ORIGINAL SCIENTIFIC PAPER

DOI: 10.59245/ps.35.2.4
UDK: 159.953 (615.07)
Received: March, 2025

ADNAN FAZLIĆ[*], IRMA DELJKIĆ[**], RAY BULL[***]

Verbalisation and Eyewitness Identification: Criterion Shift without Accuracy Loss in Bosnian Gen Z

Abstract

This study presents a planned replication of Schooler and Engstler-Schooler's (1990) core experiments to examine the effects of verbalisation on eyewitness identification in a Bosnian Generation Z sample. Across two experiments, verbalising the target's appearance did not reduce correct identifications but consistently shifted response tendencies: participants produced fewer false identifications and more 'not present' choices. These outcomes support the criterion shift account, indicating that verbalisation alters decision thresholds rather than impairing recognition memory, and align with large-scale replication findings. Post-decision confidence was higher for correct than false identifications, while confidence for rejections did not differ reliably, echoing recent syntheses of the confidence–accuracy relationship. A timing manipulation (delay before vs. after verbalisation) yielded weak-to-moderate differences in effect size but did not alter the overall pattern. The results demonstrate that the criterion shift effect generalises to a Bosnian Gen Z cohort, with practical implications for eyewitness procedures, emphasising conservative decisions and immediate confidence recording.

Keywords: *eyewitness identification, verbal overshadowing, criterion shift account, Generation Z, Bosnia and Herzegovina.*

[*] Adnan Fazlić, Faculty of Criminal Justice and Security Studies, University of Sarajevo, Bosnia and Herzegovina.
[**] Irma Deljkić, Faculty of Criminal Justice and Security Studies, University of Sarajevo, Bosnia and Herzegovina.
[***] Ray Bull, Department of Law and Social Sciences, University of Derby, England, United Kingdom.

1. INTRODUCTION

Eyewitness identification plays a pivotal role in legal and scientific contexts, yet its reliability hinges on the evidence collection methods of law enforcement (Wells et al., 2020). Practices such as double-blind lineups and post-identification confidence assessments are recommended to improve accuracy and fairness (Fitzgerald et al., 2021), whereas policy reforms are urged to reduce wrongful convictions (Wells, 2020). Various factors can skew the confidence-accuracy relationship, necessitating caution (Berkowitz et al., 2020; Kovera & Evelo, 2021). Police often require verbal descriptions from eyewitnesses—sometimes legally mandated before lineup identification—prompting research into their impact on performance (e.g., Akan et al., 2021; Kovera, 2024; Lockamyeir et al., 2020; Steblay & Wells, 2023; Tyler et al., 2023; Wilcock et al., 2008; Meissner et al., 2007), with evidence suggesting verbalisation may impair recognition memory.

The verbal overshadowing effect, first identified by Schooler and Engstler-Schooler (1990), refers to the paradox whereby describing a face can impair subsequent recognition. Their experiments demonstrated that this effect persisted across different lineup conditions and became more pronounced when participants provided detailed descriptions or engaged in longer reflection. Four main theoretical accounts have been advanced. *Recoding interference* proposes that verbal description generates a linguistic trace that competes with, and partially overwrites, the original visual memory. *Transfer-inappropriate retrieval* (TIR) suggests that verbalisation induces an analytic, feature-based processing mode that mismatches the holistic nature of face recognition, thereby impairing retrieval. *Retrieval-based interference* (RBI) emphasises that verbal reporting disrupts access to or monitoring of the original memory trace (sometimes producing source confusion), thereby reducing recognition accuracy. Finally, the *criterion shift* account argues that verbalisation does not degrade memory itself but shifts decision thresholds towards conservatism, increasing 'not present' responses and reducing guessing (Schooler & Engstler-Schooler, 1990; Schooler, 2002; Meissner et al., 2001; Clare & Lewandowsky, 2004). Meta-analyses by Meissner and Brigham (2001) and Alogna et al. (2014) confirm the robustness of the effect, although some studies highlight conflicting outcomes or context-specific variability (e.g. Clifford, 2003; Havard & Memon, 2012).

However, most prior work on verbal overshadowing predates the emergence of Generation Z (Gen Z), individuals born between 1995 and 2012, whose formative experiences of rapid technological change and shifting social dynamics may have reshaped cognitive processes (Szymkowiak et al., 2021). This cohort is marked by advanced visual processing, a preference for visual-mediated learning, and distinctive communication styles, alongside challenges such as shorter attention spans and potential limitations in social intelligence (Gabrielova & Buchko, 2021; Dimock, 2019; Seemiller & Grace, 2016, 2017). They also tend to show greater risk aversion than earlier generations (Seemiller & Grace, 2018), while demonstrating confidence and self-reliance that encourage reliance on their own abilities (Barhate & Dirani, 2021).

These attributes may shape how verbalisation interacts with memory and decision-making in eyewitness identification. To address this, the present study replicates core elements of Schooler and Engstler-Schooler (1990, Experiments 1 and 4) with a Bosnian Gen Z sample. Our contribution is twofold: we reproduce key tests of verbal overshadowing with procedural fidelity and transparent reporting, and we document the criterion shift pattern in a culturally

underrepresented context (Bosnia and Herzegovina) using a cohort increasingly salient in investigative practice. Crucially, this study does not aim to establish generational differences in the presence or magnitude of verbal overshadowing, but rather to test whether robust patterns generalise to this population and setting.

2. EXPERIMENT 1

Experiment 1 was a modified replication of Experiment 4 by Schooler and Engstler-Schooler (1990), including only experimental and control groups concerned with face verbalisation. It comprised two groups: an experimental group with a verbal description task and a control group without one. We aimed to examine whether significant intergroup differences in correct and incorrect answers would occur (as expected from the retrieval-based interference and transfer-inappropriate retrieval accounts). Our primary goal was to determine whether verbal descriptions significantly affect identification accuracy and how these influences manifest. Additionally, we examined whether verbalisation affects the type of error (false identification vs. not present) during the identification task (according to the criterion-shift account).

