Received: 20 January 2022 | Revised: 12 October 2022 | Accepted: 18 October 2022
DOI: 10.1002/acp.4012

RESEARCH ARTICLE

WILEY

Multisensory encoding of names via name tags facilitates remembering

Carolyn A. Murray[1] | Maisy Tarlow[1] | Jesse Rissman[1,2] | Ladan Shams[1,3]

[1]Department of Psychology, University of California, Los Angeles, California, USA

[2]Department of Psychiatry & Behavioral Sciences, University of California, Los Angeles, California, USA

[3]Department of Bioengineering and Neuroscience Interdepartmental Program, University of California, Los Angeles, California, USA

Correspondence
Carolyn A. Murray, Department of Psychology, University of California, Los Angeles, 1285 Psychology Building, Los Angeles, CA 90095, USA.
Email: camurray@ucla.edu

Funding information
University of California, Los Angeles Faculty Senate

Abstract

Associating names to faces can be challenging, in part because this task lacks an inherent semantic relationship between a face and name. The current study seeks to understand whether bolstering names with cross-modal cues—specifically, name tags—may aid memory for face and name pairings. In a series of five experiments, we investigated whether the presentation of congruent vocalized and written names at encoding might benefit subsequent cued recall and recognition memory tasks. The results showed that participants, cued with a picture of a face, were more likely to recall the associated name when those names were encoded with a name tag (a congruent visual cue) compared to when no supporting cross-modal cue was available. The findings were consistent with a benefit of multisensory encoding, above any effect from the availability of independent unisensory traces, extending previous findings of multisensory learning and memory benefits to a naturalistic associative memory task.

KEYWORDS
associative memory, multisensory benefit, multisensory processing, name learning

1 | INTRODUCTION

As any individual who has been to a large gathering can attest, remembering the association between names and faces is a challenge. While associative memory tasks tend to be among the most challenging in a laboratory setting, memory for names is one that is considered especially so, in part due to the lack of an inherent, semantic relationship between a face and name (e.g., there is not anything about one face that makes it seem more "Hannah" than another). This makes learning names a challenge, which has been the topic of much past research (for examples, see Brooks et al., 1993; Cohen & Faulkner, 1986; McWeeny et al., 1987).

Previous studies have tested a number of different approaches to improve recall of names associated with particular faces (we will provide a brief overview, but see Brédart, 2019 for a more comprehensive review). Spacing the learning of the names has been shown to improve recall performance (Carpenter & DeLosh, 2005), as has retrieval practice of particular names (Morris et al., 2005). However, these previously studied approaches may be difficult to use in the real world; for example, one probably cannot control how many people they meet at a conference, let alone the spacing between these meetings. Semantic associations or mental imagery devices, such as creating a mnemonic around the name and associating this with a physical feature or fact about the person, have also been shown to improve how well names are remembered (e.g., McCarty, 1980). Comparisons of these techniques, however, show that mnemonic techniques are less effective than spacing (Morris et al., 2005; Neuschatz et al., 2005), and so may benefit from additional supporting cues.

More recent research has begun to tap into the connection between the sensory content available at the time of encoding and the conditions present during later retrieval. Previous work has shown that encoding an audiovisual of a person talking is more effective for subsequent recognition of their voice than encoding the voice alone, showing the superiority of audiovisual encoding over auditory encoding in auditory recognition (von Kriegstein & Giraud, 2006). Learned congruence between a face and voice has been reported to speed recognition of a familiar face-voice pair, compared to an incongruent audiovisual pairing (O'Mahony & Newell, 2012). Interestingly, it has also been found that regions of the brain involved in audiovisual integration—for creating an association between congruent audio and visual cues—are activated more strongly during encoding for faces that will later be remembered than for those that are forgotten (Lee et al., 2017), so perhaps multisensory stimuli can support recall of face-name associations.

Facilitation of memory by utilizing multiple sensory cues would be consistent with a few memory models, most notably with dual-coding theory (see Clark & Paivio, 1991; Paivio, 1991 for reviews), wherein providing verbal and non-verbal representations (that often occur across different senses) can facilitate memory. Another similar model is the cognitive theory of multimedia learning (see Mayer, 2014 for overview), wherein presentation of stimuli across verbal and pictorial working memory channels allows for better learning. In general, encoding information across different channels can provide more routes by which a memory can be accessed. This would seemingly support findings in the multisensory research literature that multisensory information can improve memory, as multisensory information provides information through at least two senses, while unisensory information can provide only one sensory route to a memory. However, this particular framework fails to make a distinction between having information available across multiple senses and unified multisensory experiences, where congruency (temporal, spatial, structural, semantic, etc.) between stimuli can lead to the creation of integrated multisensory representations (e.g., Butler et al., 2012; Ernst & Bülthoff, 2004; Lacey et al., 2009; Laurienti et al., 2005; Shams & Kim, 2010; Spence, 2007). The integration of cues from multiple sensory modalities can result in overall improvement of the sensory signals, by, for example, uncertainty reduction leading to improved precision and/or accuracy. We seek to investigate if multisensory integration mechanisms, in particular, are able to support remembering face-name associations, beyond any benefit provided by multiple unisensory traces.

Previous work indicates there may be a memory benefit to presenting stimuli with meaningful and congruent cross-modal sensory inputs (see Matusz et al., 2017; Shams & Seitz, 2008 for an overview). For instance, studies have shown that object images are recognized better when they are originally presented with their iconic sound compared to when they are presented without sound, even when only the visual cue is presented at test (Lehmann & Murray, 2005). Similarly, auditory recognition is better for objects originally presented together with congruent images compared to audio-alone encoding (Moran et al., 2013), or to presenting the sound with a meaningless visual stimulus (Thelen et al., 2015). Improvements to recognition memory performance were also shown to extend to written words accompanied by audio of those words (Heikkilä et al., 2015; Heikkilä & Tiippana, 2016). While the exact mechanisms by which multisensory encoding benefits recall or recognition remain unexplained (though see proposed mechanisms in Shams & Seitz, 2008), electroencephalographic (EEG) signals measured during memory retrieval begin to diverge at a relatively early stage of processing for visual versus audiovisual information (Murray et al., 2004), indicating that multisensory stimulus encoding may involve distinct processes not triggered by unisensory encoding. This would suggest that there is a distinct benefit to using multisensory cues as opposed to multiple unisensory ones, which could provide an avenue to boost memory performance in everyday tasks.

The present research seeks to expand upon these findings in a number of ways, by exploring how such mechanisms could be translated into benefitting naturalistic memories for face-name associations. Of particular note in the case of name memory, where the face and name share no semantic information, providing a visual cue that is semantically congruent with the auditory cue may prove to be beneficial. In-person introductions inherently engage multiple senses in a cross-modal associative learning task, the association between a name and a face. However, each component of the association is presented in only one modality—the face is visual, and the name is auditory. To bolster memory performance for the association between a face and name, it could be beneficial to enhance each of those components by making it multisensory, and thus creating a multisensory representation for each component. While there is not a simple way to transform seeing a face into a multisensory experience (short of touching a face, which is seldom socially acceptable), the auditory presentation of the name (i.e., the spoken name) could be augmented with a visual representation, by for example, the addition of a name tag. When name tags are presented in one's native language, they provide a visual component to an introduction that is congruent specifically with the auditory information being given. Name tags thus provide a natural correspondence with the spoken name and are an ideal cue for testing whether multisensory stimulus presentation can aid with associative memory tasks.

Here, in a series of experiments, we systematically investigate the role of multisensory presentations in associative memory. We present a multisensory representation of a name through the use of vocalized names and congruent name tags, to see if a multisensory stimulus presentation would aid face-name memory. In Experiment 1, we test if presenting a name tag during an introduction will improve memory for names when participants are later probed with previously-encountered faces. In Experiments 2–4, we alter initial stimulus presentation to rule out the influence of visual text guiding attention, lip reading, and duration of time spent with the name on cued recall improvement when a name tag is provided. In Experiment 5, we test whether the synchrony of the auditory name and the visual tag—with synchrony being an important factor in multisensory integration—is useful above merely providing more information, to investigate if the memory improvement is mediated by multisensory integration. If multisensory stimulus presentation is generally helpful for this process, then performance with congruent and synchronous name tag presentation with the name should improve memory performance above and beyond the baseline in each experiment.

