PUTTING THE NAME TO THE FACE: IMPROVING NAME–FACE MEMORY ASSOCIATIONS THROUGH CONCEPTS OF BIZARRENESS AND PRODUCTION

BY
DANIEL NIENHUIS

A Graduate Thesis
submitted to the Department of Psychology. Neuroscience & Behaviour in Partial Fulfillment of the Requirements for a Master's of Experimental Psychology Degree

McMaster University

2021-08-18

MASTER OF SCIENCE (PSYCHOLOGY; 2021)
MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: PUTTING THE NAME TO THE FACE: IMPROVING NAME–FACE MEMORY ASSOCIATIONS THROUGH CONCEPTS OF BIZARRENESS AND PRODUCTION

AUTHOR: Daniel Nienhuis

PRINCIPLE INVESTIGATOR: Dr. Sandra Monteiro

CO-INVESTIGATORS: Dr. Hanae Davis, Dr. Lisa Lorentz, Brendan Stanley

NUMBER OF PAGES: 46

Abstract

Memory for names and faces is especially unpredictable. We have all experienced the embarrassment of failing to remember an individual's name despite meeting them only minutes before. Many theories have been suggested to explain these frustrating moments (Burton & Bruce, 1993; O'Mahony & Newell, 2012). However, despite value of these theories, many neglect to explain how to improve name–face memory. Mnemonic devices are powerful memory aids that improve encoding and subsequent memory recall (McCabe et al., 2013). Given the effectiveness of mnemonics, name–face memory recall may be improved with a novel mnemonic device.

Previous research by Patel (M.Sc.) at McMaster University investigated one such mnemonic device: the house bunny effect (HBE). The HBE combines elements of the bizarreness effect (Cox & Wollen, 1981) and the production effect (Quinlan & Taylor, 2013) and predicts that repeating an individual's name in a bizarre voice at the time of encoding (i.e., when meeting a new individual) improves name–face recall. However, contrary to this prediction, bizarre name production did not improve name–face memory recall compared to a normal voice production control. More importantly, a non-statistically significant trend in the opposite direction was observed—bizarre name production at the time of encoding hindered name–face memory recall (Patel, 2020).

Given this finding, we present two studies: an online conceptual replication of the HBE and a study that further elucidates the mechanisms behind the HBE. The results of these two studies will help determine how name–face memory is influenced by bizarre name production at the time of encoding (i.e., the HBE mnemonic).

Acknowledgements

I would like to thank Dr. Sandra Monteiro, Dr. Lisa Lorentz, Brendan Stanley, and Dr. Hanae Davis for their unshakable support and guidance. Without the support of these incredible mentors and colleagues, this project never would have been completed in a single year. I would also like to thank the undergraduate research assistants for their willingness and persistence in collecting the data. I will be forever grateful to everyone who has supported me and pushed me to complete my one-year Master's.

Table of Contents

List of Figures & Tables

Introduction

Imagine you are at a networking event: your singular goal is to remember the names and faces of new people. Throughout the course of the event, a dozen people have introduced themselves to you; however, much to your own frustration, the moment each of these people leaves the conversation, it feels as if they take their names with them. Trying to put their name to their face becomes an arduous and embarrassing guessing game. After all, you just met this person a few minutes ago, how could you have possibly forgotten their name? As annoying as these encounters are, we may find solace in the fact that many people struggle with name–face memory—we all have anecdotal experience to support this observation. Given this common memory deficit, many theories have been produced to explain the difficulty in recalling a name for a given face. These theories range from abstract cognitive models of name–face memory formation (e.g., Interactive Activation Model; Burton & Bruce, 1993), to perceptual observations (O'Mahoney & Newell, 2012). No matter the theory or the branch of psychology, one shared idea is that names are not directly associated or connected with a face. Names carry no predictive or descriptive value of an individual; therefore, names are not connected to a face. This mental segregation of names and faces is an ideal target to improve name–face memory. Can we improve name–face memory by enhancing the mental connection between a name and a face? Mnemonic devices are popular memory aids that improve encoding and memory retrieval. Name–face memory associations may be improved through the application of mnemonic devices.

Mnemonic devices improve encoding and aid in memory retrieval. Memory athletes (individuals who compete against each other to memorize and recall different types of information under certain time constraints) rely on the memory benefits of mnemonics (Dresler et al., 2017). For example, memory athletes often use the "Memory Palace" mnemonic, also known as the method of loci (MoL; McCabe, 2015), to remember serially presented lists. The MoL requires users to place each to-be-remembered item along a mental pathway. If an athlete is required to remember the order of a deck of 52 playing cards, then they might place each card along a mental representation of a nature trail they regularly walk. These athletes then mentally walk through their pathway to recall the order of the cards. The MoL works best for serially presented items (e.g., a grocery list, deck of cards; McCabe, 2015), but additional mnemonic devices are less restrictive. Chunking and the keyword method are two additional mnemonic examples that can be applied to less ordered information. Chunking describes the process where large groups of information are broken down into smaller groups of closely related items (Gobet et al., 2001). For example, attempting to remember a phone number like 15194533081 is difficult, but separating this number into 1-519-453-3081 is much easier to remember. Chunking can be applied to a wide variety of information, including sentences, words, and images. Lastly, the keyword method is a mnemonic device best used for remembering novel terms (McDaniel & Pressley, 1984). This method requires the user to create a mental connection between the definition of the to-be-remembered word, and a *keyword* that sounds similar to the to-be-remembered word. For example, to remember the word “carlin”, a user of the keyword method might imagine an old woman entering a car. Carlin means old woman and sounds like car. Therefore, remembering “Car” and “Old woman” produces the correct word, “Carlin” (McDaniel & Pressley, 1984). If mnemonic devices can improve memory for such a diverse range of information (i.e., serially presented lists to single items), then perhaps mnemonic devices can also be applied to name–face memory.

A few name–face memory mnemonics have been investigated. The expanding schedule and key-feature mnemonics are just two examples. The expanding schedule mnemonic requires a user to repeatedly recall the to-be-remembered name–face pair over an increasing window of time (Bredart, 2019; Morris & Fritz, 2000). If you were attempting to remember a name of some one you just met, you might try to recall their name immediately after, a few minutes after, and a few hours after first meeting them. An extension of the expanding schedule mnemonic is the name game (Morris & Fritz, 2000). The name game starts by one individual introducing themselves. After the first person says their name, then the second person in line repeats the first person's name and introduces themselves. This pattern of recall and introduction continues until everyone in the group has introduced themselves. The first person then must recall all the names of the group members. The expanding schedule name–face mnemonic (and by extension, the name game) has been shown to be an extremely robust name–face memory aid in controlled settings (Landauer & Bjork, 1978; Morris & Fritz, 2000). Lastly, the key-feature mnemonic requires a user to make a connection between a key facial feature and mental image of a scenario involving the individual's name (McCarty, 1980). For example, if you were to meet a Daniel who had a large forehead, then you might create a mental image of a dam, connecting a dam to the name Daniel. Next time you were to meet this Daniel, you would remember the large forehead, and remember that dam is close to the name Daniel. Both these name–face mnemonics have been extensively investigated in laboratory settings; however, real word application is potentially difficult (McCarty, 1980; Patton, 1994).

Current name–face mnemonics have several downsides. A viable name–face mnemonic device needs to be efficient, easy to teach, and applicable in any social setting with any name or face. The expanding schedule mnemonic is inefficient, requiring the rehearsal of the name–face pair over a long time period. In addition, the key-feature mnemonic is difficult to teach, and is limited to names and faces that are easily associated by a key-feature. Not all faces will have a distinct facial feature that easily connects to the name. Remembering “Dan with a forehead like a dam” is much easier than “Keith with no distinct features”. Meeting a new person can be a mentally taxing event; having to make a mental connection during a conversation is additional stress. Any additional name–face mnemonic devices need to be easy to apply in a wide range of social interactions and require very little mental effort. One such device was proposed by the popular media.

The 2008 comedy/romance, “The House Bunny”, introduced an interesting name–face mnemonic device. Shelly (played by Anna Farris) claims that repeating the to-be-remembered name out loud in a bizarre voice (e.g., imagine a batman voice) improves her memory of that name and face. This proposed mnemonic device is both efficient—requiring only a single repetition at the time of encoding—and simple to teach. Given these qualities, Patel (2020) investigated the validity of the so-called House Bunny Effect (HBE). The HBE is supported by the scientific literature. The bizarreness and the production effects support the claim that bizarre name production at the time of name–face encoding enhances name–face recall. The bizarreness effect states that bizarre and/or strange items are more easily remembered as opposed to common items (Cornaldi, De Beni, & Pra Baldi, 1988). This memory benefit is presumed to be due to the relative distinctiveness of bizarre memories compared to normal memories (Cox & Wollen, 1981). Distinct memories are much more salient than common memories, thus, easier to recall. The production effect states that items that are produced or spoken are more easily remembered (Quinlan & Taylor, 2013). The production effect is again presumed to rely on the distinctiveness of the items. A sentence that is sung will be remembered better than a sentence that is read quietly because of the distinct memory cues available when singing (Quinlan & Taylor, 2013). The combination of these effects supports the idea that producing (speaking) a name in a bizarre voice will enhance memory for name–face memory. However, Patel and colleagues found a numerical trend in the opposite direction: bizarre name production at the time of encoding hindered cued recall performance compared to normal voice production (Figure 1). This finding is in direct contraction to the predictions of the bizarreness and production effects. The HBE research group hypothesized that this memory hinderance was due to an increased cognitive load that produced a dual task cost. Producing the bizarre voice was both physically and mentally taxing. The original HBE study required participants to complete two tasks: first they had to create the bizarre voice and second, they had to form a name–face memory association. The dual-task cost literature states that performance will be worse on a task when it is completed at the same time as a second task (Logie et al., 2007). This cognitive demand to produce the voice may have taken attentional memory resources away from the to-be-remembered name–face associations. It is likely that the cognitive load was less for normal voice trials producing the observed numerical difference in cued memory recall scores. The original HBE study discovered a surprising memory trend, however, several experimental design choices may have contributed to these results.

