbal overshadowing due to the absence of verbal overshadowing in own-race voices.

Secondly, the study provides greater support for the TIPS hypothesis of verbal overshadowing than the recoding interference hypothesis. When recognition was reliant on featural aspects of the stimuli, the featural processing mode induced by the verbal description task was appropriate for recognition and no verbal overshadowing occurred. In contrast, when recognition relied to a larger extent on holistic processing, the featural processing mode induced by the description task was inappropriate for recognition and a verbal overshadowing effect was seen.

Finally the study provides preliminary information on how an individual's cognitive style may affect voice recognition, particularly as the nature of the voice changes from a normal voice to one less easily identified by characteristics unique to that speaker. This has practical implications for voice recognition in earwitness testimony, as it may be possible to minimise verbal overshadowing effects by ensuring the phrase used during lineup is as close as possible to that originally spoken by the perpetrator. In addition, it is worth considering that people's cognitive style may also affect their ability to identify voices, with individuals having an analytic style performing better when featural aspects of a stimulus can be used, and those with a holistic style showing improved performance when there is less featural information available. However, this finding must be considered in the light of methodological weaknesses in the experiment. By asking participants to complete the cognitive styles analysis program and undertake a maze task during the experiment

itself, the effects of these activities on the participants' cognitive processing is not clear (e.g. Finger (2002) found a maze task could release verbal overshadowing). While it is less likely to have confounded the results of the verbalisation group (both activities took place before the verbalisation task), as with most other verbal overshadowing studies, we cannot predict the effect of these filler activities on the cognitive processing mode of the control group.

REFERENCES

- Bartlett, J., Searcy, J. H., & Abdi, H. (2003). What are the routes to face recognition? In M. A. Peterson, & G. Rhodes (Eds.), *Perception of faces, objects and scenes* (pp. 21–52). New York: Oxford University Press.
- Brown, C. L., & Lloyd-Jones, T. J. (2003). Verbal overshadowing of multiple face and car recognition: effects of within- versus across-category verbal descriptions. *Applied Cognitive Psychology, 17*, 183–201.
- Bull, R., & Clifford, B. R. (1999). Earwitness testimony. *Medicine, Science, and the Law, 39*(2), 120–127.
- Clifford, B. R. (1980). Voice identification by human listeners: on earwitness reliability. *Law & Human Behavior, 4*, 373–394.
- Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893–909.
- Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: why descriptions impair face recognition. *Memory and Cognition, 25*(2), 129–139.
- Doty, N. D. (1998). The influence of nationality on the accuracy of face and voice recognition. *The American Journal of Psychology, 111*(2), 191–214.
- Fallshore, M., & Schooler, J. W. (1995). The verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*, 1608–1623.
- Finger, K. (2002). Mazes and music: using perceptual processing to release verbal overshadowing. *Applied Cognitive Psychology, 16*(8), 887–896.
- Finger, K., & Pezdek, K. (1999). The effect of cognitive interview on face identification accuracy: release from verbal overshadowing. *Journal of Applied Psychology, 84*(3), 340–348.
- Huss, M. T., & Weaver, K. A. (1996). Effect of modality in earwitness identification: memory for verbal and nonverbal auditory stimuli presented in two contexts. *The Journal of General Psychology, 123*(4), 277–287.
- Kitagami, S., Sato, W., & Yoshikawa, S. (2002). The influence of test-set similarity in verbal overshadowing. *Applied Cognitive Psychology, 16*(8), 963–972.
- Kneller, W., Memon, A., & Stevenage, S. (2001). Simultaneous and sequential lineup: decision processes of accurate and inaccurate eyewitnesses. *Applied Cognitive Psychology, 15*(6), 659–671.
- Lane, S. M., & Schooler, J. W. (2004). Skimming the surface: verbal overshadowing of analogical retrieval. *Psychological Science, 15*(11), 715–719.
- Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*(3), 556–564.
- Littlemore, J. (2001). An empirical study of the relationship between cognitive style and the use of communication strategy. *Applied Linguistics, 22*(2), 241–265.
- Macrae, C. N., & Lewis, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science, 13*(2), 194–196.
- Meissner, C. A. (2002). Applied aspects of the instructional bias effect in verbal overshadowing. *Applied Cognitive Psychology, 16*(8), 911–928.
- Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15*(6), 603–616.
- Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition, 29*(1), 176–186.

- Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language*, 35, 231–245.
- Melcher, J. M., & Schooler, J. W. (2004). Perceptual and conceptual training mediate the verbal overshadowing effect in an unfamiliar domain. *Memory and Cognition*, 32(4), 618–631.
- Ormerod, D. (2001). Sounds familiar? Voice identification evidence. *Criminal Law Review*, August, 595–622.
- Perfect, T. J. (2003). Local processing bias impairs lineup performance. *Psychological Reports*, 93, 393–394.
- Perfect, T. J., Hunt, L. J., & Harris, C. M. (2002). Verbal overshadowing in voice recognition. *Applied Cognitive Psychology*, 16(8), 973–980.
- Peterson, E. R., Deary, I. J., & Austin, E. J. (2003). The reliability of Riding's cognitive style analysis test. *Personality and Individual Differences*, 34, 881–891.
- Read, D., & Craik, F. I. M. (1995). Earwitness identification: some influences on voice recognition. *Journal of Experimental Psychology: Applied*, 1(1), 6–18.
- Riding, R. J. (2000). *Cognitive styles analysis—research applications*. Birmingham: Learning & Training Technology.
- Riding, R. J. (2001). *Cognitive styles analysis—research administration*. Birmingham: Learning & Training Technology.
- Riding, R. J. (2003). On the assessment of cognitive style: a commentary on Peterson, Deary, and Austin. *Personality and Individual Differences*, 34, 893–897.
- Riding, R. J., & Cheema, I. (1991). Cognitive styles—an overview and integration. *Educational Psychology*, 11(3), 193–215.
- Ryan, R. S., & Schooler, J. W. (1998). Whom do words hurt? Individual differences in susceptibility to verbal overshadowing. *Applied Cognitive Psychology*, 12(Special Issue), S105–S125.
- Schooler, J. W. (2002). Verbalisation produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, 16(8), 989–997.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, 22, 36–71.
- Schooler, J. W., Fiore, S. M., & Brandimonte, M. A. (1997). At a loss from words: verbal overshadowing of perceptual memories. *The Psychology of Learning and Motivation*, 37, 291–340.
- Schooler, J. W., Ryan, R. S., & Reder, L. (1996). The costs and benefits of verbally rehearsing memory for faces. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M. Johnson (Eds.), *Basic and Applied Memory Research: Practical Applications* (Vol. 2, pp. 51–65). Mahwah, NJ: Erlbaum.
- Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 904–915.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: a meta-analytic comparison. *Law & Human Behavior*, 25(5), 459–474.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of experimental Psychology*, 46A(2), 225–245.
- Tanaka, J. W., & Farah, M. J. (2003). The holistic representation of faces. In M. A. Peterson, & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes* (pp. 53–74). New York: Oxford University Press.
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, 25(5), 583–589.
- Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, 14(2), 89–103.
- Westerman, D. L., & Larsen, J. D. (1997). Verbal-overshadowing effect: evidence for a general shift in processing. *American Journal of Psychology*, 110(3), 117–128.
- Yarmey, A. D. (1995). Earwitness speaker identification. *Psychology, Public Policy and Law*, 1(4), 792–816.

- Yarmey, A. D. (2001a). Earwitness descriptions and speaker identification. *Forensic Linguistics*, 8(1), 113–122.
- Yarmey, A. D. (2001b). Expert testimony: does eyewitness memory research have probative value for the courts? *Canadian Psychology*, 42(2), 92–100.
- Zhang, L. (2002). Thinking styles: their relationships with modes of thinking and academic performance. *Educational Psychology*, 22(3), 331–348.