2 | EXPERIMENT 1

In this experiment, we examined whether name tags can improve memory of names using a within-subject design in which during the

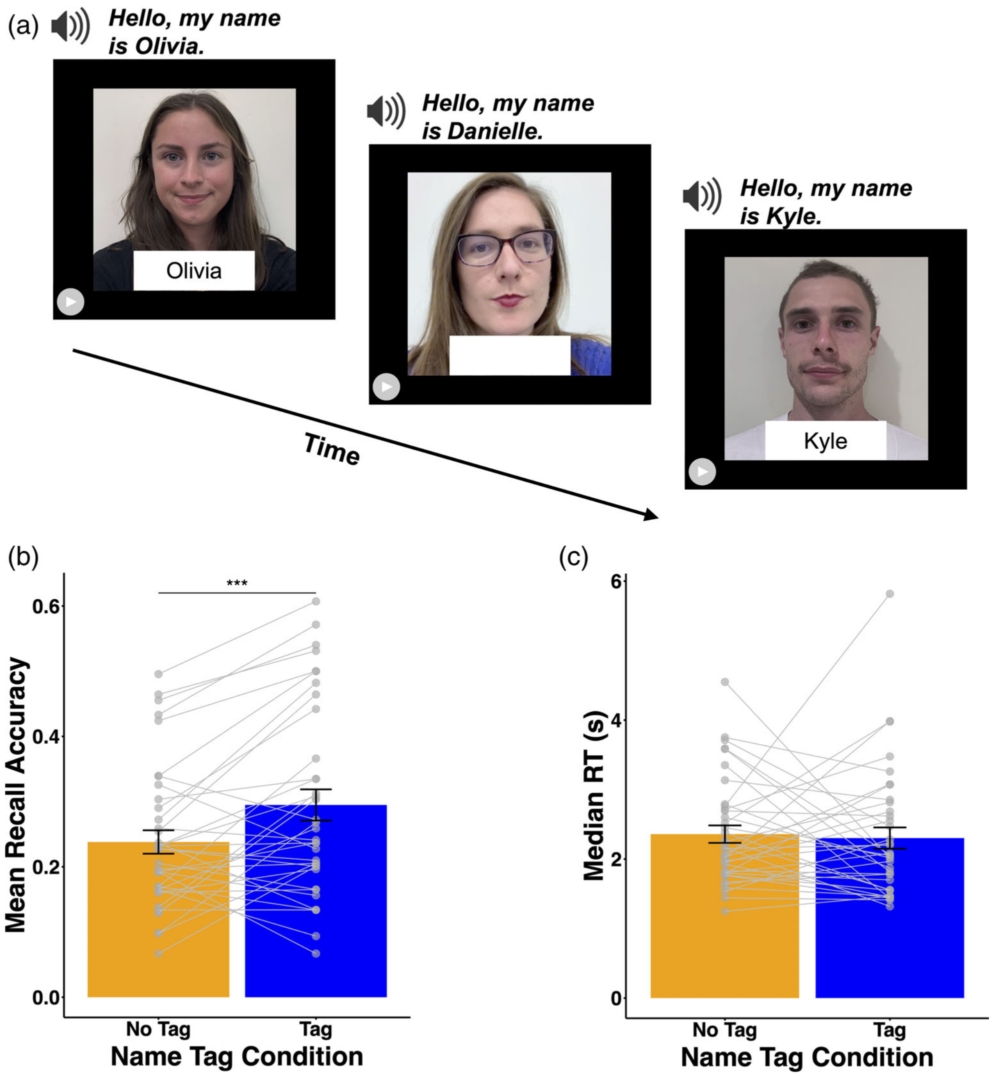

FIGURE 1 Experiment 1. (a) A diagram of the encoding procedure. Participants were presented with 15 videos per block, half with a name tag and half without. (b) Recall performance. There was a significant main effect such that participants recalled a higher proportion of names presented with a name tag compared to those presented without. The overlaid scatter plot represents individual participant scores. (c) Reaction time (for correct responses, in seconds) for the recall task measured as the first keystroke made in the response. Error bars are standard errors

encoding phase half of the trials included a name tag and half of them did not. We hypothesized that the addition of this visual information would improve the recall of names.

2.1 | Methods

2.1.1 | Participants

Participants were 38 undergraduate students (22 females) at the University of California, Los Angeles. Average participant age was 19.49 years ($SD = 1.07$), and all reported normal or corrected-to-normal sight and hearing, except for one participant who reported that they did not have corrected-to-normal sight, but reported no difficulty observing the stimuli on the computer screen and were thus included in the analyses. Additionally, 30 of these participants were native English speakers. The remaining 8 were fluent in English. Initial analyses indicated that the results did not differ if non-native speakers were excluded, so those participants were kept in the analyses for this and the follow-up experiments. Two participants were excluded from analyses due to computer errors resulting in incomplete session data.

Written informed consent was obtained from each participant and experimental procedures were reviewed and approved by the UCLA Institutional Review Board.

2.1.2 | Materials

Experimental stimuli were 60 brief video clips (1–2 s duration) of young adults (age 18–22; half male, half female) captured from the chest up against a white background. In each video, the speaker introduced themself with the phrase "Hello, my name is [name]." Names presented during these videos were selected from the most common first names given to male and female children in the United States between 1990 and 1999 as reported by the United States Social Security Administration, so all of the names would have similar familiarity to participants.

A white rectangle acting as a name tag was placed over the chest and neck of each individual video, but did not obscure the mouth. This remained in the same location for the duration of the experiment. During half of the trials, this rectangle remained blank, presenting no additional name information to the participants (the “no tag” level of the name tag condition). In the other half of the trials, black text spelling the name given in the video was presented for the duration of the video in this white rectangle (the “tag” level of name tag condition). See Figure 1 for examples of both conditions.

Experimental stimuli were presented using PsychoPy software (Peirce et al., 2019) on a Mac Mini computer.

2.1.3 | Procedure

The 60 videos were presented to participants across 4 blocks. During each block, participants were shown 15 videos with a mix of genders. As this meant that there were an uneven number of trials in each block, the first and third block had 8 tag trials, and the second and fourth blocks had 7 tag trials. Individuals within each block were presented in a random order, and the order of these blocks of individuals were also randomized.

In each block, participants were first given an encoding phase (Figure 1a), where they were presented with each of the 15 videos in that block. To ensure participants were attending to the videos, they were asked to make a button-press response to report the gender of the speaker after each video, using “1” to indicate a male speaker and “2” to indicate female speaker. Participants were not informed that they would be tested later on their memory of the names. After seeing and reporting the gender for all 15 videos in the block, participants were given a 3-min break during which they were asked to close their eyes and relax. At the end of this delay, participants were given a cued recall test of the name. They were presented with a still image from each of the videos they had seen before the delay, in a randomized order, and prompted to type in the name they remembered being associated with that person. Still images were created from the final frames of each video, and were selected such that the faces had closed lips, to remove any facial cues for sounds in the name. Participants were given 10 s to recall and type the name; after the 10 s, the experiment would advance to the next question.

After each cued recall attempt, participants were asked to rate their confidence in their memory for the name on a scale from 1 (low confidence) to 4 (high confidence). After being tested on all 15 names and providing confidence ratings, participants were given a 1-min break before moving on to the next block.

2.1.4 | Analysis

Participant responses were rated by three blind raters for correctness, as well as by computer test-matching. The human- and computer-based scoring did not qualitatively alter the results, so human ratings were used to allow for spelling errors and alternative name spellings. Human raters were instructed to rate a response as correct if the typed response was an alternative spelling of a name, if an answer was cut off by the response time limit and could not reasonably be mistaken for another name, or if the name typed was a shortened version of the correct name that could not be mistaken for another name. If any of the raters judged a participant's attempt as correct, the response was marked as correct for final analysis. Raters largely agreed with one another, such that all three raters matched their judgments on 98% of responses.

Reaction time (RT) was collected for each key press in the typed response, and was analyzed using the first key input from the participant. As reaction times were non-normal in their distribution, median values on correct trials were used in the analyses. In the case where participants did not have any correct responses in one condition, they were removed from pair-wise analyses.

Data used in these analyses have been made available in a GitHub repository (https://github.com/murray-carolynA/Data_MultisensoryNametagStudy).

2.2 | Results

Results from the attention check (i.e., the gender judgment task) during the encoding phase showed high accuracy across all participants ($M = 98.6\%$, $SD = 2.4\%$), indicating they were attending to the stimuli at encoding. As such, all trials were included in the final analyses.