Design choices of the original HBE study require correcting in a conceptual replication. The original HBE study was designed by the Multisensory Perception (MSP) Lab at McMaster University. The MSP lab was very interested in where participants were focusing during their memory task, so they required all participants to use an eye-tracking rig. No a-priori predictions about eye positions were made, and no statistically significant trend of eye movements was observed. The eye tracking rig was cumbersome and required a few minutes of calibration. It is possible the addition of the eye-tracking rig contributed to the overall

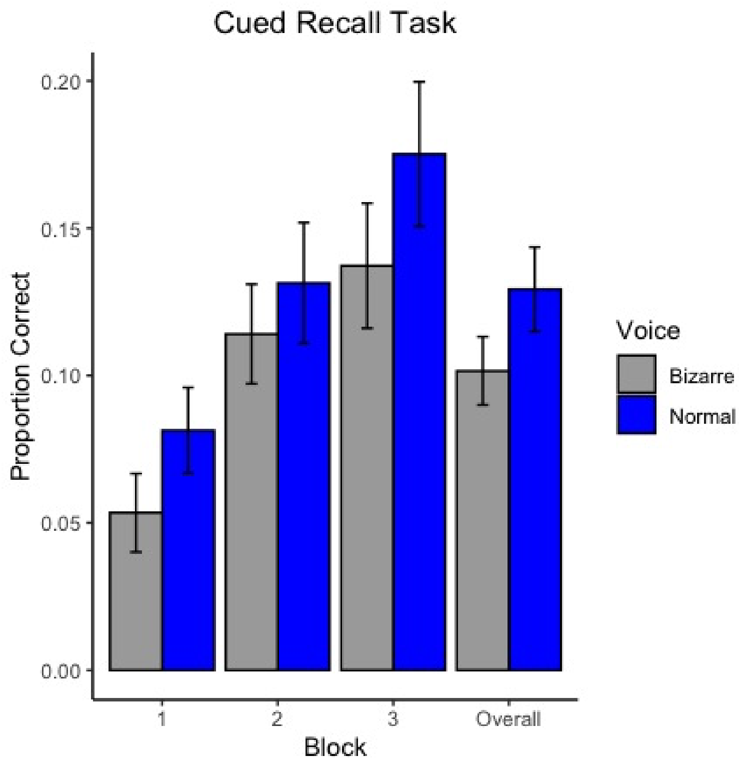


Figure 1: Cued recall task results (Patel (2020), pp. 11). Bizarre voice production at the time of encoding hindered recall of name–face pairs compared to a normal voice production control. This negative trend was not statically significant.

cognitive demands of the task, leading to the statistically insignificant effect. Our conceptual replication omitted the eye-tracking task to cut down on any confounding factors. In addition, participant cued recall performance was low in the original study (.10-.15 proportion correct recalled; Figure 2). This poor average performance reflects the difficulty of the task. Each learning block contained 15 trials: 15 names paired with 15 faces. Remembering 15 name–face pairs is an extremely difficult task. In an attempt to raise cued recall performance, we decreased the number of name–face pairs presented in each block. Through a series of pilot studies, we settled on learning block sizes of 10 trials with 8 tested name–face pairs in each cued recall block. Ten name–face pairs in each of four learning blocks improved cued recall performance while still producing enough data for statistical analysis. Next, the original HBE study only included female Caucasian faces, female names, and female participants. This design choice was based on unpublished research from the MSP lab that suggested there was a sex difference in name–face memory performance. However, a mnemonic device must be applicable to a wide range of social situations. In our conceptual replication, we included a wider ranger of ethnicities, male names, male faces, and male participants. Lastly, Patel (2020) only included a cued recall task and an old/new recognition task for faces. In addition to these two tasks, we included a free recall task for names. It is possible that bizarre name production influences memory for names alone. Only using a cued recall task and recognition task would fail to target this possibility. By using all three memory tasks, we hope to investigate the effects of bizarre name production on name–face memory (cued recall task), name memory (free recall task), and face memory (recognition task). Patel (2020) produced an extremely interesting finding with bizarre name production producing a trend toward a memory cost rather than a memory benefit, however, we hope to further their findings through a more ecologically valid online conceptual replication.

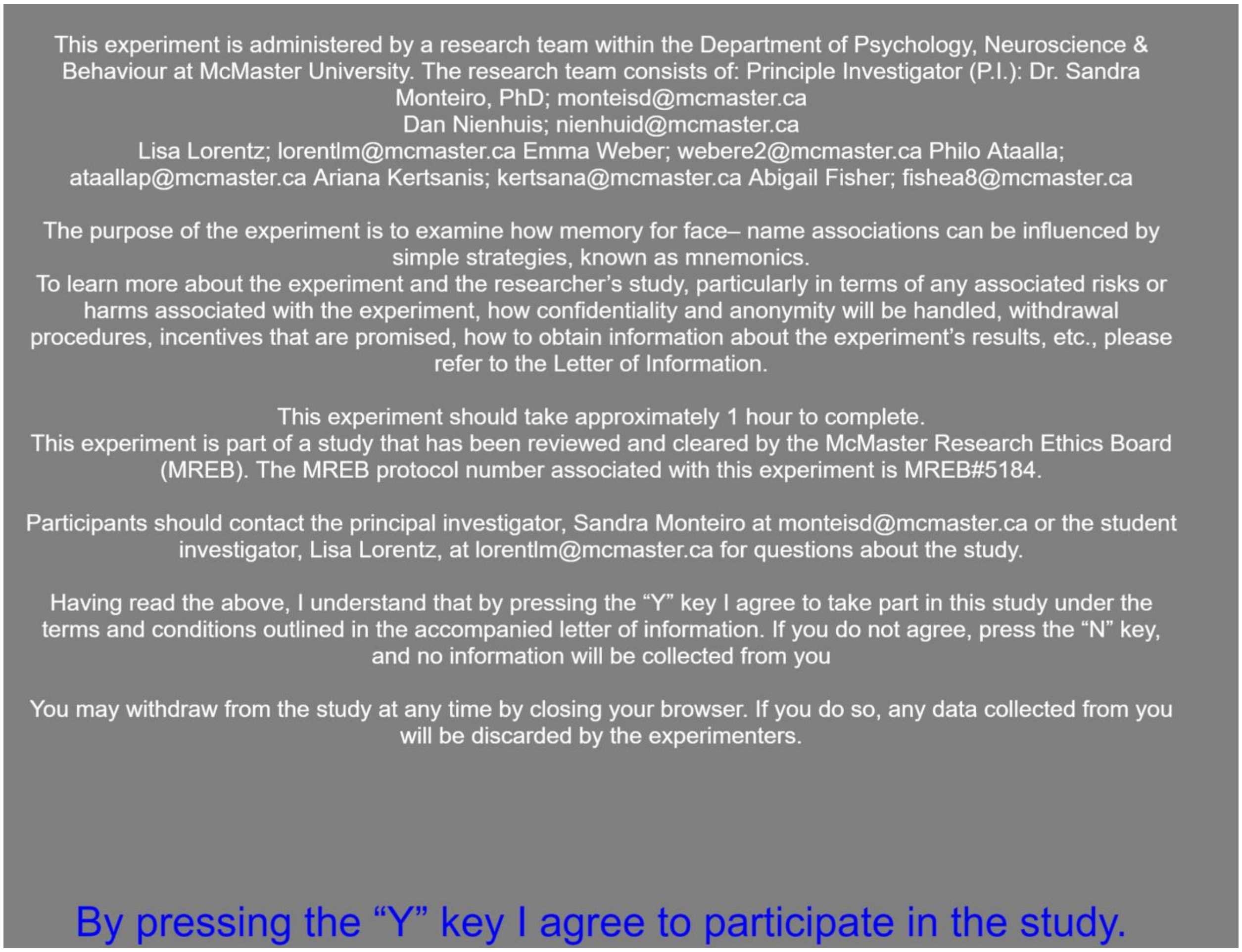
Here we produce a series of studies to further elaborate on the findings of Patel (2020). Study one was designed as a conceptual replication of the original HBE memory study. Participants were presented with a series of name–face pairs and instructed to repeat the names back in either a bizarre or normal voice. Their memory for the name–face pairs was tested in subsequent memory tasks. Study two more directly evaluated the influence of cognitive load using a passive task manipulation of study one. Participants were no longer required to produce the words in a bizarre or normal voice, rather, they simply heard computer generated bizarre and normal voices. Their name–face memory was again tested in subsequent memory tasks. Patel (2020) suggested that bizarre voice production hindered name–face memory because of the cognitive effort required to produce the bizarre voice. If this theory is true, then the omission of the voice production requirement would improve memory for bizarre name–face trials. The results of study one and two will provide insight into how bizarre name production influences name–face associations and name–face memory recall.

Methods

Study 1: Online Conceptual Replication of the HBE

Participants

Thirty-two (n = 32) McMaster University undergraduate psychology students participated in this online memory study (3 males, 29 females; mean age = 18.5 years, SD = 1.1 years). Participation was voluntary but restricted to individuals with corrected or corrected-to-normal vision and audition by self-report. Participation was compensated with one course credit.



This experiment is administered by a research team within the Department of Psychology, Neuroscience & Behaviour at McMaster University. The research team consists of: Principle Investigator (P.I.): Dr. Sandra Monteiro, PhD; monteisd@mcmaster.ca
Dan Nienhuis; nienhuid@mcmaster.ca
Lisa Lorentz; lorentlm@mcmaster.ca Emma Weber; webere2@mcmaster.ca Philo Ataalla; ataallap@mcmaster.ca Ariana Kertsanis; kertsana@mcmaster.ca Abigail Fisher; fishea8@mcmaster.ca

The purpose of the experiment is to examine how memory for face– name associations can be influenced by simple strategies, known as mnemonics.
To learn more about the experiment and the researcher's study, particularly in terms of any associated risks or harms associated with the experiment, how confidentiality and anonymity will be handled, withdrawal procedures, incentives that are promised, how to obtain information about the experiment's results, etc., please refer to the Letter of Information.

This experiment should take approximately 1 hour to complete.
This experiment is part of a study that has been reviewed and cleared by the McMaster Research Ethics Board (MREB). The MREB protocol number associated with this experiment is MREB#5184.

Participants should contact the principal investigator, Sandra Monteiro at monteisd@mcmaster.ca or the student investigator, Lisa Lorentz, at lorentlm@mcmaster.ca for questions about the study.

Having read the above, I understand that by pressing the "Y" key I agree to take part in this study under the terms and conditions outlined in the accompanied letter of information. If you do not agree, press the "N" key, and no information will be collected from you

You may withdraw from the study at any time by closing your browser. If you do so, any data collected from you will be discarded by the experimenters.