Experiment 1 involved a 20-minute delay between providing a verbal description and the identification task, to be compared with Experiment 2 (no delay), partly to test whether support could be found for findings from recent studies, such as Wilson et al. (2018). In Experiment 1 (as in Experiment 2), attention was paid to the possible effect of verbalisation on post-decision confidence and whether the confidence level is a reliable indicator of correct identification (Clare & Lewandowsky, 2004; Wixted & Wells, 2017).

2.1. Method

Participants. The sample consisted of undergraduate students from the University of Sarajevo - UNSA ($N = 98$; 55 – male, 43 – female; mean age = 18.90 years, $SD = 0.634$).[1]

Overview and design. Participants were randomly assigned to one of the two groups: experimental ($N = 49$) or control ($N = 49$). Experimental sessions included 1 to 10 participants. None had participated in relevant previous experiments.

Procedure. The first task was identical for both groups: participants watched a 44-second video recording of a simulated bank robbery. Afterwards, the experimental group had five minutes to write a description of the target face's appearance. The control group spent five minutes listing European capitals. Both groups then completed a 20-minute distractor task involving a crossword puzzle, similar to Schooler and Engstler-Schooler's (1990) original experiments. Immediately afterwards, participants performed a face recognition task using a photo lineup of 8 faces, including the perpetrator. They chose a number from 1 to 8 or 9 for 'not

[1] Students at the University of Sarajevo, Faculty of Criminal Justice and Security Studies participated in this research for course credit. This method of engaging participants in experimental research is common in most of the registered replication studies (Alogna et al. 2014). Students from other faculties at the University of Sarajevo voluntarily participated.

present' in the provided paper. Finally, participants indicated their post-decision confidence on a scale of 1 (completely unconfident) to 7 (completely confident).

2.2. Results and Discussion

Identification accuracy. Table 1 shows the identification performance data for the experimental and control groups in Experiment 1. To assess whether group membership (experimental vs. control) was related to identification accuracy in Experiment 1, a chi-square test was performed. In the experimental condition, 53.1% of participants made correct identifications compared with 49.0% in the control condition, while false identifications were recorded for 46.9% and 51.0% of participants, respectively. The Pearson chi-square test showed no significant association between condition and accuracy, χ^2 (1, N = 98) = 0.16, p = .686, with a trivial effect size (φ = .041). Fisher's exact test yielded the same conclusion (p = .840). Thus, verbalisation did not meaningfully affect whether participants correctly or incorrectly identified the perpetrator in this experiment.

Table 1: Response Type (%) and Self-Report Confidence (M, SD) for experimental and control groups in Experiment 1

	Experimental condition	
Response Type	Experimental group (N = 49)	Control group (N = 49)
Correct Identification	53.1	49.0
False Identification	20.4	38.8
Not Present	26.5	12.2
Confidence		
M	5.25	5.20
SD	1.17	1.56

We also examined the impact of verbalisation on the type of error (false identification or rejected identification). A chi-square test regarding the two groups and the two types of error found a significant effect, χ^2 (1) = 5.30, p = 0.021, ϕ = -0.332, and the phi-coefficient (ϕ), which in this case was used as the effect size measure, revealed that this effect of verbalisation decreasing false identifications and increasing rejected identifications was moderately large.

Thus, Experiment 1 did not support the findings of Schooler and Engstler-Schooler (1990), the two meta-analyses mentioned above, or other studies supporting the transfer-inappropriate retrieval and retrieval-based interference accounts (e.g. Dodson et al., 1997). Instead, our results align with those of Clare and Lewandowsky (2004), who support the criterion-shift account of the type of identification error. They found that in optional-choice lineups (i.e., a 'not present' option is available), where verbal description preceded identification performance, participants became more reluctant to identify the perpetrator.

Confidence. A univariate analysis of variance (ANOVA) was conducted to examine the effects of group (experimental vs. control) and identification accuracy (correct vs. false)

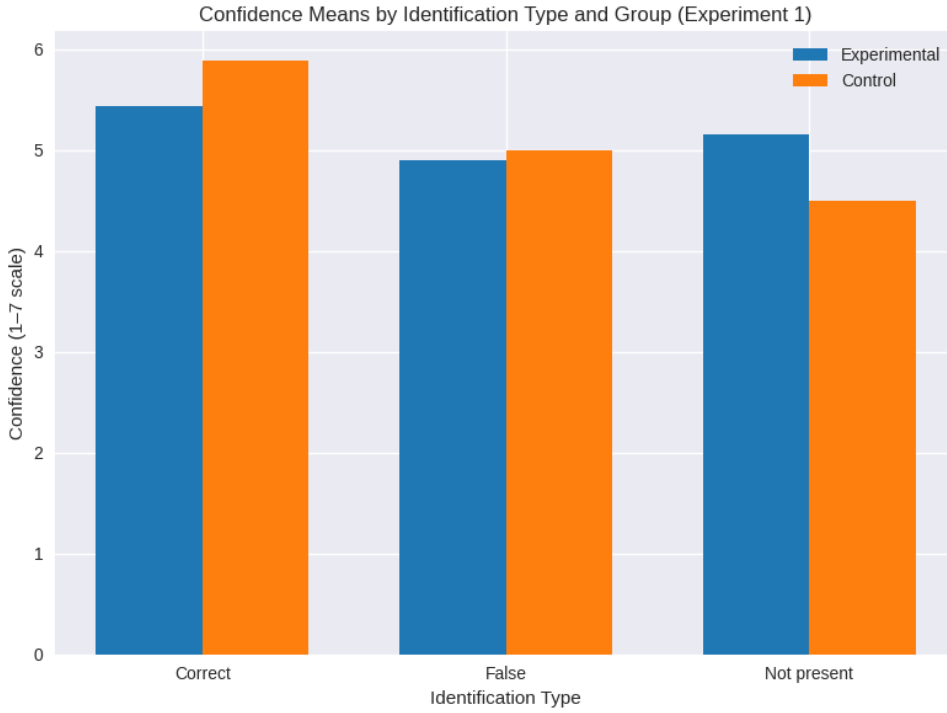
on post-decision confidence. Levene's test indicated that the assumption of homogeneity of variances was met, $F(3,93) = 0.09$, $p = .964$.