Initial analyses showed that performance differences between name tag conditions persisted across blocks, regardless of whether participants did not know their memory would be tested (as in block 1) or if they did (all subsequent blocks), so analyses collapsed performance across blocks. (Block-wise analysis of performance for this and all following experiments have been included in the Data S1.) Descriptive statistics of participant performance are printed in Table 1. Pairwise one-way t-test comparison of accuracy between the two conditions showed superior recall performance in the tag condition over the no-tag condition ($t(36) = 3.59$, $p < .001$, Cohen's $d = 0.44$); calculated as in Cohen, 1988, such that

$$d = \frac{M_{\text{tag}} - M_{\text{notag}}}{\sqrt{\frac{SD_{\text{tag}}^2 + SD_{\text{notag}}^2}{2}}}.$$

Pairwise one-way t-test comparison of RT for correct responses showed no significant effect of name tag condition ($t(36) = 0.32$, $p = .75$). The results were the same when including all trials.

Confidence results generally tracked the accuracy data across experiments, and can be found in the Data S1.

2.3 | Interim discussion

Results from this experiment indicate that participants do perform better when they are given a semantic visual cue congruent with the

TABLE 1 Descriptive statistics for accuracy and reaction time, Experiments 1–5

		Recall accuracy	Recall RT	Recognition accuracy	Recognition RT
Experiment 1	Name tag	0.29 (*0.15*)	2.30 (*0.93*)	N/A	N/A
	No tag	0.24 (*0.11*)	2.36 (*0.76*)	N/A	N/A
Experiment 2	English tag	0.25 (*0.13*)	2.21 (*0.59*)	N/A	N/A
	Armenian tag	0.19 (*0.11*)	2.67 (*1.00*)	N/A	N/A
Experiment 3	Name tag	0.27 (*0.15*)	2.11 (*0.79*)	N/A	N/A
	No tag	0.19 (*0.13*)	2.71 (*1.33*)	N/A	N/A
Experiment 4	Name tag	0.31 (*0.13*)	2.25 (*0.45*)	0.54 (*0.15*)	2.96 (*0.81*)
	No tag	0.27 (*0.14*)	2.34 (*1.03*)	0.51 (*0.15*)	3.04 (*0.91*)
Experiment 5	Synchronous	0.43 (*0.17*)	3.09 (*0.98*)	0.64 (*0.17*)	2.40 (*0.76*)
	Asynchronous	0.39 (*0.14*)	3.24 (*1.12*)	0.62 (*0.13*)	2.67 (*0.85*)

Note: Means for accuracy are reported as a proportion correct and reaction times (RT) is reported as the average of the median response times for correct responses in seconds. Standard deviations for each measure are provided in parentheses.

auditory stimuli, even if these stimuli were not available at the time of retrieval. This would seem to generally support the utility of multisensory stimulus presentation for this type of recall task.

However, there are alternative explanations for the observed superiority of the tag condition. For example, It has been shown that objects presented with an accompanying irrelevant stimulus in a different modality can improve memory for that object relative to objects presented alone (Matusz et al., 2017). Alternatively, the presence of the name tag may have increased the salience of the visual stimuli, and therefore led to higher arousal in the tag condition, compared to the no tag condition that contained a blank rectangle. To investigate if the mere presence of an additional visual in the form of a name tag could explain improved performance in the tag condition, we conducted a second experiment.

3 | EXPERIMENT 2

In this experiment we investigated whether the superior memory performance in the previous experiment was due to the difference in visual salience in the two conditions. We compared the performance between two name tag conditions that had equal visual salience and only differed in the semantic content. In one condition, the name tag could be read and understood by the participants (Latin alphabet, hereafter called the English condition), and in the control condition it was written in an unfamiliar alphabet (Armenian alphabet, hereafter called the Armenian condition) that was unfamiliar and incomprehensible to the participants. If the difference in performance observed in the previous experiment was due to visual salience of the tag, then that difference should disappear in this experiment. On the other hand, if the superiority of the tag condition was due to the additional visual semantic cue, then we should observe a superior performance of the English name tag over the Armenian name tag.

3.1 | Methods

3.1.1 | Participants

Participants were 41 undergraduate students at the University of California, Los Angeles, with an average age of 19.66 years ($SD = 1.86$). Two participants were excluded from analyses because they knew individuals in the videos from everyday life by different names, leaving 39 participants in the analysis (30 female). All participants except one reported normal or corrected-to-normal vision, and all reported having normal hearing and being unable to read Armenian. The participant who reported not having corrected-to-normal vision reported no difficulty seeing the stimuli on the computer screen, and was kept in the analyses.

3.1.2 | Materials & procedure

Videos and the name tag format matched those in Experiment 1 except in the no tag condition. In this experiment, to control for the visual saliency of having a name tag, the no tag condition was replaced by a condition with a name tag written in an alphabet unfamiliar to the participants. In this half of trials, the name was written in the Armenian alphabet, so the size and shape of the letters would be similar to the names written in a familiar alphabet (see Figure 2a), but the congruency between the visual and audio signals would not be present for participants.

The procedure and data analysis matched that of Experiment 1.

3.2 | Results

Pairwise *t*-test comparison of cued recall accuracy showed superior performance in the English name-tag condition compared to the

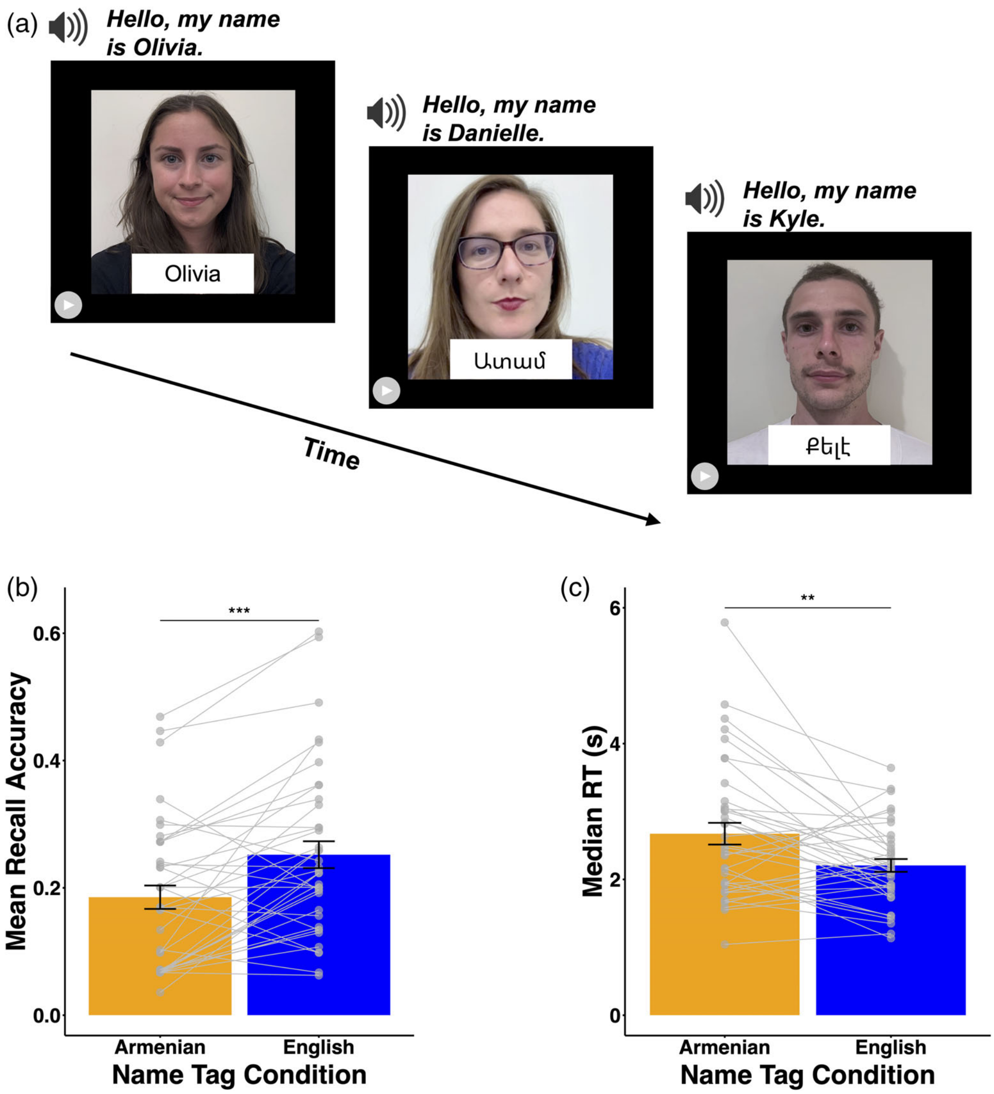

FIGURE 2 Experiment 2. (a) A diagram of the encoding procedure for experiment 2. Armenian names are presented on the trials lacking a congruent name tag to control for the visual saliency of having a name tag with writing. (b) Recall performance. Participants showed significantly higher recall for names when the face was initially presented with an English name tag than with an Armenian name tag. ***$p < .001$. The overlaid scatter plot represents individual participant scores. (c) Reaction time (for correct responses, in seconds) for the recall task measured as the first keystroke made in the response. Error bars are standard errors

Armenian name tag condition, $t(38) = 4.21$, $p < .001$, Cohen's $d = 0.54$, (Figure 2b). Median reaction time for correct responses also showed an effect of tag, $t(38) = 3.16$, $p = .003$, Cohen's $d = 0.57$, such that median response time was shorter when recalling names originally presented with an English tag than an Armenian one (Figure 2c). The results were qualitatively the same when including all trials (including incorrect responses).