By pressing the "Y" key I agree to participate in the study.

Figure 2: Online consent form. This form was presented to each participant prior to the start of the study. Pressing the "Y" key signified that they had given their consent to take part in the study. A PDF copy of the consent form was emailed to the participant following the completion of the study. If the participant did not consent, the study was terminated with no further data collection. The contact information of the researchers has been censored for their privacy.

Prior to the start of the study, participants gave their online consent to take part in the study (Figure 2). Handedness, age, sex at birth, and the need for corrective lenses (e.g., glasses or contacts) were recorded for each participant. These demographic data were obtained by self report. The current study was approved by the McMaster Research Ethics Board (MREB; #5184) and funded by NSERC Canada. All participant data were stored on secure servers at McMaster University.

Apparatus

The current study was coded in and run on Pavlovia® 2020.2.8 (Pierce et al., 2019). A link to the online platform was sent to each participant prior to their appointment slot. The study was run on each participants' personal electronic device. Permitted devices were limited to computers (e.g., desktop and laptops) and omitted tablets and smartphones. Participants were encouraged to use the speakers built into their device; however, headphones were permitted if required.*

Design & Stimuli

The memory for a series of name–face pairs was tested for each participant. Participants were presented with a series of names and faces and were instructed to repeat each name in either a bizarre or normal voice. Their name–face memory was tested using cued recall, free recall, and recognition tasks.

A 2x2x4 mixed factor design was used. The first factor "voice" contained two levels: bizarre and normal. The factor of voice described how the participants had to pronounce the

*Due to COVID-19 restrictions, it was expected that participants would take part in the study from the comfort of their home. Therefore, given that additional familial noises (e.g., children, dogs, family members) were likely, headphones were permitted to ensure each participant could hear the auditory stimuli.

presented name: either a bizarre or normal pronunciation. Second, the factor of "sex" had two levels: male and female. This factor pertains to the sex of the faces presented to the participants during the learning, cued recall, and recognition task phases. Lastly, the factor of "ethnicity" included four levels: African American, Asian, Latino, and Caucasian. This factor describes the ethnicities of the faces presented to the participants during the learning, cued recall, and recognition task phases.

Faces were obtained from the Chicago Face Database (CFD: Developed by the University of Chicago). The database can be downloaded for free for research purposes (https://chicagofaces.org/default/). Sixty-four female and sixty-four male faces (sixteen faces from each ethnicity) were randomly selected from a pool of 597 faces.

Names were obtained from the top 100 names in the United States in the year 2000. Only easily pronounceable and spellable names were included (e.g., Benjamin and Daniel). Voice recordings were developed by author BS and colleagues at McMaster University using the stock apple voice synthesizer app.

The eight experimental learning trials from each block included a randomized order of four male and four female faces. Of these eight faces, two were randomly selected from each ethnicity. Lastly, four trials were "Bizarre" (i.e., participants repeated the name in a low, gravely voice) and four were "Normal" trials (i.e., participants repeated the name in a normal voice).

The two filler trials (first and last trials) from each learning block included a randomized order of one male and one female face. These faces were randomly selected from the four ethnicities. One "Bizarre" and one "Normal" trial was used.

Procedure

At the time of this study, COVID-19 was a public health risk. In keeping with the social distancing guidelines outlined by the Canadian government and McMaster University, the study was run in an online format; no in person contact was permitted. However, an experimenter trained to run the study monitored each participant via a zoom video conference. The experimenter was present to answer any questions the participants might have had and to ensure the study continued as expected. Due to privacy concerns, video cameras were turned off and no auditory or visual recordings were made.

An experimenter began a zoom call with each participant at the time of their scheduled appointment. Participants were asked to sit in a quiet and comfortable space and to position their electronic device at a comfortable distance from their face. To ensure that the speakers from their device were working, a test monotone was played over their speakers. Participants were told to adjust their audio settings to a comfortable level. Once the participant was comfortable with the volume level, they were free to continue with the study.

The study involved five phases: learning blocks, cued recall task, distractor task, free recall task, and a recognition task (Figure 3).

Learning & Cued Recall Task Phases

Participants were instructed to attend to a white central fixation cross. The cross was presented for a duration of 0.5 s and signalled the start of each trial. Following the removal of the fixation cross, a face was presented on the screen. Simultaneously an audio recording stating “Hello, my name is _____” was played. After the completion of the audio recording, the word “Bizarre” or “Normal” was presented below the face. The presented word cued the participant to

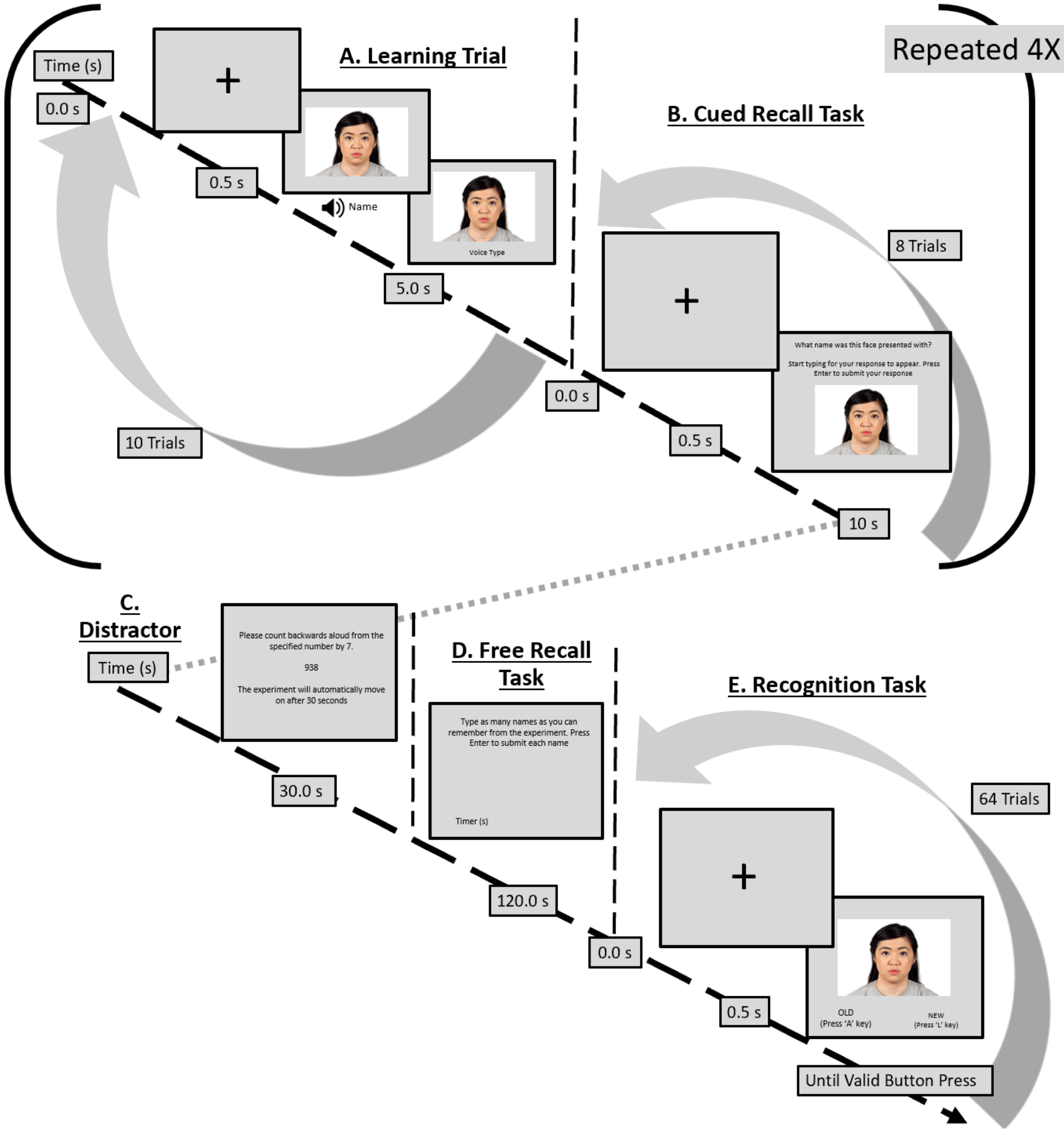


Figure 3. Flow diagram: study one. The study consisted of five phases: learning, cued recall task, distractor task, free recall task, and the recognition task. **A.** Example of a learning trial. Ten trials were presented in each of four blocks. After each block, a cued recall task was presented. The first and last trials of each learning block were filler trials and omitted from the cued recall and recognition tasks due to primacy and recency effects.

B. Example of a cued recall task trial. Eight trials were presented in each of four blocks. These eight trials were the faces presented in the previous learning block (minus the filler trials). Participants were instructed to type the name that was presented with the face from the previous learning block. **C.** Example of the distractor task. Participants counted out loud from 938 by intervals of 7. **D.** Example of the free recall task. Participants were instructed to recall any of the names presented in any of the previous four learning blocks. Order of names recalled did not matter. **E.** Example of the recognition task. Participants identified each face as either “old” (presented in one of the learning blocks) or “new” (never seen during the experiment). Sixty-four trials were presented: thirty-old and thirty-four new. No filler trial faces were used during this task.

use either a bizarre or a normal voice to repeat the presented name. If a “Bizarre” cue was given, participants had to verbally repeat the presented name using a low, gravely voice. Participants were required to try and use the same bizarre voice throughout the experiment. If a “Normal” cue was given, participants were asked to use their normal speaking voice to repeat the name. Each face was presented for a duration of five seconds.

Ten faces were presented in each of four learning blocks. Of the ten trials in each block, eight were considered experimental trials while the first and last trials were considered filler trials. Experimental trial faces were presented again during the cued recall task and recognition task phases. In contrast, filler trial faces were not presented again in any following phases. Filler trials were included to control for primacy and recency effects (Laming, 2010).

Immediately following each block of learning, a cued recall task was presented. Eight trials were presented in each of the four recall blocks. These eight trials were the eight experimental trial faces from the previous learning phase presented in a randomized order.

Each trial began with a white fixation cross which was presented for 0.5 seconds. Following the fixation cross, a face was presented for ten seconds. Participants were asked to type the name that was presented with the face from the previous learning phase. If participants did not remember the name, they were instructed to provide a best guess. The trial timed out after 10 seconds. In total, thirty-two cued recall trials were presented to each participant.