The analysis revealed that the main effect of group was not significant, $F(1, 93) = 0.31$, $p = .577$, $\eta^2 = .003$, indicating that verbalisation did not influence overall confidence ratings. In contrast, the main effect of identification accuracy was significant, $F(1, 93) = 8.24$, $p < .01$, $\eta 2 = .081$, indicating that confidence differed depending on whether participants made a correct or a false identification. Specifically, participants who correctly identified the perpetrator ($M = 5.65$, $SD = 1.17$) reported significantly higher confidence than those who made false identifications ($M = 4.96$, $SD = 1.22$), as confirmed by Tukey's Honestly Significant Difference (HSD) test. According to Cohen's (1988) criteria, the partial eta-squared value indicates a medium effect size. Finally, the interaction between group and identification accuracy was not significant, $F(1, 93) = 1.52$, $p = .220$, $\eta^2 = .016$, suggesting that the confidence–accuracy relationship was consistent across both experimental and control conditions.

Furthermore, a univariate ANOVA was conducted to examine the effects of group (experimental vs. control) and identification type (correct, false, not present) on post-decision confidence. Levene's test confirmed homogeneity of variances, $F(5, 91) = 0.44$, $p = .821$. The main effect of group was not significant, $F(1, 91) = 0.02$, $p = .886$, $\eta^2 < .001$, indicating that verbalisation did not influence confidence ratings. In contrast, identification type had a significant effect, $F(2, 91) = 4.57$, $p = .013$, $\eta^2 = .091$, with participants who made correct identifications ($M = 5.65$, $SD = 1.17$) reporting higher confidence than those who made false identifications ($M = 4.97$, $SD = 1.35$); confidence for "not present" choices ($M = 4.95$, $SD = 1.03$) was intermediate and not significantly different from either. The interaction between group and identification type was not significant, $F(2, 91) = 1.27$, $p = .285$, $\eta 2 = .027$, indicating that the confidence–accuracy relationship was consistent across both conditions. Overall, confidence was primarily determined by identification accuracy, with higher confidence accompanying correct identifications, while verbalisation did not undermine confidence, and rejections appeared to reflect cautious but deliberate decision-making.

The finding that verbalisation had no significant influence on post-decision confidence aligns with Clare and Lewandowsky (2004). The relationship between the post-decision confidence level and identification accuracy, in which high-confidence identifications can be considered accurate, is consistent with the findings of Wixted and Wells (2017).

Figure 1: Mean post-decision confidence ratings by identification type (correct, false, not present) and group (experimental vs. control) in Experiment 1. Bars represent mean values (M).



3. EXPERIMENT 2

Experiment 2 replicated the first experiment conducted by Schooler and Engstler-Schooler (1990). We examined the verbal overshadowing effect by comparing participants' identification performance when the task was performed immediately after a verbal description versus 20 minutes after viewing the target face. We also examined how verbalisation influences participants' post-decision confidence and explored the relationship between confidence and identification accuracy.

3.1. Method

Participants. The sample consisted of undergraduate students from the University of Sarajevo - UNSA (N = 102; 58 – male, 44 – female; mean age = 19.45 years, SD = 1.14).

Overview and design. Participants were randomly assigned to the experimental (N = 50) or control group (N = 52). Sessions included 1-10 participants; none had taken part in similar experiments previously.

Procedure. Participants viewed a 44-second simulated bank robbery video and then completed a 20-minute distractor task (crossword puzzle). Following this, the experimental

group spent five minutes writing a description of the target's face, while the control group listed European capitals for the same duration. Immediately afterwards, all participants undertook a lineup identification task using an eight-photo simultaneous lineup containing the perpetrator (response options: 1–8 or 9 = 'not present'), followed by a post-decision confidence rating on a 1–7 scale. The key procedural distinction was that Experiment 2 introduced the delay before verbalisation (or the control task) and lineup, whereas Experiment 1 introduced the delay after verbalisation (or the control task) and before the lineup.

3.2. Results and Discussion

Identification accuracy. In Experiment 2, a chi-square analysis was used to test whether identification accuracy (correct vs. false) differed between the experimental and control groups. Correct identifications were made by 48.0% of participants in the experimental condition and 42.3% in the control condition, while false identifications occurred in 52.0% and 57.7% of cases, respectively. The Pearson chi-square test showed no significant relationship between group and identification accuracy, χ^2 (1, N = 102) = 0.33, p = .564, with a negligible effect size (φ = .057). Fisher's exact test produced the same outcome (p = .691). These results indicate that verbalisation did not significantly influence the likelihood of producing a correct versus false identification in Experiment 2.

Furthermore, to examine whether verbalisation influenced the type of identification error, a chi-square test compared false identifications with "not present" responses across groups. In the experimental condition, 57.7% of errors were false, and 42.3% were rejections, whereas in the control condition, 83.3% were false and only 16.7% were rejections. The Pearson chi-square test indicated a significant association between group and error type, χ^2 (1, N = 56) = 4.49, p = .034, with a moderate effect size (φ = −.283). Fisher's exact test confirmed this result (p = .043). These findings suggest that verbalisation reduced the likelihood of false identifications and increased the tendency to reject identification, consistent with a criterion-shift pattern.