3.3 | Interim discussion

Results from this experiment show that semantically congruent visual stimuli can facilitate remembering names, and rules out the role of visual salience and arousal as the underlying mechanism for this facilitation. These findings are consistent with the hypothesis of multisensory integration as an underlying mechanism for improved memory.

It is important to note that in the control conditions for the previous two experiments (no name tag or Armenian name tag) there are two cues available about the name: the acoustic cue (the voice) and the lip movement cue. Previous work has shown that lip reading may provide important multisensory cues, and can assist with disambiguating sounds (Bernstein et al., 2004). The written name (name tag) information can help encoding the face-name association in two different ways: by disambiguating (reducing the uncertainty) of the lip movement cue or by disambiguating the auditory cue. In order to gain insight into which process is occurring, we conducted the following experiment.

4 | EXPERIMENT 3

The goal of this experiment was to gain insight into the role of the lip movement cue in the facilitation effect of name tags observed in previous experiments. To investigate whether the observed facilitation effect stems primarily from the disambiguation of lip movements by the name tag (both visual cues, but one perceptual and the other semantic) videos were replaced by still images in this experiment, to remove lip reading

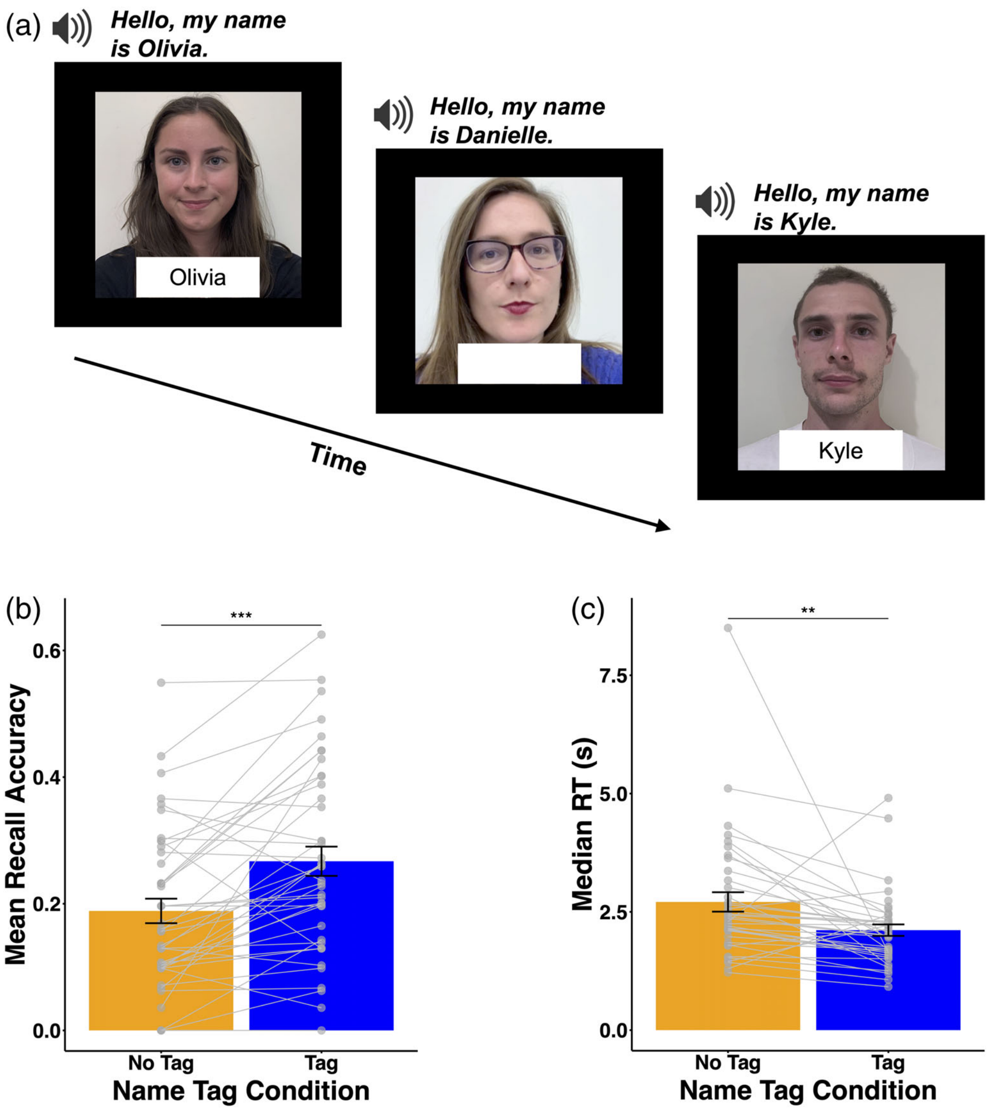

FIGURE 3 Experiment 3. (a) A diagram of the modified encoding procedure for Experiment 3 (still pictures instead of videos). (b) Recall performance. Participants showed higher average recall performance for names encoded with a name tag than for those encoded without. The overlaid scatter plot represents individual participant scores. (c) Reaction time (for correct responses, in seconds) for the recall task. Measured as the first keystroke made in the response. Error bars are standard errors

cues. To the extent that the benefit of the name tag cue stems from its interaction with the lip movement cue, in this experiment the effect should disappear or be weakened. Conversely, if the benefit of name tag stems primarily from interaction with the auditory cue or just by providing an additional source of information without interacting with the other cues, then the effect should remain the same here.

4.1 | Methods

4.1.1 | Participants

A total of 44 participants (37 female), who were all undergraduate students at the University of California, Los Angeles, were enrolled. Participants had an average age of 19.58 years ($SD = 2.11$), and all reported normal or corrected-to-normal sight and hearing. Thirty-five reported being native English speakers, and all reported being fluent in the language. Two participants were excluded from analysis due to computer issues interrupting the experiment.

4.1.2 | Materials & procedure

Materials had one major change from the preceding experiments: the video of the speaker was replaced by a still image of the individual from the video, to remove the ability of participants to use lipreading to help with the task. The images were taken from the end of each video, selected so the speakers' lips were closed and provided no cues for what sounds the individuals may have been speaking. Each image was presented for the duration of the video it was replacing, and the audio that accompanied it was taken from the original video.

The procedure and data analysis matched that of the first experiment, where the tag was either blank or had an English tag.

4.2 | Results

Accuracy results for the recall task were very similar to those of the previous experiments (Figure 3). Pairwise on-way *t*-test analyses showed higher recall for names originally presented with a name tag compared to those presented with no tag ($t[41] = 4.64$, $p < .001$, Cohen's $d = 0.57$). Median correct reaction time also showed an effect of name tag ($t[37] = 2.82$, $p = .008$, Cohen's $d = 0.55$), such that recall responses to names originally presented with a name tag were faster than those for names originally presented with no tag. The results were qualitatively the same when all reaction times were used.

4.3 | Interim discussion

Results for Experiment 3 indicate that the observed superiority of the tag condition over no tag is not due (at least entirely) to interaction with the lip-reading cue. The findings were consistent with those of Experiments 1 and 2, supporting the interpretation that multisensory mechanisms may be able to explain the improved recall performance when name tags are presented at encoding. However, it should be noted that the amount of time participants were exposed to each name differed between the two conditions: when name tags were presented, participants were aware of the name much earlier than in the no tag condition. This difference in duration could lead to improved performance from longer exposure to the visual cue, rather than any multisensory mechanisms. As such, Experiment 4 was designed to keep name exposure times equal between the tag and no tag conditions.