After the completion of four cycles of learning and cued recall task phases, participants started the distractor task phase.

Distractor Phase

Participants were instructed to count backwards from 938 by a seven-digit interval for twenty seconds. A distractor task was required to minimize any mental rehearsal of the names and faces presented during the most recent learning and cued recall task phases. After twenty seconds, the experiment automatically moved onto the free recall phase.

Free Recall Phase

Participants were instructed to type all the names they could recall from all the previous learning phases. Order of the names did not matter. The experiment automatically moved onto the recognition task phase after 120 seconds.

Recognition Task Phase

The recognition task phase required participants to identify faces as either “old” or “new”. If they thought the face was presented in a previous learning phase, participants pressed “A”, identifying the face as old. Similarly, if they thought the face had never been presented before, participants pressed “L”, identifying it as new. Sixty-four faces were presented: thirty-two faces were old while thirty-two faces were new. The thirty-two “old” faces consisted of all the experimental trial faces from each of the four learning blocks presented in a randomized order.

The start of each trial was signified by a white fixation cross that was presented for 0.5 seconds. Following the removal of the fixation cross, a face was presented. Participants pressed “A” to signify an old face or pressed “L” to signify a new face. The next trial began immediately following a response.

After the completion of all sixty-four recognition task trials, participants were verbally debriefed and thanked for their time. A copy of their consent form and a short debrief form were emailed to each participant immediately following the meeting. The compensation in the form of course credit was transferred to them shortly after the completion of the study.

Results

Voice Accuracy Coding

Undergraduate researchers PA, EW, AK, and AF manually recorded the voice accuracy (Bizarre voice used on a bizarre trial and vice versa) of each participant during the study. Only trials in which the participants used the correct voice were included for statistical analysis. Across all thirty-two participants, only 85 trials (8.3%) were omitted from analysis due to the usage of an incorrect voice.

Response Coding

Typed cued and free recall name responses were manually encoded by the researchers. Names that were spelled incorrectly but sounded like the correct name (e.g., Connor and Connar) were marked as correct. Names that were only partially complete (e.g, Ben for Benjamin or Arian for Arianna) were flagged for further evaluation by the research group. Wrong names and repeated names were labelled as incorrect. All statistical analyses were completed using RStudio© (version 1.3.959). To maintain inter-rater reliability, prior to coding all the data files, the research group met to discuss what an incorrect, correct, and “flagged” response looked like. Undergraduate researchers PA, EW, AK, and AF coded the data files for the participants they personally supervised. All flagged responses were run past graduate student supervisor LL.

Cued Recall Performance

A paired-sample t-test was run between the bizarre and normal voice production cued recall responses. A significant effect of voice type (i.e., bizarre and normal) at the time of encoding on cued recall performance was observed ($t(17) = -3.63$, $p = .002$). Memory performance on the cued recall task was significantly worse for bizarre production trials than normal production trials (Figure 4). Bizarre production did not produce a memory benefit, rather, bizarre name production at the time of name–face memory encoding hindered name–face cued recall performance. We successfully replicated the findings of the original HBE study through a conceptually identical online study. Producing a name in a bizarre voice at the time of encoding produced a name–face memory cost compared to a normal voice production control.

Free Recall Performance

A paired-sample t-test was applied to the bizarre and normal name production trials for the free recall name responses. Names were only included in the analysis if the correct voice was used at encoding. No statistically significant effect of voice type production on name free recall was observed ($t(31) = -.75$, $p = .46$). Bizarre name production at the time of encoding did not improve or hinder name free recall as compared to a normal name production condition (Figure 5).

Recognition Task Performance

A paired-sample t-test was used to analyse the difference between the bizarre and normal name production at the time of encoding for old/new recognition judgments. No statistically significant difference in corrected hits (Hits: correctly identifying an old face – False Alarms: incorrectly identifying a new face as old) was observed between bizarre and normal name production trials ($t(63) =.61$, $p = .55$). Bizarre name production at the time of encoding did not improve or hinder old/new face recognition judgments as compared to a normal name production condition (Figure 6).

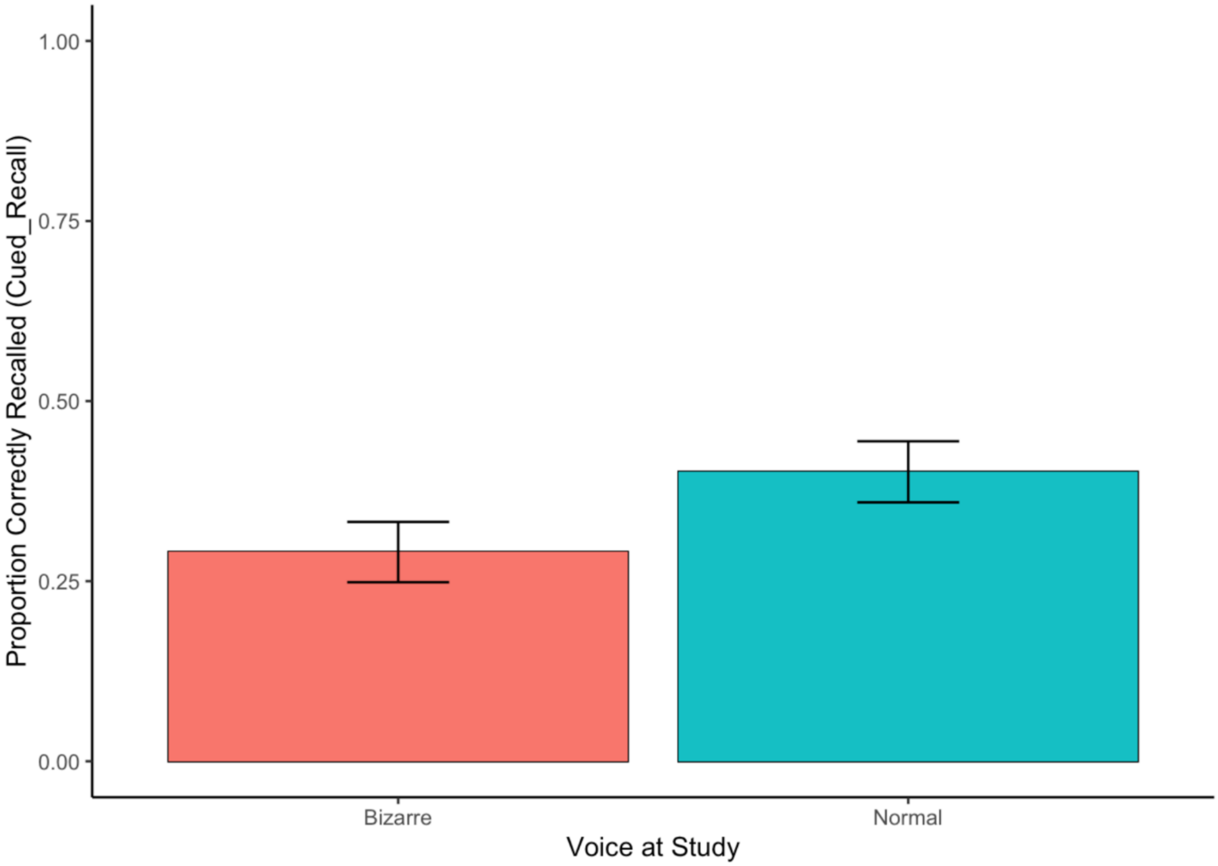


Figure 4: Cued recall task performance: study one. Proportion correctly recalled for bizarre and normal production conditions during the cued recall task. Bizarre voice production at the time of encoding significantly hindered cued recall performance for name–face pairs compared to a normal voice production control.

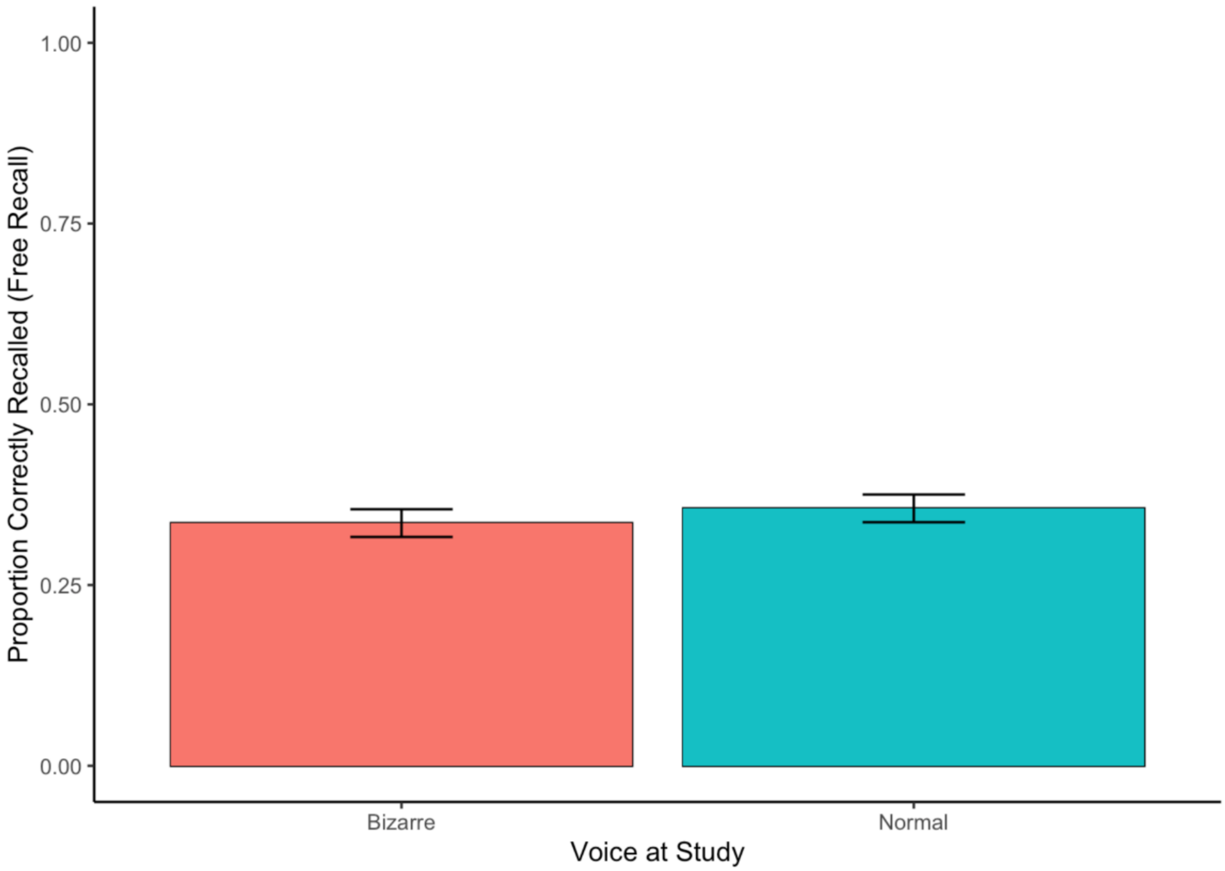


Figure 5: Free recall task performance: study one. Proportion correctly recalled for bizarre and normal production conditions during the free recall task. The number of names recalled from each condition did not differ significantly. Bizarre voice production at the time of encoding did not influence free recall performance compared to a normal voice control condition.