Thus, our Experiment 2 did not confirm the findings of Schooler and Engstler-Schooler (1990), Meissner and Brigham (2001), Alogna et al., (2014), nor the transfer-inappropriate retrieval and retrieval-based interference accounts. Experiment 2 was consistent with a criterion shift account (Clare & Lewandowsky, 2004), but the effect was weaker when the identification task was performed immediately after the verbal description. This result contrasts with Wilson et al. (2018), who found that the verbal overshadowing effect occurred when the identification task was performed immediately after the verbal description and did not occur with a 20-minute delay between the verbal description and the identification task.

Table 2: Response Type (%) and Self-Report Confidence (M, SD) for experimental and control groups in Experiment 2

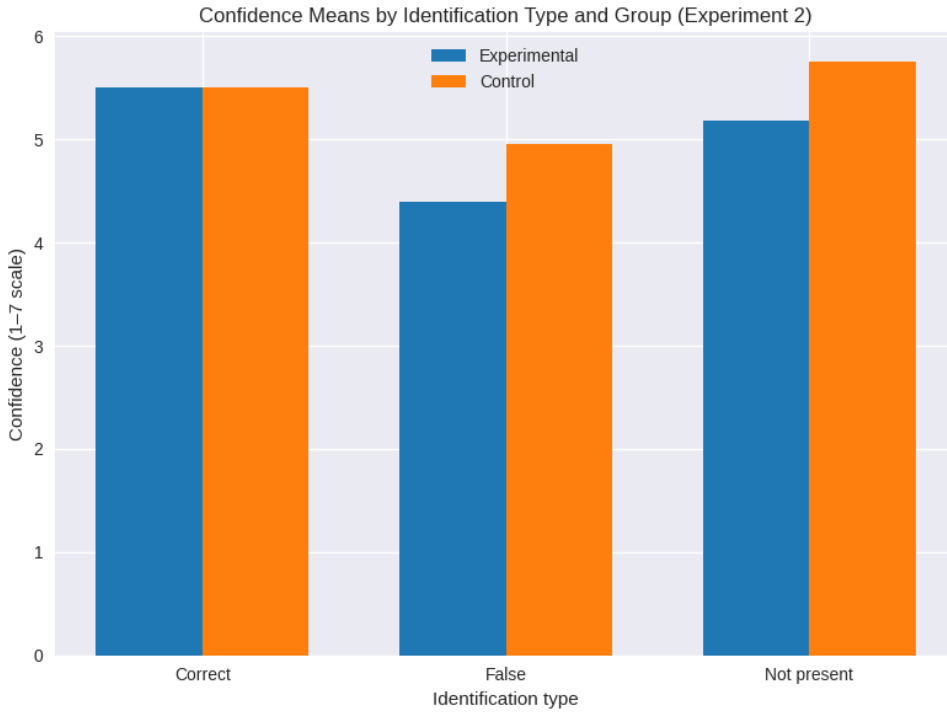
Response Type	Experimental condition	
	Experimental group (N=50)	Control group (N=52)
Correct Identification	48.0	42.3
False Identification	30.0	48.1
Not present	22.0	9.6
Confidence		
M	5.10	5.25
SD	1.34	1.50

Confidence. A univariate ANOVA was conducted to examine the effects of group (experimental vs. control) and identification accuracy (correct vs. false) on post-decision confidence. Levene's test confirmed that the assumption of homogeneity of variances was met, $F(3, 97) = 0.74$, $p = .530$. The main effect of group was not significant, $F(1, 97) = 0.36$, $p = .550$, $\eta^2 = .004$, indicating that verbalisation did not influence confidence ratings overall. In contrast, identification accuracy had a significant effect, $F(1, 97) = 4.54$, $p = .036$, $\eta^2 = .045$, with participants who made correct identifications ($M = 5.50$, $SD = 1.33$) reporting higher confidence than those who made false identifications ($M = 4.91$, $SD = 1.46$). In accordance with Cohen's criteria of effect size (Cohen, 1988), the partial eta squared value of .045 represents a small-to-medium effect size. The interaction between group and identification accuracy was not significant, $F(1, 97) = 0.36$, $p = .550$, $\eta^2 = .004$. Taken together, these results show that confidence was primarily determined by identification accuracy, while group membership and the interaction between group and accuracy did not significantly affect confidence.

Furthermore, a univariate ANOVA was carried out to assess the impact of group membership (experimental vs. control) and identification type (correct, false, not present) on post-decision confidence. Levene's test indicated that the assumption of equal error variances was satisfied, $F(5, 95) = 1.20$, $p = .317$. The analysis showed no significant main effect of group, $F(1, 95) = 1.21$, $p = .274$, $\eta^2 = .013$, suggesting that verbalisation did not alter overall confidence levels. By contrast, identification type produced a significant effect, $F(2, 95) = 3.83$, $p = .025$, $\eta^2 = .075$. Participants who correctly identified the perpetrator ($M = 5.50$, $SD = 1.33$) reported greater confidence than those who made false identifications ($M = 4.75$, $SD = 1.52$), while confidence for "not present" responses ($M = 5.33$, $SD = 1.23$) fell between the two and did not differ significantly. Based on Cohen's (1988) guidelines, the partial eta squared value of .075 reflects a medium effect size, indicating that identification accuracy exerted a meaningful influence on confidence ratings.

In general, Experiment 2 replicated the overall findings of Experiment 1.

Figure 2: Mean post-decision confidence ratings by identification type (correct, false, not present) and group (experimental vs. control) in Experiment 2. Bars represent mean values (M).



4. GENERAL DISCUSSION

The present study examined the verbal overshadowing effect originally outlined by Schooler and Engstler-Schooler (1990). Across two experiments, verbalisation did not reduce correct identifications. However, it consistently altered error patterns: participants who verbalised their observations were more likely to reject lineups with a 'not present' response, whereas those in the control condition more often produced false identifications. This reliable shift in response tendencies, observed without a decline in accuracy, aligns with criterion shift predictions and mirrors findings from prior replications, including Jong et al. (2016), Simmons et al. (2014), Susa and Swanner (2017), and Eggleston et al. (2016). In line with large-scale replication efforts, our results therefore indicate stability in the phenomenon while extending the evidence base to a Bosnian Gen Z cohort, a population of growing relevance in investigative practice.