5 | EXPERIMENT 4

The objective of this experiment was to equate the duration of time in which the name of the speaker is available to the participant across conditions to test whether the observed superiority of the tag condition was due to the longer duration of the name information being available in the tag condition.

5.1 | Methods

5.1.1 | Participants

Participants for this experiment were 49 undergraduate students (39 female) at the University of California, Los Angeles, with a mean age of 19.06 years ($SD = 0.87$). All participants reported normal or corrected-to-normal vision and normal hearing, and 7 reported being non-native speakers of English, but were fluent and so kept in for analyses.

5.1.2 | Materials & procedure

Videos and the name tag format matched those in Experiment 1, except all videos were cut such that the introduction ("Hello, my name is") was removed, leaving only the name. This meant that the tag and video were on screen for only the duration of the stated name.

The experimental procedure was similar to those of Experiment 1, with a few notable changes. As task performance had been, overall, somewhat low in previous experiments, the number of names participants were asked to learn per block was reduced from 15 to 10, and the number of blocks increased from 4 to 6. Moreover, at test, the confidence rating task was replaced by a recognition memory task, to probe if recognition would benefit from multisensory encoding as well as recall. Participants were given the same 10 s to type a response to the recall prompt as in Experiment 1, and then given a 5-alternative multiple-choice recognition test for the name, using the same image as a prompt. The 5 names selected for the recognition test included the correct name and 4 alternatives that had been presented in the same block. To ensure this task would not be trivial and 5 probable names would exist, blocks now consisted of the same gender of speaker in all videos, resulting in 3 blocks of female and 3 blocks of male speakers. The assignment of male or female speakers to blocks was pseudorandom between participants, as was block order. As all of the speakers within one block were of the same gender, the encoding task of recognizing the gender was removed for this experiment.

5.2 | Results

Recall results in this experiment (Figure 4b–c) largely follow those of the previous experiments: pairwise one-way t-test results showed that average cued recall performance was higher for names originally presented with a name tag than for names presented without a tag ($t[49] = 3.11$, $p = .003$, Cohen's $d = 0.32$). Median correct reaction time in the recall task showed no significant effect of name tag ($t[49] = 0.59$, $p = .56$). Results were qualitatively the same when using all trials.

On the recognition task, there was a marginal effect of name tag condition ($t[49] = 1.83$, $p = .07$, Cohen's $d = 0.22$), such that names originally presented with a tag were remembered more often than names originally presented with no tag. There was no significant difference between name tag conditions on median correct response time in the recognition task ($t[49] = 0.75$, $p = .45$). These results were qualitatively the same when analyzed using response times from both correct and incorrect trials.

5.3 | Interim discussion

Experiment 4 further supports that the addition of a visual stimulus congruent with auditory stimulus improves performance in cued recall for names, even if the presentation of the congruent visual stimulus matches the length of the auditory stimulus. Interestingly, recognition of the names does not show a similar benefit in accuracy, though the data trends such that recognition accuracy is somewhat higher for faces originally presented with a tag compared with those that were not.

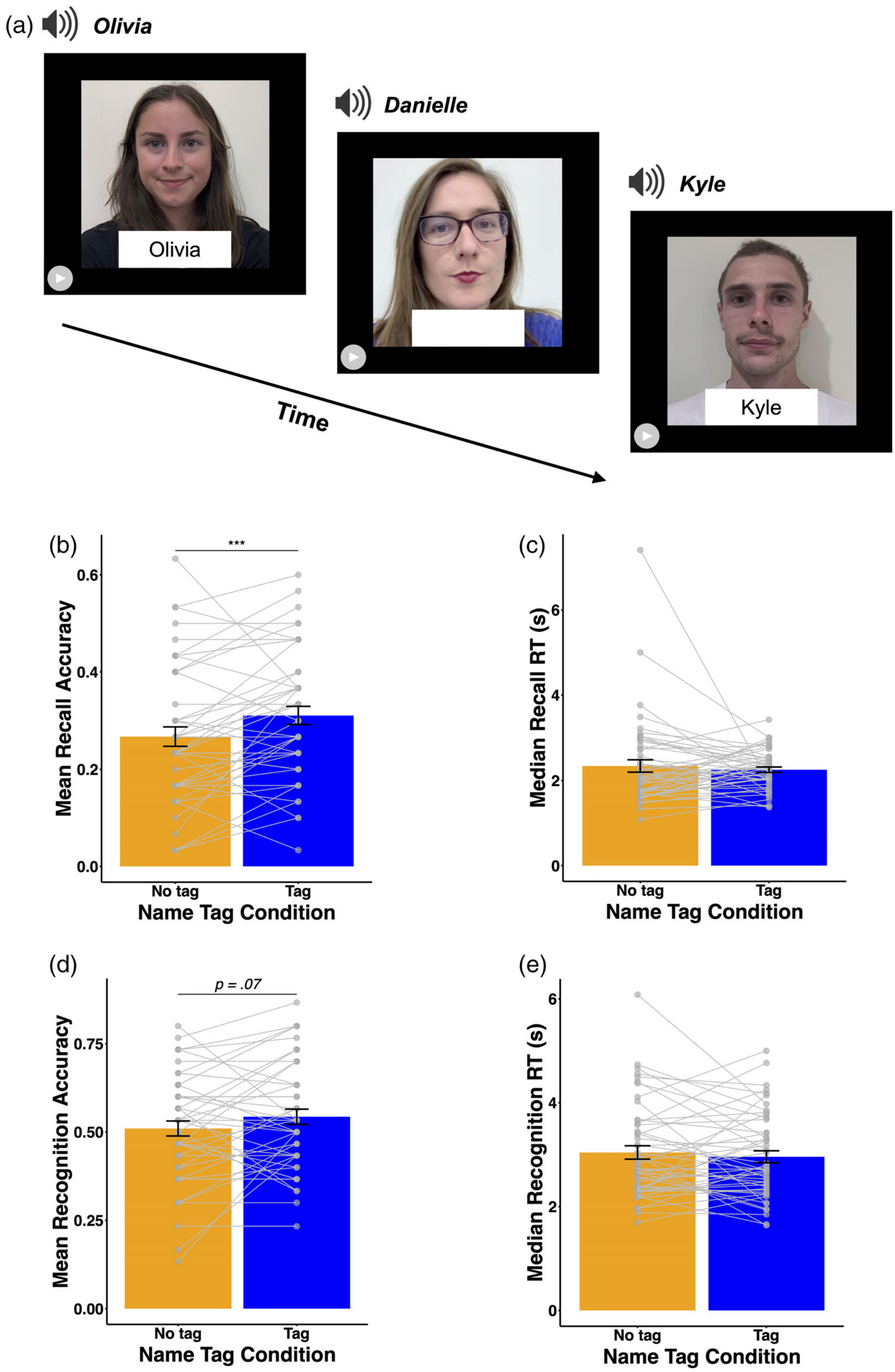

FIGURE 4 Experiment 4. (a) A diagram of the modified encoding procedure for Experiment 4. This design ensures the onset time and subsequent presentation duration of the name tag and stated name are matched. (b) Recall performance. There was a main effect of name tag, such that presenting a name tag produced higher average recall than if no tag was present during encoding. The overlaid scatter plot represents individual participant scores. (c) Reaction time (for correct responses, in seconds) for the recall task measured as the first keystroke made in the response. (d) Recognition performance. Participants showed no significant difference in recognition performance based on the tag condition. (e) Reaction time (for correct responses, in seconds) for the recognition task. Error bars are standard errors

6 | EXPERIMENT 5

Experiments 1–4 establish that addition of a name tag improves the recall of names. However, two distinct underlying mechanisms could mediate this facilitation (see Figure 5). One possibility is that the name tag could be serving as an additional memory trace that would aid recall by providing a second redundant route to the desired information (i.e. the name), Figure 5a. Alternatively, the tag cue provides a multisensory representation of name by combining with the audio (and maybe also lip movements; Figure 5b) and a richer encoding of name-face association. To tease apart these two potential mechanisms, in this fifth experiment we compared two conditions that were equal in the number of “traces” during encoding, but one condition allows for multisensory integration to occur, whereas the