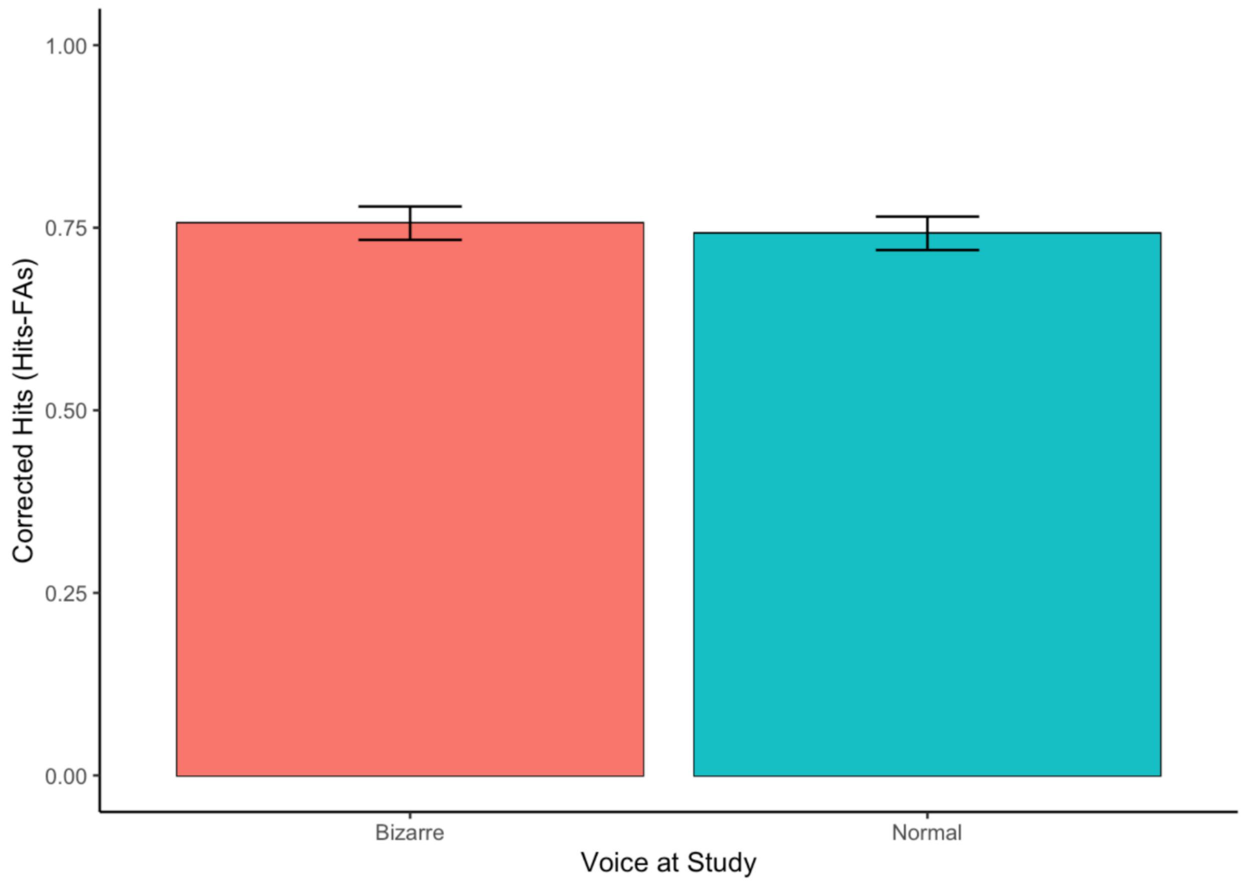


Figure 6: Recognition task performance: study one. Corrected-hits for the bizarre and normal voice production conditions. The measure of corrected-hits represents the proportion of hits (correctly identifying a face as old) minus the false alarms (incorrectly labelling a new face as old). Bizarre voice production at the time of encoding did not significantly influence recognition performance compared toa normal voice production control.

Discussion

Study One: Online Conceptual Replication

We successfully replicated the findings of Patel (2020) while using an online conceptual replication. Producing a name in a bizarre voice leads to a name – face memory cost compared to normal name production. Interestingly, we did not observe any significant difference in free recall or recognition task performance across bizarre or normal voice production trials. Producing a name in a bizarre voice does not influence name or face memory alone compared to normal production. This observation may be partly explained by the cognitive load/dual task cost theory proposed by Patel (2020). Producing a name in a bizarre voice requires additional cognitive resources compared to simply producing a name in a normal voice. This additional cognitive demand for bizarre production may leave fewer cognitive resources to create strong name–face associations. In contrast, the relatively simple task of speaking normally leaves sufficient cognitive resources to create strong name–face memories. This difference in available cognitive resources may explain the observed memory cost for bizarre production trials. Study two investigates the validity of the cognitive load/dual task theory.

Study 2: Passive Task Manipulation

In study one, we successfully replicated the results of Patel (2020): bizarre name production at the time of name–face encoding hinders memory recall in a subsequent cued recall task. However, study one failed to address the mechanisms behind the negative impacts of bizarre name production. As previously stated, a dual task cost driven by a large cognitive load was theorized as the explanation behind these results. Producing the bizarre voice required significant mental resources, leaving few cognitive resources to create strong name–face memory associations. To further investigate the validity of

the cognitive load-dual task cost theory, study two was designed as a passive task. Participants no longer had to produce the bizarre voice, rather, they simply heard the names in a bizarre or normal voice. If the cognitive load theory is true, then removing the demanding task of producing the bizarre voice will improve cued recall performance for bizarre names compared to study one.

Methods

Study two was identical to study one except for one change. Participants were no longer required to repeat the names back in either a bizarre or normal voice. Rather, participants simply heard names presented in either a bizarre or normal voice (Figure 7).

Participants

Thirty-two (n = 32) McMaster University undergraduate psychology students participated in study two (8 males, 24 females; mean age = 19.1 years, SD = 1.8 years). Participation was voluntary but restricted to individuals with corrected or corrected-to-normal vision and audition by self-report. Participation was compensated with one course credit or $*15/hr*.

Prior to the start of the study, participants gave their online consent to take part in the study (Figure 2). Handedness, age, sex at birth, and the need for corrective lenses (e.g., glasses or contacts) were recorded for each participant. These demographic data were obtained by self report. The current study was approved by the McMaster Research Ethics Board (MREB) and funded by NSERC Canada. All participant data were stored on secure servers at McMaster University.

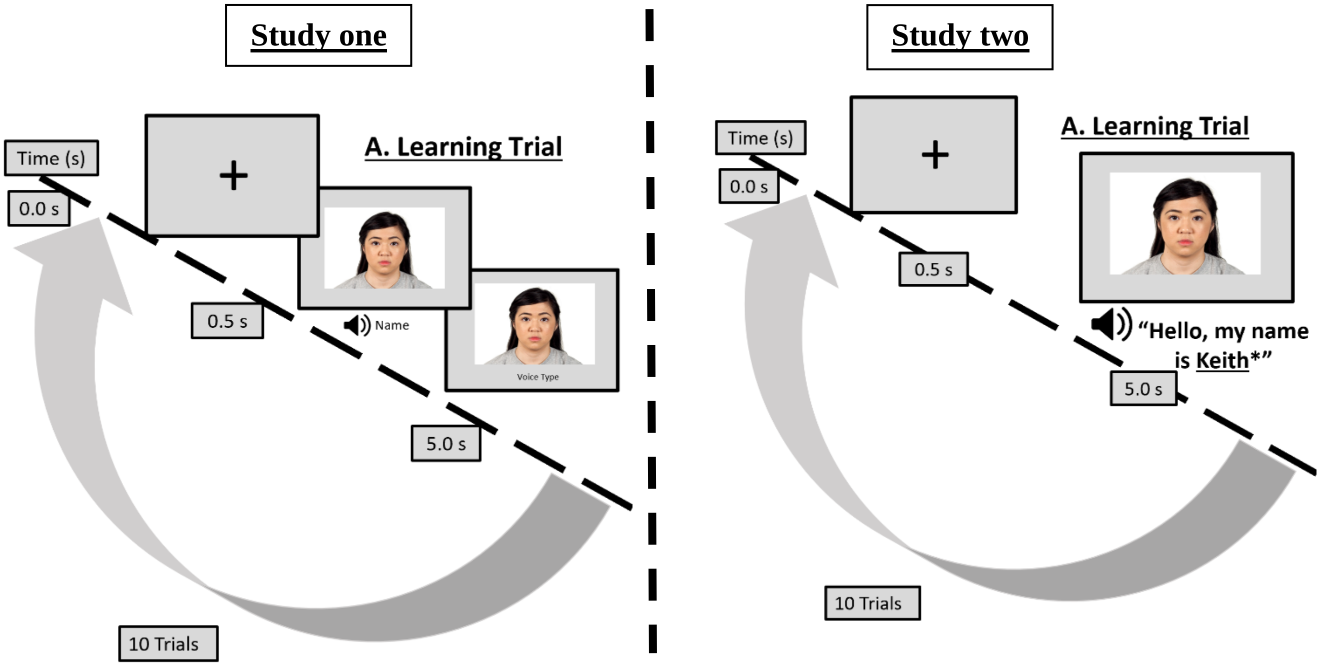


Figure 7: Flow diagram: study two methods change. Visual representation of the single change in study two compared to study one. Participants were no longer instructed to repeat the presented names in the learning blocks. Instead, participants passively heard computer-generated names that stated "Hello, my name is _____". *The name was presented in either a bizarre or normal voice. The rest of study two was structured exactly like study one.