Our findings lend partial support to the criterion shift account proposed by Clare and Lewandowsky (2004), which argues that verbalising a non-verbal stimulus alters the types of errors made during identification tasks. Unlike the transfer-inappropriate retrieval and retrieval-based interference accounts, which predict reduced accuracy, the criterion shift account anticipates a change in error patterns without necessarily affecting hit rates. Specifically, individuals who verbalise a stimulus before recognition appear to adopt a more conservative decision-making criterion, raising the threshold for matching their visual memory

to lineup faces. Because accurate identification depends on strong memory traces, this effect is most pronounced among individuals with weaker memory. For those with limited experience in tasks such as describing faces or participating in eyewitness interviews, the process can feel cognitively demanding. Without a reference point to evaluate the reliability of their verbal descriptions, such individuals are more vulnerable to the verbal overshadowing effect, increasing the likelihood of rejecting identification and selecting the 'not present' option.

These findings can also be considered within the framework of Generation Z's cognitive and behavioural characteristics. A substantial proportion of participants (42–53%) achieved accurate identifications, potentially reflecting the advanced visual skills noted by Seemiller and Grace (2017) and Djafarova and Bowes (2021), and even suggesting the presence of 'super recognisers' (Portch et al., 2024). Conversely, the identification errors observed may be linked to traits such as shorter attention spans and reduced social skills, consistent with Szymkowiak et al. (2021), Seemiller and Grace (2018), and Buhalis and Karatay (2022). Moreover, the risk-averse and pragmatic tendencies characteristic of Gen Z, particularly among those with weaker memory who provided verbal descriptions, could explain their preference for selecting the 'not present' option rather than risking a false identification. This cautious decision-making aligns with the insights from Seemiller and Grace (2018) and Barhate and Dirani (2021), underscoring how generational tendencies may shape eyewitness performance. Importantly, however, we do not interpret these findings as evidence of Gen Z-specific mechanisms; rather, they indicate that the criterion shift pattern generalises to a Bosnian Gen Z sample. The generational context remains useful for hypothesis generation, but the present data are consistent with established replication outcomes.

To better understand this phenomenon, it is important to consider Winkielman and Schwarz's (2001) notion that individuals base their self-assessment of task performance on past experiences, thereby forming specific expectations about their abilities. In the context of eyewitness identification, this means that witnesses evaluate their descriptive skills and memory capacity based on prior experiences, which, in turn, shape their expectations for identification accuracy. Moreover, these experiences and expectations are intrinsically linked to emotional responses; for example, limited experience with eyewitness identification and the high stakes of the decision can induce anxiety or social stress (Daudelin-Peltier et al., 2017; Pezdek et al., 2021). This emotional impact may compromise recognition memory, leading witnesses to choose what they perceive as the "least harmful option"—rejecting identification altogether. Supporting this view, Goette et al. (2015) found that both the personality of the decision-maker (e.g., anxious versus non-anxious) and the surrounding context significantly influence the decision-making process.

The error-pattern shift was numerically larger when a 20-minute delay followed verbalisation (Experiment 1) than when the delay preceded verbalisation (Experiment 2). However, the difference in phi coefficients indicates a weak-to-moderate overall effect and does not imply a robust timing dependency. Because hit rates were not reduced, the pattern is better explained by criterion shift than by transfer-inappropriate retrieval or retrieval-based interference.

A shift in the decision-making criterion must be examined in relation to post-decision confidence. Pezdek et al. (2021) found that acute stress does not impair the relationship between confidence and accuracy, while Wixted and Wells (2017) demonstrated that higher post-decision confidence often predicts accurate identification. Our study partially supports

these findings, as participants who made correct identifications exhibited significantly higher confidence than those who made false identifications. Interestingly, however, the confidence levels of participants who rejected identification were not significantly different from those of participants with correct identifications. Wixted and Wells (2017) caution that high post-decision confidence is not an infallible indicator of accuracy, particularly when memory may be compromised. They further suggest that a low confidence level, when reported honestly, should be regarded as a potential red flag for unreliable identification and duly noted.

Face recognition and memory retrieval are inherently complex, making the verbal overshadowing effect challenging to replicate. Meta-analyses by Meissner and Brigham (2001) and Alogna et al. (2014) reveal that while verbalisation produces a modest effect in one design and a larger effect in another, both are smaller than those observed in the original study (Pohl, 2022), thereby casting doubt on the robustness of the effect. Although Alogna et al. (2014) reported a robust verbal overshadowing effect (small, but replicated across multiple labs), this claim lacks sufficient support (Yarkoni, 2020). Moreover, these cognitive processes are influenced by factors such as encoding methods, context, and cultural background (Langlois, 2020), a critical consideration given that our study is the first conducted in Bosnia and Herzegovina. Additionally, studies in Slavic-speaking regions have not confirmed the effect (Sučić, 2011; Kabzińska, 2012). Finally, publication bias, often referred to as the file drawer effect, may contribute to an overrepresentation of positive findings, further complicating our understanding of this phenomenon (Gao, 2020).

4.1. Research Limitations

This research is limited by the artificial nature of experimental conditions, which often fail to capture the emotional, cognitive, and social dynamics of real eyewitness situations (Chae, 2010). In particular, the Schooler and Engstler-Schooler (1990) protocol may not accurately reflect real-world scenarios due to its fixed time delays between observation, verbal description, and identification, as well as its reliance on simultaneous lineups instead of sequential ones, a factor that can affect decision-making and increase false identifications (Meissner et al., 2005; Mickes & Wixted, 2015). Furthermore, while high post-decision confidence may indicate accurate identification even in non-pristine conditions (Wixted et al., 2018), there remains uncertainty about the optimal timing and measurement of such confidence (Wade et al., 2018). Future research should examine whether the criterion-shift account applies in target-absent lineups, assess the presence of the verbal overshadowing effect in sequential lineups, and further clarify the reliability of post-decision confidence as an accuracy indicator in more realistic settings.