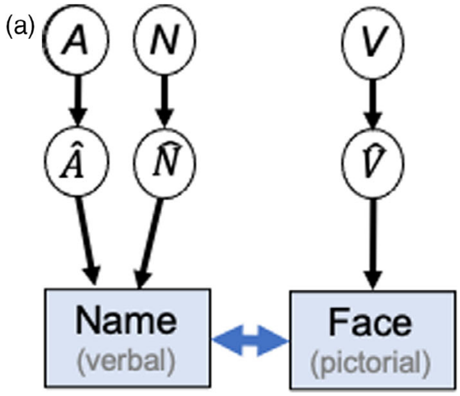

FIGURE 5 Two possible mechanisms underlying the name tag facilitation of face-name associative memory. *A* denotes auditory input provided by the spoken name. *N* denotes the visual input (text) provided by the name-tag. *V* denotes the visual input provided by the video of the individual. *A*, *N*, and *V* are noisy sensory inputs, and $\widehat{A}$, $\widehat{N}$, and $\widehat{V}$ represent the perceptual estimates. (a) If multisensory mechanisms are not utilized, $\widehat{A}$ (spoken name) and $\widehat{N}$ (written name) independently provide information about name (verbal representation), without interacting with each other perceptually. (b) If multisensory mechanisms are utilized, both $\widehat{A}$ and $\widehat{N}$ provide information about name and each now provide an improved estimate of name due to integration with the other sensory stimulus, as depicted by arrows from both sensory stimuli to each of the perceptual estimates

other condition does not. This was achieved by manipulating the relative timing of the cues, because it is well established that temporal congruency between cues plays an important role in integration of cues (see Calvert & Thesen, 2004; Shams & Kim, 2010). In both conditions, the same cues (video, audio, name tag) were presented, however, in one condition the audio and tag were presented simultaneously, and in the other condition the tag followed the audio with a delay that is expected to disrupt integration. If the benefits of name tags derive exclusively from their provision of an additional memory trace, then we would expect to see equal performance across conditions. In contrast, a multisensory framework would predict that simultaneity between the audio and congruent visual stimuli would be necessary to receive a memory benefit, and therefore we should see better performance in the synchronous condition.

6.1 | Methods

6.1.1 | Participants

A total of 38 participants (24 female), who were all undergraduate students at the University of California, Los Angeles, were enrolled. Participants had a mean age of 20.89 years ($SD = 3.36$), and all reported normal or corrected-to-normal sight and hearing. Thirty-five reported being native English speakers, and all reported being fluent in the language. Four participants were excluded from analysis due to remembering zero names in either condition during any block of the experiment.

6.1.2 | Materials & procedure

Materials were the still images from Experiment 3, as these reduce the influence of congruency between lipreading and the visual name tag from playing a role in participants performance. Still images were presented for 5.5 s with the audio from the original videos played starting at the visual stimulus onset. At the bottom of the image, placed over the neck and torso for the duration of the stimulus, as in Experiments 1–4, was a white rectangle. Both name tag conditions in this experiment present a name tag, and differ in when the tag is displayed: synchronously with the name, or asynchronously. In the synchronous condition, the name is visible starting simultaneously with the still image and audio, and, in the asynchronous condition, the name is visible beginning 2.5 s after the start of the presentation of the still image. In both cases, the visual name will be presented for 2.5 s.

Blocks are organized as in Experiment 4: a total of 6 blocks containing 10 same-gender speakers and names to remember in each block, with tests of both cued recall and recognition for the names given after a 3-min delay.

6.2 | Results

Recall results in this experiment (Figure 6) largely follow those of the previous experiments: a pairwise one-way *t*-test showed that recall performance was higher in the multisensory synchronous condition compared to performance in the asynchronous condition ($t[33] = 2.27$, $p = .03$, Cohen's $d = 0.23$). There was no significant effect of name tag condition on median correct recall time, $t(32) = 0.82$, $p = .42$. This effect was qualitatively the same when all response times were included.

Recognition results showed no significant effect of name tag condition, $t(33) = 0.90$, $p = 0.37$. However, there was a significant effect of name tag condition on recognition response time ($t(33) = 2.27$, $p = .03$, Cohen's $d = 0.35$), such that participants on average had faster median responses to names originally presented synchronously

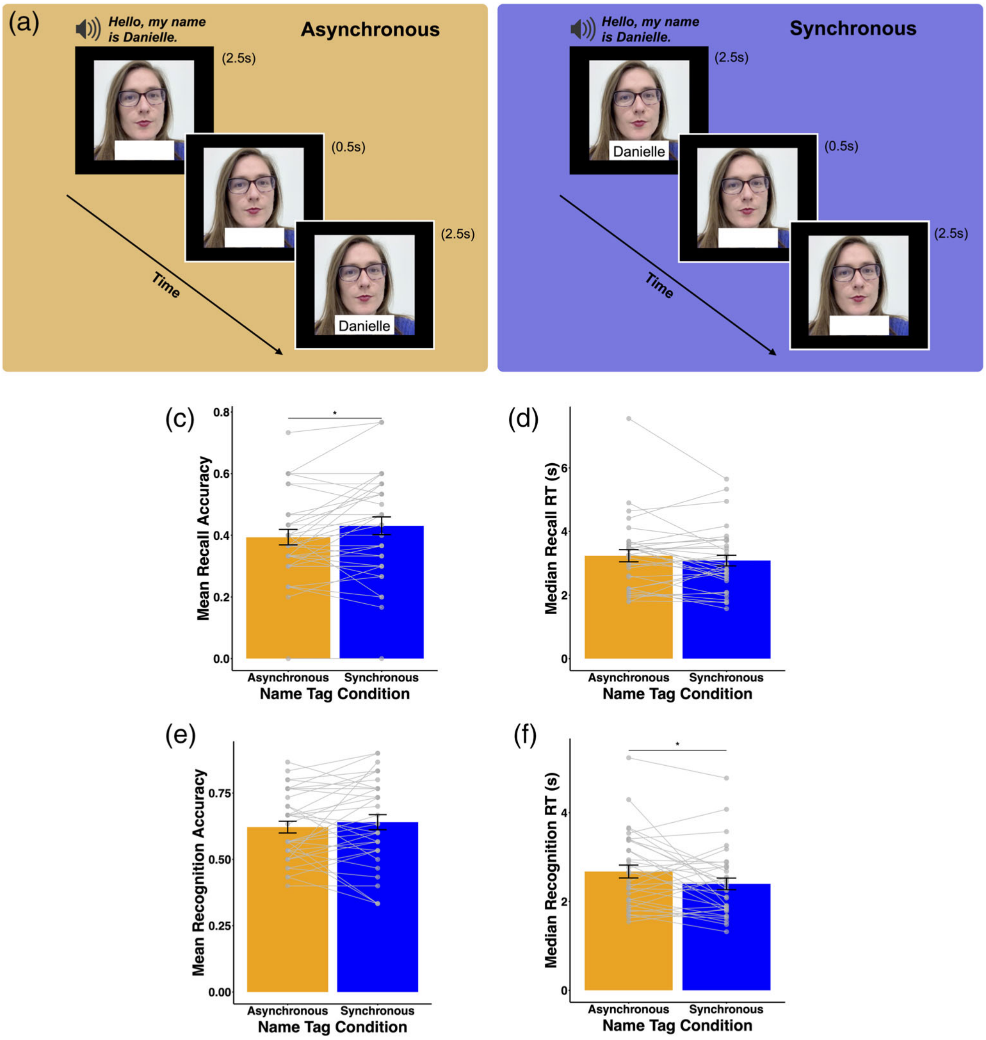

FIGURE 6 Experiment 5. (a) Schematic of asynchronous condition. The asynchronous condition presented the name tag after a 0.5 s delay. (b) Schematic of synchronous condition. The synchronous condition presented the name tag and audio synchronously. (c) Recall performance. There was a significant main effect of condition on recall. The overlaid scatter plot represents individual participant scores. (d) Reaction time (for correct responses, in seconds) for the recall task, measured as the first keystroke made in the response. (e) Recognition performance. There was no main effect of condition on recognition. (g) Reaction time (for correct responses, in seconds) for the recognition task. Error bars are standard errors

than asynchronously. These results were qualitatively different when all response trials were included, such that there was no significant difference between name tag conditions when correct and incorrect responses were used ($t(33) = 1.31$, $p = .20$).

7 | DISCUSSION

In this study, we investigated memory of people's names using naturalistic stimuli of videos in which speakers introduced themselves, as is often the case in daily life. Remembering people's names in this context amounts to an associative memory task in which the brain encodes an association between a visual face/body and an auditory presentation of a name (although the lip movements of the speaker may also contribute to this encoding). Because there is no inherent relationship between one's name and one's face, the learning and retention of this association is non-trivial, especially when tasked with the learning of multiple face-name pairs within a short period of time, which is often the case when we attend a party or a professional event.