Stimuli & Apparatus

Bizarre audio files were created by using the built in “Old Male” and “Old Woman” voice filters in *Voxal*. The Old Male voice filter added a low pass filter set at 800 Hz, a voice pitch shifter set at 50% (while preserving formants), vibrato at 600 Hz with a 0.50 depth (semitone), and amplification with a gain percentage of 180%. The Old Woman filter added a high pass filter set at 1250 Hz, a voice pitch shifter set at 130.5% (while preserving formants), and vibrato at 10 Hz with a 0.90 depth (semitone). Normal voice audio files were not altered.

Results

Response Coding

Typed responses were coded exactly like study one. No trials were omitted due to voice accuracy as the participants were not required to produce the names.

Cued Recall Performance

A paired-sample t-test was run between the bizarre and normal names that were correctly recalled during the cued recall task. A significant memory effect of names heard in a bizarre voice and normal voices during encoding was observed ($t(31) = -2.55$, $p = .016$). Memory performance on the cued recall task was significantly worse for bizarre voice trials than normal voice trials (Figure 8). Our passive task manipulation—removing the requirement to produce the bizarre or normal voice at the time of encoding—did not improve memory for bizarre voice conditions as predicted by the cognitive load-dual task cost theory proposed by the original HBE study.

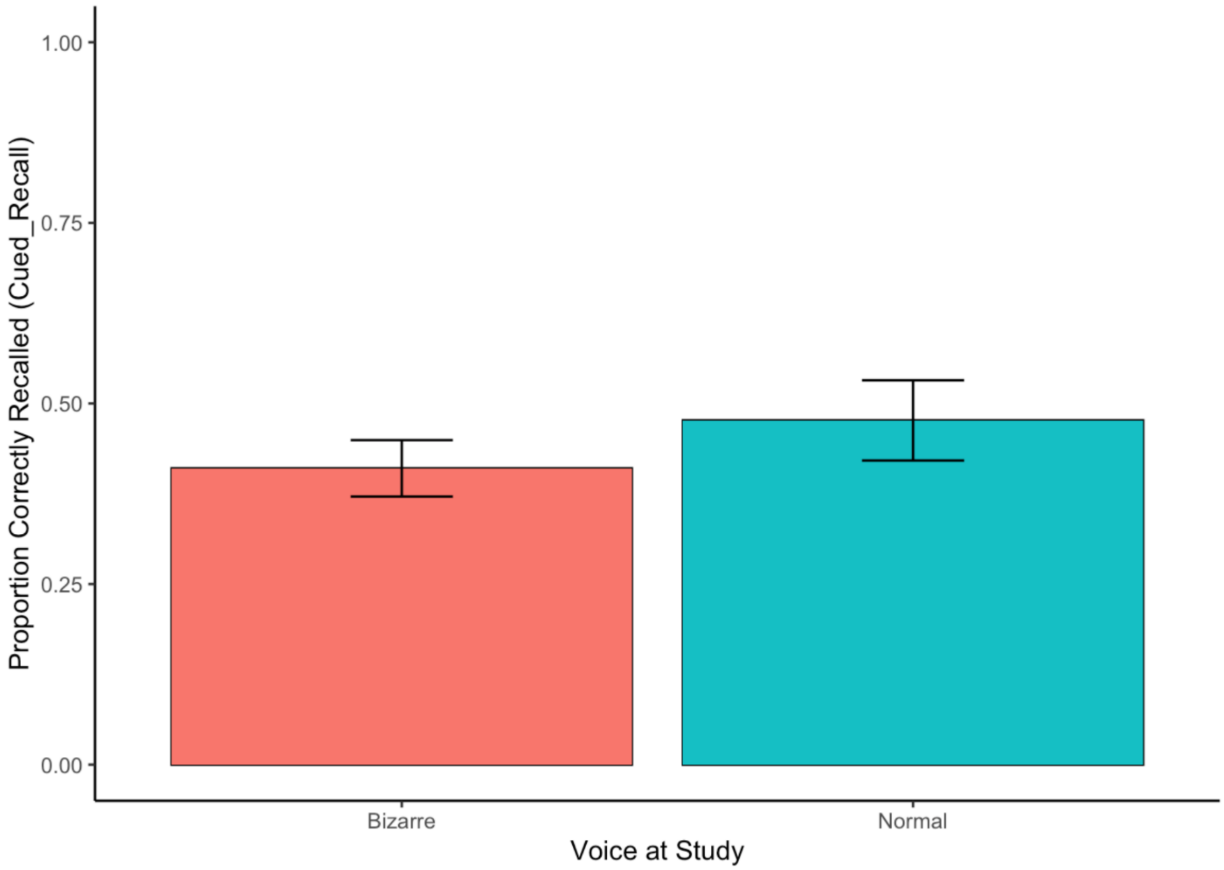


Figure 8: Cued recall task performance: study two. Proportion correctly recalled for bizarre and normal passive conditions during the cued recall task. Hearing a name presented in a bizarre voice at the time of encoding significantly hindered cued recall performance for name–face pairs compared to hearing a name pronounced in a normal voice.

Free Recall Performance

A paired-sample t-test was applied to the bizarre and normal voice trials for the free recall name responses. A statistically significant memory effect was observed ($t(31) = -2.98$, $p = .006$). Hearing a name in a bizarre voice at the time of encoding significantly decreased the number of names recalled compared to a normal voice control (Figure 9). Study one did not produce any statistically significant differences in free recall scores across bizarre and normal voice production trials. Hearing names in a bizarre voice appears to hinder free recall performance more than producing names in a bizarre voice.

Recognition Performance

A paired-sample t-test was used to analyse the difference between the bizarre and normal voice trials at the time of encoding for old/new recognition judgments. A statistically significant difference in corrected hits (Hits: correctly identifying an old face –false alarms: incorrectly identifying a new face as old) was observed between bizarre and normal name production trials ($t(31) = -2.06$, $p = .048$). Hearing a name in a bizarre voice at the time of encoding significantly decreased recognition performance compared to a normal voice control (Figure 10). Again, study one did not produce any statistically significant differences in recognition scores across bizarre and normal voice production trials. Hearing names in a bizarre voice appears to hinder recognition performance more than producing names in a bizarre voice.

Discussion

Study Two: Passive Task Manipulation

The memory cost observed in Patel (2020) and replicated in study one was theorized as a dual task cost produced by an increased cognitive load. Under the dual task cost theory

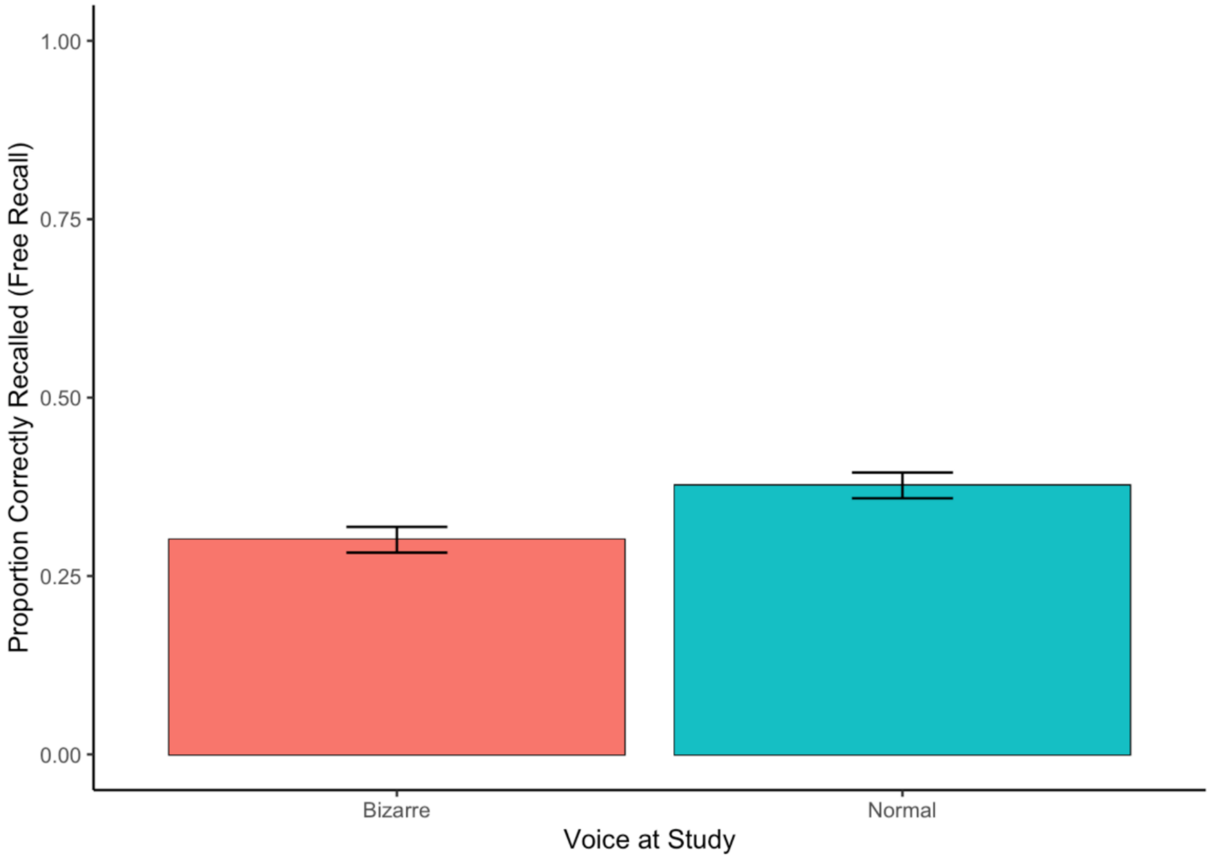


Figure 9: Free recall task performance: study two. Proportion correctly recalled for bizarre and normal passive conditions during the free recall task. The number of names recalled from each condition was statistically different. Hearing a name produced in a bizarre voice significantly hindered name–face memory recall compared to hearing a name produced in a normal voice.