4.2. Practical Implications

Our findings have important implications for law enforcement eyewitness identification procedures. Consistent with the criterion-shift account, verbalisation prior to identification does not enhance accuracy but does reliably alter decision thresholds, encouraging eyewitnesses, particularly those with weaker memory, to adopt a more conservative strategy. This shift

reduces guessing and lowers the risk of false identifications, thereby protecting innocent individuals in target-absent contexts while still allowing confident witnesses to make accurate identifications. Agencies should therefore anticipate more rejections and fewer false positives when verbalisation precedes identification, provided that procedures remain rigorous, employing double-blind administration, clear pre-identification instructions, reduced delays, and immediate confidence recording. Crucially, confidence measures should be captured at the time of decision and interpreted alongside lineup quality indicators, ensuring that verbalisation is understood as a mechanism for shifting response criteria rather than improving memory accuracy.

4.3. Future research

Future research should examine the verbal overshadowing effect in settings that more closely resemble real-life situations, such as by including longer periods between witnessing and recalling events and incorporating real-world stress factors to better reflect the experiences of actual eyewitnesses. Moreover, research should examine how cultural and linguistic differences affect memory retrieval and decision-making to understand their impact on identification accuracy. Considering Generation Z's unique cognitive and behavioural characteristics, such as improved visual processing abilities, shorter attention spans, and a tendency to avoid risks, future research should assess how these traits affect their effectiveness as eyewitnesses. Determining whether Gen Z members are more prone to the verbal overshadowing effect or whether their technological proficiency influences their ability to recognise faces could lead to more customised investigative methods.

Finally, studying the relationship between confidence after a decision and identification accuracy, especially within the framework of best-practice procedures, will be essential for developing more effective protocols to reduce false identifications and enhance the reliability of eyewitness testimony.

5. CONCLUSION

The present study replicated key elements of Schooler and Engstler-Schooler's (1990) verbal overshadowing paradigm with a Bosnian Gen Z sample, providing novel evidence from a culturally underrepresented population. Across two experiments, verbalisation did not impair recognition memory but consistently shifted response criteria, increasing lineup rejections and reducing false identifications while leaving hit rates stable. Confidence analyses further confirmed the robustness of the confidence-accuracy relationship, showing that witnesses remained able to calibrate confidence to accuracy. Taken together, these findings support the criterion-shift account of verbal overshadowing and highlight practical implications for investigative practice: verbalisation prior to identification should be understood as a mechanism that encourages conservative responding rather than as one that impairs memory.

REFERENCES

1. Akan, M., Robinson, M. M., Mickes, L., Wixted, J. T., & Benjamin, A. S. (2021). The effect of lineup size on eyewitness identification. *Journal of Experimental Psychology: Applied*, *27*(2), 369-392. https://doi.org/10.1037/xap0000340
2. Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, Š., Birch, S., Birt, A. R., Bornstein, B. H., Bouwmeester, S., Brandimonte, M. A., Brown, C., Buswell, K., Carlson, C., Carlson, M., Chu, S., Cislak, A., Colarusso, M., Colloff, M. F., Dellapaolera, K. S., Delvenne, J.-F., Zwaan, R. A. (2014). Registered Replication Report: Schooler and Engstler-Schooler (1990). *Perspectives on Psychological Science*, *9*(5), 556-578. https://doi.org/10.1177/1745691614545653
3. Barhate, B., & Dirani, K. M. (2021). Career aspirations of generation Z: a systematic literature review. *European Journal of Training and Development*, *46*(1/2), pp. 139-157. https://doi.org/10.1108/EJTD-07-2020-0124
4. Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2020). Convicting with confidence? Why we should not over-rely on eyewitness confidence. *Memory*, *30*(1), 10–15. https://doi.org/10.1080/09658211.2020.1849308
5. Buhalis, D., Karatay, N. (2022). Mixed Reality (MR) for Generation Z in Cultural Heritage Tourism Towards Metaverse. In J. L., Stienmetz, B. Ferrer-Rosell, & D. Massimo (Eds.), *Information and Communication Technologies in Tourism 2022. ENTER 2022* (pp. 16-27). Springer. https://doi.org/10.1007/978-3-030-94751-4_2
6. Chae, Y. (2010). Application of Laboratory Research on Eyewitness Testimony. *Journal of Forensic Psychology Practice*, *10*, 252-261. https://doi.org/10.1080/15228930903550608
7. Clare, J., & Lewandowsky, S. (2004). Verbalising Facial Memory: Criterion Effects in Verbal Overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 739-755. https://doi.org/10.1037/0278-7393.30.4.739
8. Clifford, B. R. (2003). The verbal overshadowing effect: In search of a chimera. In G. Vervaeke, M. Vanderhallen, P. J. van Koppen & J. Goethals (Eds.), *Much ado about crime: Chapters in psychology and crime* (pp. 151-162). Politeia
9. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, Second Edition. Lawrence Erlbaum Associates.
10. Daudelin-Peltier, C., Forget, H., Blais, C., Deschênes, A., & Fiset, D. (2017). The Effect of Acute Social Stress on the Recognition of Facial Expression of Emotions. *Scientific Reports*, *7*, 1036, https://doi.org/10.1038/s41598-017-01053-3
11. Dimock, M. (2019). *Defining generations: Where Millennials end and Generation Z begins*. https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins
12. Djafarova, E., & Bowes, T. (2021). 'Instagram made Me buy it': Generation Z impulse purchases in fashion industry. *Journal of Retailing and Consumer Services*, *59*, 102345. https://doi.org/10.1016/j.jretconser.2020.102345
13. Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, *25*(2), 129-139. https://doi.org/10.3758/bf03201107
14. Eggleston, C., Gilbert, E. A., & Lai, C. K. (2016). *RRR-Schooler-Eggleston*. https://osf.io/b4g79/
15. Fitzgerald, R. J., Rubínová, E., & Juncu, S. (2021). Eyewitness Identification Around the World. In A. M. Smith, M. P. Toglia, & J. M. Lampinen (Eds.), *Methods, measures, and theories in eyewitness identification tasks* (pp. 294-322). Taylor and Francis. https://doi.org/10.4324/9781003138105-16