A few previous studies have shown that multisensory encoding of objects or object features (e.g., motion, or voice) facilitates learning (e.g., Kim et al., 2008; Seitz et al., 2006; Shams et al., 2011; von Kriegstein & Giraud, 2006) and episodic memory (e.g., Heikkilä et al., 2015; Lehmann & Murray, 2005; Moran et al., 2013). However, these learning and memory tasks involved processing of a single feature or recognition of an object or object feature, and did not involve memory of an association. Here, we examined whether the benefit of multisensory encoding extends to associative memory. Specifically, we asked whether a multisensory encoding of a name can aid people's ability to bind that name to a face. To render the encoding of a name multisensory, we added a written representation of the name in the form of a name tag in addition to the auditory introduction given by the speaker. We then compared the memory of names in the presence and absence of name tags.

Across a series of five experiments, we found that participants, when cued with a face, were more likely to remember the associated name when that name had been encoded with a name tag, compared to when no name tag was provided. Experiment 1 showed a robust superiority of the tag condition (effect size 0.44) over the no tag condition. Experiment 2 examined whether the observed effect in Experiment 1 was due to the difference in visual saliency of the two conditions (blank vs. text below the face) by controlling for the visual saliency. In both conditions name tags were presented, but in one condition they were in English and congruent with the spoken name, and in the other condition they were in Armenian (a language that could not be understood by participants) and not congruent with the spoken English name. The English tag condition resulted in superior cued recall performance compared to the unintelligible name tag (effect size 0.54), ruling out that the difference in performance was due to visual saliency. Experiment 4 examined whether the observed effect in the earlier experiments was due to the fact that name information was available to the observers throughout the trial whereas the name information conveyed by voice was only available for a portion of trial duration. In that experiment, the presentation of the name tag during the trial was cut and matched the duration of the vocalization of the name. The name tag advantage effect persisted, ruling out the role of the difference in duration of name information as the underlying factor. These experiments collectively establish that the presentation of name tag aids memory of names by providing an additional cue for name. However the mechanism by which this additional cue facilitates face-name memory remains unclear. Experiments 3 and 5 aimed to shed light on this question.

The name tag cue is a visual semantic cue. It can interact and disambiguate (reduce the uncertainty of) the other semantic cues, namely the vocal cue and the lip-reading cue. The lip reading cue is an impoverished cue and, as such, could benefit from disambiguation in a within-modality (vision) manner when a name tag is added, bypassing multisensory mechanisms. Therefore, we asked if the interaction between name tag and the lip movements is the primary factor underlying the observed facilitation of memory. In Experiment 3, lip movement cues were eliminated by replacing videos with static images during the encoding phase. The superiority of name tag condition over no-name tag persisted with a similar effect size (effect size 0.44 with the lip movement vs. 0.55 without lip movements), suggesting that the putative enhancement of lip reading cue by name tag cannot account for the observed effect.

Finally, we aimed to gain insight into the underlying mechanism of the name tag benefit by teasing apart the role of multiple independent memory traces (Figure 5a) versus the role of integration of multisensory cues (Figure 5b). The name tag provides an additional memory trace, which can facilitate recall by providing an alternative retrieval route to access the name when cued with the face. That is, the face might trigger the retrieval of the auditory memory of the spoken name *or* the visual memory of the written name, essentially giving participants an extra chance to succeed at recalling the name. In this framework (Figure 5a), the mere existence of an additional cue is sufficient for improved recall. On the other hand, in the multisensory encoding framework (Figure 5b), the interaction between the cues and the integrated representation of the feature/object can play a key role in the richness of the encoding, thus increasing the likelihood of later recall (Shams & Seitz, 2008). More specifically, the name tag cue can be integrated with the vocal cue, resulting in a more accurate and/or more precise representation of the name. This improved name representation can strengthen the encoding of the face-name association and lead to improved memory performance.

In order to tease apart these two possible accounts, in Experiment 5, we compared two conditions in which the number of traces were equivalent, but one condition lends itself to integration of the name tag cue with other cues, whereas the other condition does not. It is well established that temporal congruency is key in integration of sensory cues, and the lower the temporal congruency the lower the probability of integration (e.g., Calvert & Thesen, 2004; Ernst & Bülthoff, 2004; Shams et al., 2002; Shams & Kim, 2010). Therefore, by manipulating the relative timing of the name tag and the vocal (and lip movement) cues, we can influence their probability of integration. It has been shown that introducing audio and visual stimulus onset asynchronies of between 150 and 250 ms reduces audiovisual speech fusion and alters brain activity in speech-processing regions of the brain (Macaluso et al., 2004; Miller & D'Esposito, 2005; van Atteveldt et al., 2007). Therefore, it is to be expected that the name tag cue would get integrated with the other name cues when it is presented synchronously and not integrated when it is presented with a delay of 500 ms. On the other hand, in both of these conditions all of the cues are available in each trial, and by delaying the name tag relative to the video, the performance may even be expected to improve according to multiple independent memory trace account: the name information which is initially encoded by voice and lip-reading, gets reinforced by the later presentation of the name tag. The results of Experiment 5 showed that the synchronous presentation of the name tag leads to better memory performance than the asynchronous presentation. This would support the multisensory integration hypothesis, that multisensory object representation itself can be helpful to memory above what would be predicted by having multiple independent sensory traces. Future research will need to probe this question further by examining the nature of multisensory interactions that promote facilitation of memory, including which sensory combinations can facilitate memory performance, and what kinds of memory tasks will benefit from multisensory integration.

The present results cannot be accounted for by the dual-coding theory or the cognitive theory of multimedia learning, according to which the combined verbal and pictorial presentation of words facilitates memory compared to verbal-alone presentations. In the present study, in all conditions, including the baseline no tag condition both verbal (name) and non-verbal (video/image) representations are available (see Figure 5). The only difference between the experimental and control conditions is the availability of *additional* verbal information (name tag), or, in the case of Experiment 5, the relative timing of the additional verbal (name tag) information.

The improved memory accuracy under multisensory stimulus presentation conditions does not seem to be as robust in the multiple-choice recognition task compared to the recall task. Experiments 4 and 5 evaluated both name recall and name recognition in response to face cues. In Experiment 4 there was a trend for a multisensory benefit in recognition, whereas in both experiments the benefit of multisensory presentation in recall was statistically significant. While previous experiments have shown multisensory benefits in recognition tasks, those experiments were structured quite differently from the current experiment. This experiment, unlike many previous multisensory memory studies, used an associative memory task. Previous multisensory research has probed memory for single items, while the current study investigated memory for an association between a name and a face. Moreover, additional experimental power may be needed to uncover statistically significant effects in the recognition task.

Also of note are that the brain mechanisms by which multisensory stimuli benefit recall performance are unclear. The current results can speak to a few different behavioral theories, but cannot distinguish between them decisively. Previous work has indicated that cross-modal interactions allow for information distributed across multiple senses to be combined into meaningful representations. This combination of senses has been found to allow for optimal processing of sensory information and can help disambiguate noisy stimulus presentation via uncertainty reduction (one signal can disambiguate another signals, leading to the improvement in precision and/or accuracy) (see Ernst & Bülthoff, 2004; Shams & Beierholm, 2010 for overviews). Multisensory stimulus presentation may also change how attention is directed and multisensory scenes are segregated (Lewkowicz et al., 2021) at the time of encoding. Which mechanism, if any of these, supports the current findings is currently unclear, and future neuroimaging research may be able to identify the neural mechanisms supporting multisensory memory benefits.

It should also be noted that, while these experiments do provide evidence for multisensory memory benefits, further research could help directly rule out the possibility that participants were using strategies during encoding that would selectively benefit the name tag conditions. For example, it is possible that participants preferentially encode the visual tag by default, and only use the auditory information when the tag is unavailable. This would mean that participants may have a switching cost as they change strategies between trials where a tag is synchronously presented with the audio as opposed to when the tag is asynchronous or absent, and the cognitive cost of this visual-to-auditory attentional switching on no-tag trials could potentially explain the benefit of tags seen in all experiments. We believe this explanation is very unlikely, particularly given the results of Experiment 4. In that experiment, the average duration of the videos was greatly reduced, such that none were longer than 2 s, and the audio presentation of the name began immediately at the start of the trial (see Videos S1, S2, S3, S4, S5, S6, S7, S8, S9, S10 for an example). If the aforementioned task switching strategy explained the full set of results, one might expect to see that any benefit would disappear if participants were denied the time needed to assess which strategy they should use, but a difference still existed between the tag and no tag conditions in Experiment 4. However, participants were not asked explicitly to describe any strategies they were using, and therefore we cannot entirely rule out that strategizing could play some role in the observed benefit. Further research using a between-subjects design (where some participants have only tag trials and others have only no-tag trials) to prevent task switching could investigate this further. Such an investigation could also probe how each level of synchrony in Experiment 5 compares to a unisensory baseline. This has been left out of Experiment 5 to maximize experimental power in testing the underlying mechanism of the observed benefit to remembering names when a tag is present—that is, whether it was due to having more sensory cues available or if multisensory integration was specifically helpful—but future experiments could investigate this relationship.