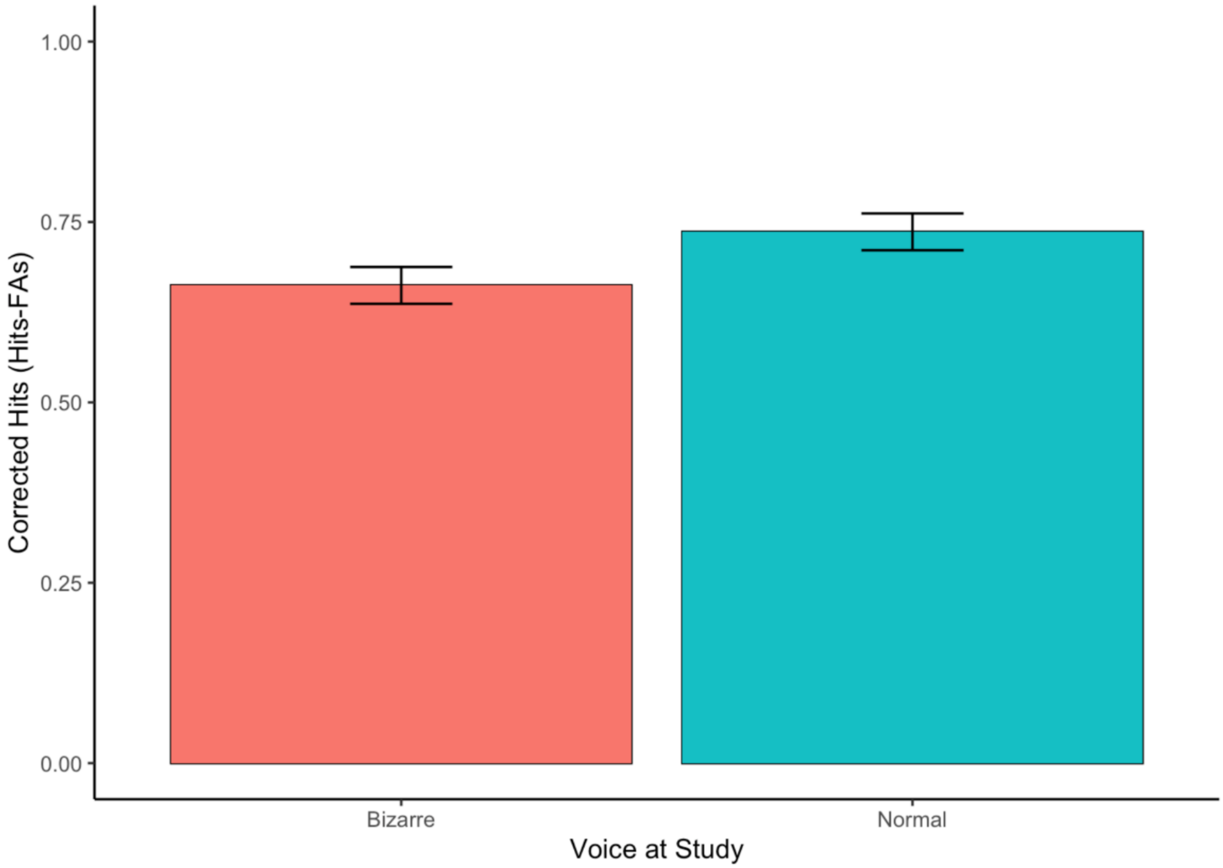


Figure 10: Recognition task performance: study two. Corrected hits for the bizarre and normal passive conditions. The measure of corrected hits represents the proportion of hits (correctly identifying a face as old) minus the false alarms (incorrectly labelling a new face as old). Hearing a name produced in a bizarre voice at the time of encoding significantly decreased recognition performance compared to hearing a name in a normal voice.

proposed by Patel (2020), producing a name in a bizarre voice required more cognitive resources compared to producing a name in a normal voice. This dedication of cognitive resources to bizarre production and not name–face memory associations led to the observed memory deficit for bizarre production trials. If the observed memory cost was due to a dual task cost, then diminishing the cognitive demands of the bizarre voice trials should rescue the name–face memory performance for bizarre trials compared to normal trials. However, even as a passive task—participants simply heard names produced in either a bizarre or normal voice—a significant memory deficit was observed for bizarre trials compared to normal trials in all three memory tasks (Figure 8-10). Hearing a name produced in a bizarre voice is sufficient to produce a memory deficit.

Given our inability to produce evidence in support of the cognitive load/dual task cost theory, we completed a series of exploratory experiment wide analyses to better understand our results. Three omnibus two-way mixed factor ANOVAs (one ANOVA for each task) were completed with the factors of experiment (Expt) and voice (Figures 11-13; Tables 1-3).

No significant differences were observed between study one and two recognition task performances when controlling for the main effect of voice (e.g., bizarre or normal) ($F(1,62) = .88$, $p = .35$; Figure 11; Table 1). Participants correctly identified old and new faces equally well across study one and two independent of the voice used. In addition, no effect of voice type on recognition task performance was observed when controlling for the main effect of experiments ($F(1,62) = .16$, $p = .69$; Figure 11; Table 1). Participant's recognition task performance did not differ with the voice type at encoding when omitting the effects of study one and two. The interaction effect trended towards significance ($F(1,62) = 3.3$, $p=.08$), however, this trend

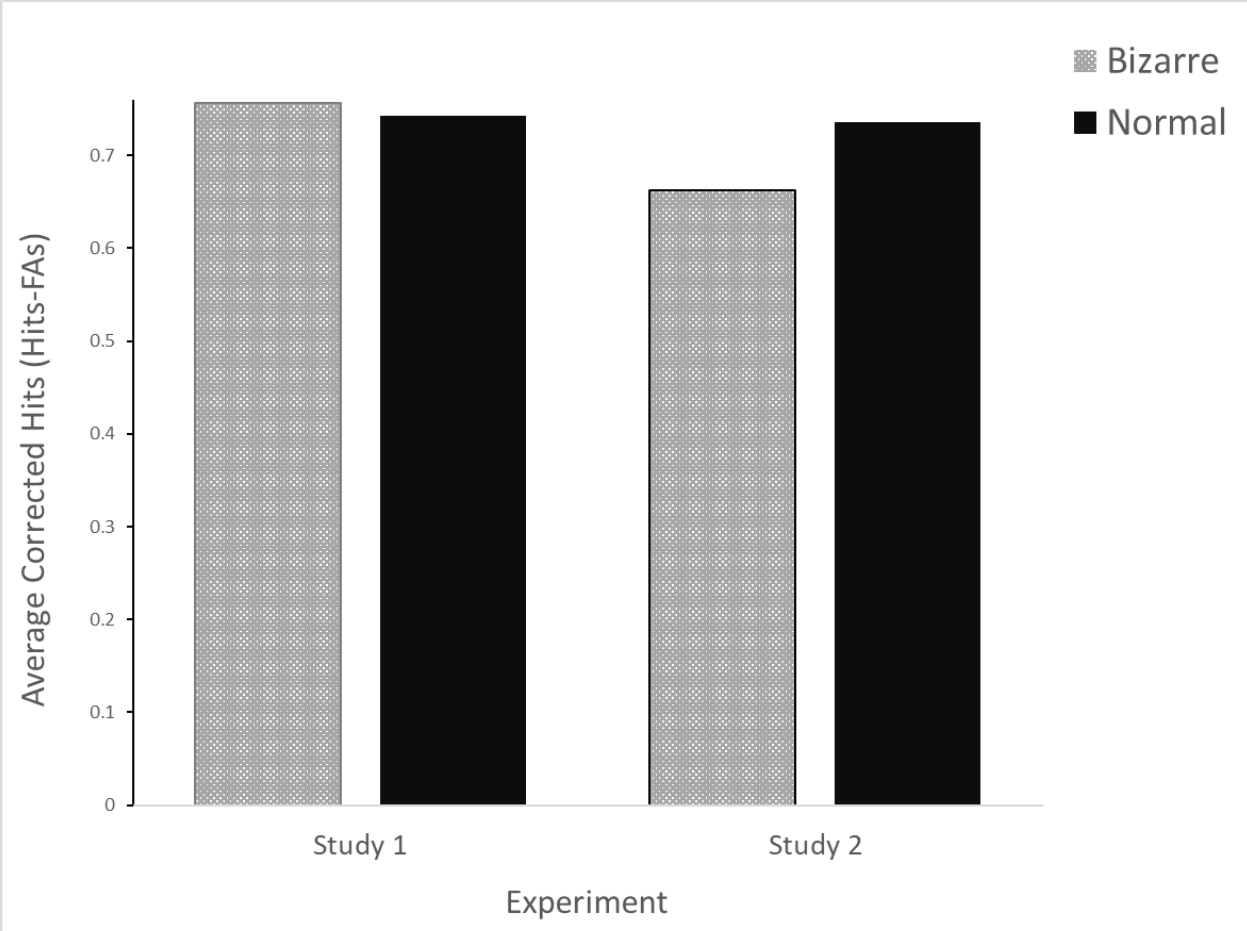


Figure 11: Experiment Wide Analysis: Recognition Task Bar Graph. Average corrected hits for study one and study two recognition task performances.

Table 1.