16. Gabrielova, K., & Buchko, A. A. (2021). Here comes Generation Z: Millennials as managers. *Business Horizons*, *64*(4), 489-499. https://doi.org/10.1016/j.bushor.2021.02.013
17. Gao, Z. (2020). Researcher biases. In B. J. Carducci, C. S. Nave, J. S. Mio, & R. E. Riggio (Eds.), *The Wiley encyclopedia of personality and individual differences* (pp. 37-41). Wiley. https://doi.org/10.1002/9781119547167.ch76
18. Goette L., Bendahan S., Thoresen J., Hollis F., & Sandi C. (2015). Stress pulls us apart: anxiety leads to differences in competitive confidence under stress. *Psychoneuroendocrinology*, *54*, 115–123. https://doi.org/10.1016/j.psyneuen.2015.01.019
19. Havard, C., & Memon, A. (2012). Facial Recognition from Identification Parades. In C. Wilkinson & C. Rynn (Eds.), *Craniofacial Identification* (pp. 86-100). Cambridge University Press.
20. Jong, J., Alogna, V. K., & Halberstadt, J. (2016). *RRR-Schooler-Jong*. https://osf.io/sqzuq/
21. Kabzińska, J. (2012). Long-term Consequences of Providing a Perpetrator's Description on a Subsequent Lineup Performance. *VARSTVOSLOVJE, Journal of Criminal Justice and Security*, 14(2), 165-174. https://www.fvv.um.si/rV/arhiv/2012-2/03_Kabzinska.pdf
22. Kovera, M. B. (2024). The role of suspect development practices in eyewitness identification accuracy and racial disparities in wrongful conviction. *Social Issues and Policy Review, 18*(1), 125–147. https://doi.org/10.1111/sipr.12102
23. Langlois, J. (2020). Language and cognitive science: How language affects reasoning and memory. *Training, Language and Culture*, *4*(3), 66-76. https://doi.org/10.22363/2521-442X-2020-4-3-66-76
24. Lockamyeir, R. F., Carlson, C. A., Jones, A. R., Carlson, M. A., & Weatherford, D. R. (2020). The effect of viewing distance on empirical discriminability and the confidence–accuracy relationship for eyewitness identification. *Applied Cognitive Psychology*, *34*(5), 1047–1060. https://doi.org/10.1002/acp.3683
25. Meissner, C. A., & Brigham, J. C. (2001). A Meta-analysis of the Verbal Overshadowing Effect in Face Identification. *Applied Cognitive Psychology*, *15*, 603–616. https://doi.org/10.1002/acp.728
26. Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The Influence of Retrieval Processes in Verbal Overshadowing. *Memory & Cognition*, *29*(1), 176-186. https://doi.org/10.3758/bf03195751
27. Meissner, C. A., Sporer, S. L., & Schooler, J. W. (2007). Person Descriptions as Eyewitness Evidence. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The Handbook of Eyewitness Psychology: Volume II – Memory for People* (pp. 1-34). Routledge.
28. Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness Decisions in Simultaneous and Sequential Lineups: A Dual-process Signal Detection Theory Analysis. *Memory & Cognition*, *33*(5), 783-792. https://doi.org/10.3758/bf03193074
29. Mickes, L., & Wixted, J. T. (2015). On the Applied Implications of the "Verbal Overshadowing Effect". *Perspectives on Psychological Science*, *10*(3), 400-403. https://doi.org/10.1080/20445911.2024.2419888
30. Pezdek, K., Abed, E., & Cormia, A. (2021). Elevated stress impairs the accuracy of eyewitness memory but not the confidence–accuracy relationship. *Journal of Experimental Psychology: Applied*, *27*(1), 158–169. https://doi.org/10.1037/xap0000316
31. Pohl, R. F. (2022). Labeling and overshadowing effect. In R. F. Pohl (Ed.), *Cognitive Illusions: Intriguing Phenomena in Thinking, Judgment, and Memory*, 3rd Edition (pp. 386-482). Routledge. https://doi.org/10.4324/9781003154730