Our findings are generally in line with previous multisensory findings, and expand those results to associative memory, and to a more naturalistic memory task. The current experiments also suggest that multisensory mechanisms can be leveraged in daily, difficult tasks to improve memory performance. These findings do not contradict previous memory theories, but rather can function as an additional tool that can be used to improve human memory in difficult situations. Traditional techniques for improving memory—including mnemonics and spatial mapping—are effective, but do require a relatively high level of sophistication and intent to employ. Using basic sensory information could be more easily and passively implemented to improve memory. This could lead to the development of new strategies, techniques, and technologies to improve everyday life and learning, even for relatively difficult associative memory tasks.

ACKNOWLEDGMENTS

This work was supported by the UCLA Faculty Senate intramural grant. We would like to thank Ashkan Arjomandi, Fatima Iqbal, Chenyang Lin, Golbarg Rahimi, Abheri Setlur, and Sarah Thraen for rating participant responses.

FUNDING INFORMATION

This work was supported by the UCLA Faculty Senate intramural grant.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

The data for these studies are available at https://github.com/murray-carolynA/Data_MultisensoryNametagStudy, and all other materials may be obtained upon request.

INFORMED CONSENT STATEMENT

The authors affirm that human research participants provided informed consent for publication of the images in Figures 1–4 and 6. (Participants recorded for the stimuli did not provide consent for their pictures to be used in publication. Thus, Figures 1–4 and 6 use other individuals staged to resemble the experimental stimuli who did consent for their images to be used.)

ORCID

Carolyn A. Murray https://orcid.org/0000-0003-2091-6094
Jesse Rissman https://orcid.org/0000-0001-8889-5539

REFERENCES

Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication*, *44*(1), 5–18. https://doi.org/10.1016/j.specom.2004.10.011

Brédart, S. (2019). Strategies to improve name learning: A review. *European Psychologist*, *24*(4), 349–358. https://doi.org/10.1027/1016-9040/a000363

Brooks, J. O., Friedman, L., Gibson, J. M., & Yesavage, J. A. (1993). Spontaneous mnemonic strategies used by older and younger adults to remember proper names. *Memory*, *1*(4), 393–407. https://doi.org/10.1080/09658219308258245

Butler, J. S., Foxe, J. J., Fiebelkorn, I. C., Mercier, M. R., & Molholm, S. (2012). Multisensory representation of frequency across audition and touch: High density electrical mapping reveals early sensory-perceptual coupling. *Journal of Neuroscience*, *32*(44), 15338–15344. https://doi.org/10.1523/JNEUROSCI.1796-12.2012

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1), 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, *19*(5), 619–636. https://doi.org/10.1002/acp.1101

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, *4*(2), 187–197. https://doi.org/10.1111/j.2044-835X.1986.tb01010.x

Cohen, J. (1988). The t test for means. In *Statistical power analysis for the behavioral sciences* (2nd ed., pp. 19–74). L. Erlbaum Associates.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. https://doi.org/10.1016/j.tics.2004.02.002

Heikkilä, J., Alho, K., Hyvönen, H., & Tiippana, K. (2015). Audiovisual semantic congruency during encoding enhances memory performance. *Experimental Psychology*, *62*(2), 123–130. https://doi.org/10.1027/1618-3169/a000279

Heikkilä, J., & Tiippana, K. (2016). School-aged children can benefit from audiovisual semantic congruency during memory encoding. *Experimental Brain Research*, *234*(5), 1199–1207. https://doi.org/10.1007/s00221-015-4341-6

Kim, R. S., Seitz, A. R., & Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One*, *3*(1), e1532. https://doi.org/10.1371/journal.pone.0001532

Lacey, S., Tal, N., Amedi, A., & Sathian, K. (2009). A putative model of multisensory object representation. *Brain Topography*, *21*(3), 269–274. https://doi.org/10.1007/s10548-009-0087-4

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, *166*(3), 289–297. https://doi.org/10.1007/s00221-005-2370-2

Lee, H., Stirnberg, R., Stöcker, T., & Axmacher, N. (2017). Audiovisual integration supports face–name associative memory formation. *Cognitive Neuroscience*, *8*(4), 177–192. https://doi.org/10.1080/17588928.2017.1327426

Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Cognitive Brain Research*, *24*(2), 326–334. https://doi.org/10.1016/j.cogbrainres.2005.02.005

Lewkowicz, D. J., Schmuckler, M., & Agrawal, V. (2021). The multisensory cocktail party problem in adults: Perceptual segregation of talking faces on the basis of audiovisual temporal synchrony. *Cognition*, *214*, 104743. https://doi.org/10.1016/j.cognition.2021.104743

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: A PET study. *NeuroImage*, *21*(2), 725–732. https://doi.org/10.1016/j.neuroimage.2003.09.049

Matusz, P. J., Wallace, M. T., & Murray, M. M. (2017). A multisensory perspective on object memory. *Neuropsychologia*, *105*, 243–252. https://doi.org/10.1016/j.neuropsychologia.2017.04.008

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

McCarty, D. L. (1980). Investigation of a visual imagery mnemonic device for acquiring face–name associations. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 145–155. https://doi.org/10.1037/0278-7393.6.2.145

McWeeny, K. H., Young, A. W., Hay, D. C., & Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology*, *78*(2), 143–149. https://doi.org/10.1111/j.2044-8295.1987.tb02235.x

Miller, L. M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience*, *25*(25), 5884–5893. https://doi.org/10.1523/JNEUROSCI.0896-05.2005

Moran, Z. D., Bachman, P., Pham, P., Hah Cho, S., Cannon, T. D., & Shams, L. (2013). Multisensory encoding improves auditory recognition. *Multisensory Research*, *26*(6), 581–592. https://doi.org/10.1163/22134808-00002436

Morris, P. E., Fritz, C. O., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for learning proper names: Expanding retrieval practice, meaning and imagery. *Applied Cognitive Psychology*, *19*(6), 779–798. https://doi.org/10.1002/acp.1115

Murray, M. M., Michel, C. M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage*, *21*(1), 125–135. https://doi.org/10.1016/j.neuroimage.2003.09.035

Neuschatz, J. S., Preston, E. L., Toglia, M. P., & Neuschatz, J. S. (2005). Comparison of the efficacy of two name-learning techniques: Expanding rehearsal and name-face imagery. *The American Journal of Psychology*, *118*(1), 79–102.

O'Mahony, C., & Newell, F. N. (2012). Integration of faces and voices, but not faces and names, in person recognition. *British Journal of Psychology*, *103*(1), 73–82. https://doi.org/10.1111/j.2044-8295.2011.02044.x

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology*, *16*(14), 1422–1427. https://doi.org/10.1016/j.cub.2006.05.048

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*(9), 425–432. https://doi.org/10.1016/j.tics.2010.07.001

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews*, *7*(3), 269–284. https://doi.org/10.1016/j.plrev.2010.04.006

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006

Shams, L., Wozny, D. R., Kim, R. S., & Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology*, *2*, 264. https://doi.org/10.3389/fpsyg.2011.00264

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, *28*(2), 61–70. https://doi.org/10.1250/ast.28.61

Thelen, A., Talsma, D., & Murray, M. M. (2015). Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition*, *138*, 148–160. https://doi.org/10.1016/j.cognition.2015.02.003

van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(4), 962–974. https://doi.org/10.1093/cercor/bhl007

von Kriegstein, K., & Giraud, A.-L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biology*, *4*(10), e326. https://doi.org/10.1371/journal.pbio.0040326

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Murray, C. A., Tarlow, M., Rissman, J., & Shams, L. (2022). Multisensory encoding of names via name tags facilitates remembering. *Applied Cognitive Psychology*, 1–15. https://doi.org/10.1002/acp.4012