32. Portch, E., Attard-Johnson, J., Estudillo, A. J., Mestry, N., & Bate, S. (2024). "Super-recognisers" and the Legal System. In E. Pica, D. Ross, & J. Pozzulo (Eds.), *The Impact of Technology on the Criminal Justice System: A Psychological Overview* (pp. 272-300). Routledge. https://doi.org/10.4324/9781003323112
33. Risius, U. M., Staniloiu, A., Piefke, M., Maderwald, S., Schulte, F. P., Brand, M., & Markowitsch, H. J. (2013). Retrieval, Monitoring, And Control Processes: A 7Tesla fMRI Approach to Memory Accuracy. *Frontiers in Behavioral Neuroscience*, *7*, 1-21. https://doi.org/10.3389/fnbeh.2013.00024
34. Schooler, J. W. (2002). Verbalisation produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, *16*(8), 989-997. https://doi.org/10.1002/acp.930
35. Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal Overshadowing of Visual Memories: Some Things are Better Left Unsaid. *Cognitive Psychology*, *22*(1), 36-71. https://psycnet.apa.org/doi/10.1016/0010-0285(90)90003-M
36. Seemiller, C., & Grace, M. (2016). *Generation Z goes to college*. Jossey-Bass.
37. Seemiller, C., & Grace, M. (2017). Generation Z: Educating and Engaging the Next Generation of Students. *About Campus*, *22*(3), 21-26. https://doi.org/10.1002/abc.21293
38. Seemiller, C., & Grace, M. (2018). *Generation Z: A Century in the Making*. Routledge. https://doi.org/10.4324/9780429442476
39. Simons, J. D., Holcombe, A. O., Schooler, J. W., Drew, A., Spellman, B., Ballard-Wood, A., Michael, R. B. (2014). *RRR-Schooler-Mturk*. https://osf.io/ez4w3/
40. Steblay, N. K., & Wells, G. L. (2023). In their own words: Verbalisations of real eyewitnesses during identification lineups. *Psychology, Public Policy, and Law, 29*(3), 272–287. https://doi.org/10.1037/law0000386
41. Sučić, I. (2011). *Točnost prepoznavanja počinitelja kaznenih djela: Ispitivanje fenomena verbalnog zasjenjivanja* [Doctoral dissertation, Filozofski fakultet u Zagrebu]. https://www.croris.hr/crosbi/publikacija/ocjenski-rad/365671
42. Susa, K. J., & Swanner, J. K. (2017). *RRR-Schooler-Susa*. https://osf.io/5vunt/
43. Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Kundi, G. S. (2021). Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, *65*, 101565. https://doi.org/10.1016/j.techsoc.2021.101565
44. Tyler, R., Towler, A., Kemp R. I., & White. D. (2023). Let's talk about faces: Identifying faces from verbal descriptions, *British Journal of Psychology*, *114*(1), 262 - 281. http://dx.doi.org/10.1111/bjop.12610
45. Wade, K. A., Nash, R. A., & Lindsay, D. S. (2018). Reasons to Doubt the Reliability of Eyewitness Memory: Commentary on Wixted, Mickes, and Fisher (2018). *Perspectives on Psychological Science, 13*(3), 339-342. https://doi.org/10.1177/1745691618758261
46. Wells, G. L. (2020). Psychological science on eyewitness identification and its impact on police practices and policies. *American Psychologist*, *75*(9), 1316–1329. https://doi.org/10.1037/amp0000749
47. Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, *44*(1), 3–36. https://doi.org/10.1037/lhb0000359
48. Wilcock, R., Bull, R., & Milne, R. (2008). *Witness identification in criminal cases*. Oxford University Press.

49. Wilson, B. M., Seale-Carlisle, T. M., & Mickes, L. (2018). The Effect of Verbal Descriptions on Performance in Lineups and Showups. *Journal of Experimental Psychology: General*, *147*(1), 113-124. https://doi.org/10.1037/xge0000354
50. Winkielman, P., & Schwarz, N. (2001). How Pleasant was Your Childhood? Beliefs about Memory Shape Inferences from Experienced Difficulty of Recall. *Psychological Science*, *12*, 176–178. https://doi.org/10.1111/1467-9280.00330
51. Wixted, J. T., & Wells, G. L. (2017). The Relationship Between Eyewitness Confidence and Identification Accuracy: A New Synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://psycnet.apa.org/doi/10.1177/1529100616686966
52. Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). In the DNA Exoneration Cases, Eyewitness Memory Was Not the Problem: A Reply to Berkowitz and Frenda (2018) and Wade, Nash and Lindsay (2018). *Perspectives on Psychological Science*, *13*(3), 343-345. https://doi.org/10.1177/1745691618769859
53. Yarkoni T. (2020). The generalizability crisis. *The Behavioral and Brain Sciences*, *45*, e1. https://doi.org/10.1017/S0140525X20001685

Sažetak

Adnan Fazlić[*], **Irma Deljkić**[**], **Ray Bull**[***]

Verbalizacija i prepoznavanje osoba: Promjena kriterija bez gubitka točnosti u bosansko-hercegovačkoj generaciji Z

Studija predstavlja replikaciju temeljnih eksperimenata Schoolera i Engstler-Schoolera (1990) radi ispitivanja učinaka verbalizacije na prepoznavanje počinitelja na uzorku bosansko-hercegovačke generacije Z. U dva eksperimenta verbalizacija izgleda ciljne osobe nije smanjila točne identifikacije, ali je dosljedno mijenjala obrasce odgovora: sudionici su davali manje pogrešnih identifikacija i češće birali opciju „nije prisutan". Dobiveni nalazi podupiru teoriju promjene kriterija, sugerirajući da verbalizacija mijenja pragove odlučivanja umjesto da narušava memoriju prepoznavanja, te se podudaraju s rezultatima velikih replikacijskih studija. Pouzdanje nakon odluke bilo je veće kod točnih nego kod pogrešnih identifikacija, dok se pouzdanje kod odbijanja nije razlikovalo, što odražava sinteze odnosa između pouzdanja i točnosti. Manipulacija s vremenom (odgoda prije u odnosu na onu nakon verbalizacije) dovela je do slabih odnosno umjerenih razlika u veličini učinka, ali nije promijenila opći obrazac. Rezultati pokazuju da se učinak promjene kriterija generalizira na bosansko-hercegovačkoj generaciji Z, s praktičnim implikacijama za procedure prepoznavanja kojima se naglašavaju konzervativne odluke i neposredno bilježenje pouzdanja.

Ključne riječi: prepoznavanje osoba, verbalno zasjenjivanje, promjena kriterija, generacija Z, Bosna i Hercegovina.

[*] Adnan Fazlić, Fakultet za kriminalistiku, kriminologiju i sigurnosne studije, Sveučilište u Sarajevu, Bosna i Hercegovina.
[**] Irma Deljkić, Fakultet za kriminalistiku, kriminologiju i sigurnosne studije, Sveučilište u Sarajevu, Bosna i Hercegovina.
[***] Ray Bull, Katedra za pravne i društvene znanosti, Sveučilište u Derbyju, Engleska, Ujedinjeno Kraljevstvo.