Omnibus F mixed factor ANOVA for recognition task performance

	Sum Sq	num Df	Error SS	den Df	F Value	Pr (>F)
Intercept	36	1	5.6	62	394	<2e–16**
Expt	.08	1	5.6	62	.88	.35
Voice	.003	1	1.2	62	.16	.69
Expt:voice	.06	1	1.2	62	3.3	.075

**Statistically significant at $p < .001$

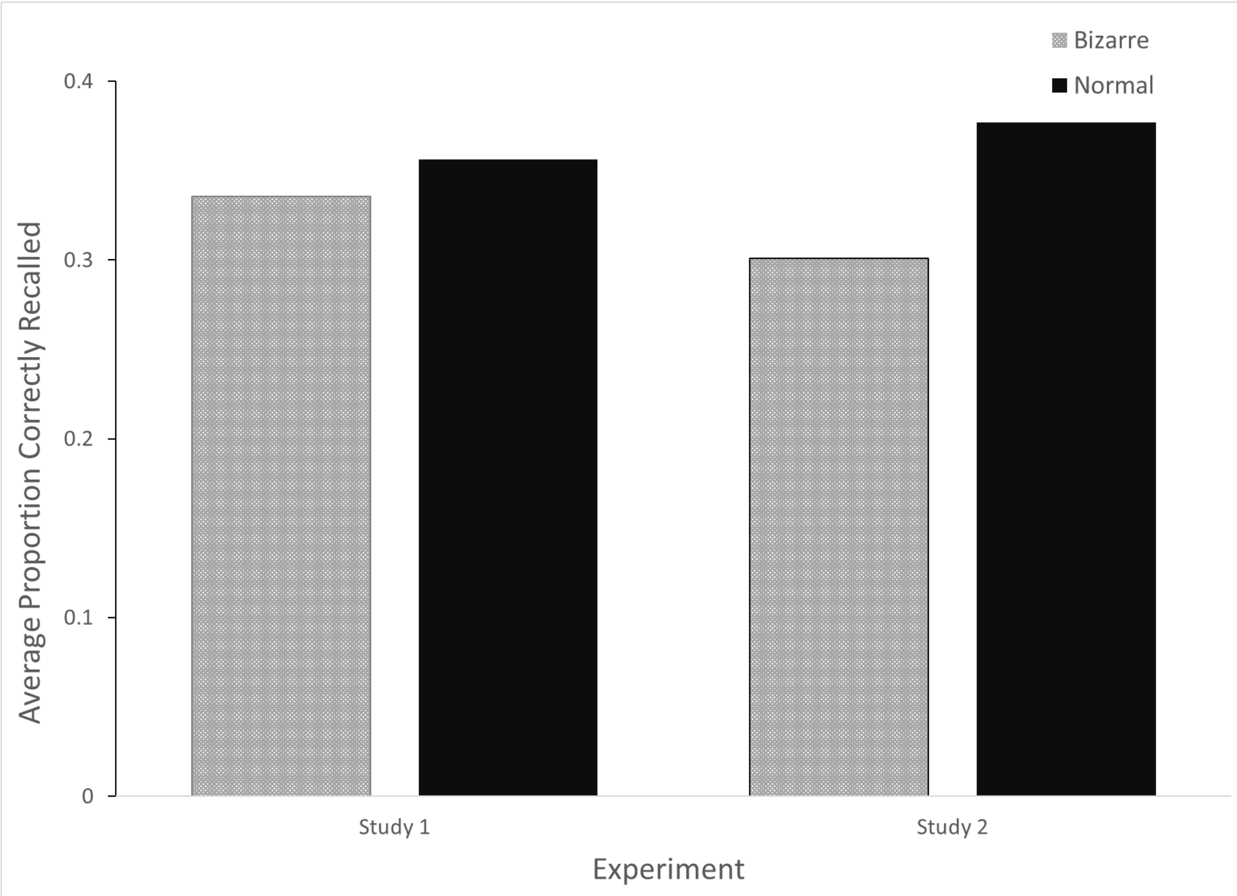


Figure 12: Experiment Wide Analysis: Free Recall Task Bar Graph. Average proportion correct recall for study one and study two free recall task performances.

Table 2.

Omnibus F mixed factor ANOVA for free recall task performance

	Sum Sq	num Df	Error SS	den Df	F Value	Pr (>F)
Intercept	7.7	1	2	62	235	<2e–16**
Expt	.002	1	2	62	.05	.83
Voice	.007	1	.69	62	.6	.44
Expt:voice	.02	1	.69	62	2.2	.14

**Statistically significant at $p < .001$

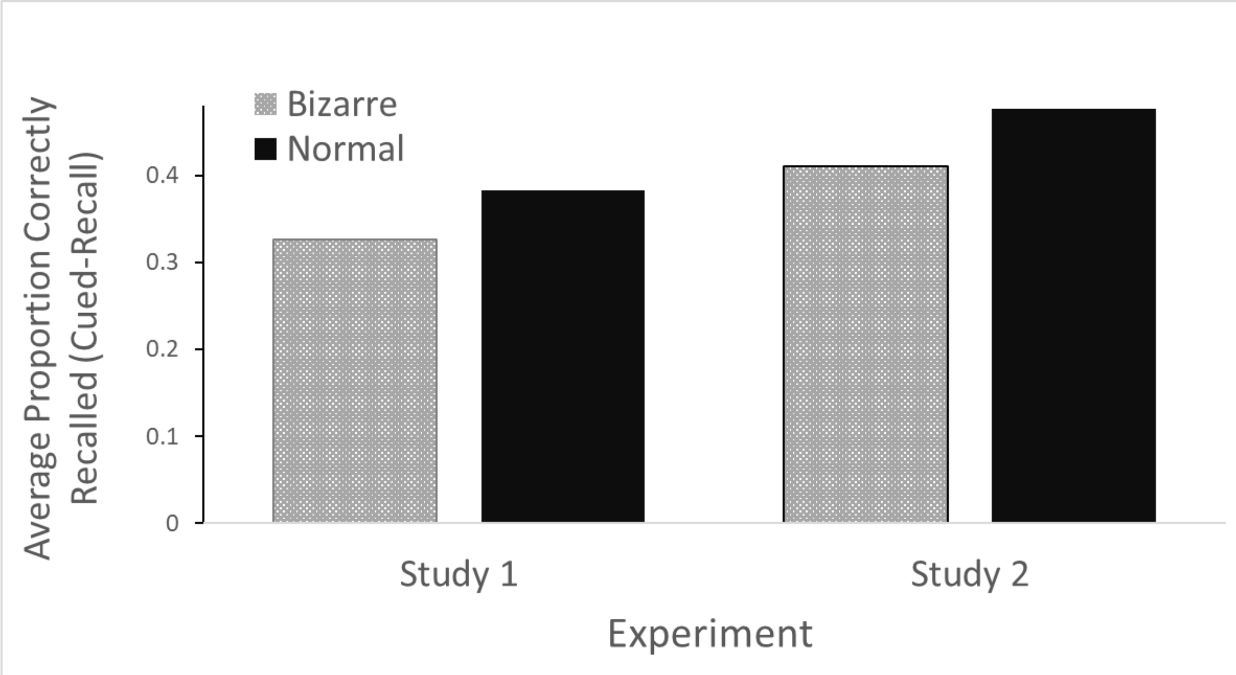


Figure 13: Experiment Wide Analysis: Cued Recall Task Bar Graph. Average proportion correctly recalled for study one and study two cued recall task performance.

Table 3.

Omnibus F mixed factor ANOVA for cued recall performance

	Sum Sq	num Df	Error SS	den Df	F Value	Pr (>F)
Intercept	8	1	3.1	62	159	<2e–16**
Expt	.25	1	3.1	62	5	.03*
Voice	.05	1	.6	62	5.1	.03*
Expt:voice	.0009	1	.6	62	.09	.77

**Statistically significant at p < .001
*Statistically significant at p < .05

suggests that bizarre name performance was selectively inhibited in study two, contrary to our study two goal of improving bizarre name performances to normal name levels (Figure 11).

No statistically significant effect of experiment was observed on free recall task performance when controlling for the main effect of voice ($F(1,62) = .05$, $p = .83$; Figure 12; Table 2). Participants recalled a similar number of names for study one compared to study two. Similarly, no significant effect of voice type on free recall task performance was observed ($F(1,62) = .6$, $p = .44$; Figure 12; Table 2). Participants recalled a similar number of bizarre names compared to normal names when collapsing across study one and two.

A statistically significant difference in cued recall task performance was observed between study one and two when controlling for the main effect of voice ($F(1,62) = 5$, $p = .03$; Figure 13; Table 3). Participants recalled more name–face pairs in study two compared to study one. This finding would suggest that our passive task manipulation in study two did make the cued recall task easier. Omitting the production requirement may have successfully diminished the cognitive load, allocating additional resources towards creating name–face memories. However, this memory benefit was observed for both normal and bizarre voice conditions (Figure ###). This finding would suggest that the cognitive load of production is not driving the observed memory deficit. Lastly, a statistically significant effect of voice type on cued recall task performance was observed when controlling for the main effect of experiment ($F(1,62) = 5.1$, $p = .03$; Figure 13; Table 3). Participants correctly recalled more normal name–face pairs than bizarre name–face pairs when omitting the factor of experiment.

The passive task manipulation in study two successfully improved cued recall task performance compared to study one, perhaps through a diminished cognitive load. However, considering we still observed a statistically significant memory deficit for bizarre names in study two, we can conclude that the cognitive load/dual task theory is not driving this effect. A currently unknown aspect of bizarre names would appear to be producing our memory deficit.

General Discussion

Patel (2020) observed a non-significant trend toward a cost to producing a name in a bizarre voice on name–face memory (Figure 1). This memory cost would appear to be in contradiction to the well-known bizarreness (Cox & Wollen, 1981) and production effects (Quinlan & Taylor, 2013) and less importantly, in direct contradiction of the predictions made by the *House Bunny* (2008). Given these contradictions, the findings reported by Patel (2020) required additional study.

Using a modified online conceptual replication, we successfully replicated the findings of Patel (2020). Producing a name in a bizarre voice at the time of encoding significantly hindered name–face memory recall compared to a normal voice control ($t(17) = -3.63$, $p = .002$; Figure 4). No similar cost to name or face memory alone was observed in the free recall or recognition task results (Figure 5-6). Bizarre verbal production hinders name–face memory recall, however, a dual task cost created by an increased cognitive load (as suggested by Patel (2020)) does not appear to explain this memory cost.

We did not find evidence in support of the cognitive load/dual task cost theory proposed by Patel (2020). The memory cost for names produced in a bizarre voice was theorized to be due to a dual task cost (Logie et al., 2007). It was assumed that producing a name in a bizarre voice required more cognitive resources than simply producing a name in a normal voice. This increased cognitive load for the bizarre name condition may have prevented the formation of strong name–face memory associations for bizarre voice trials. Under the cognitive load/dual task cost theory, decreasing the cognitive load of the bizarre condition would diminish the memory cost observed for the bizarre name production trials. However, diminishing the task demands did not rescue name–face memory recall for bizarre trials as shown in study 2.

Study two was designed as a passive version of study one: participants simply heard names produced in either a bizarre or normal voice. It was theorized that removing the production requirement would reduce the cognitive load of the task. We did not observe an increase in name–face memory recall for the bizarre name condition compared to the normal name condition (Figure 8). However, as shown by the experiment wide analyses, cued recall performance was significantly better in study two than study one, independent of voice type (Table 3). We successfully made the cued recall task easier in study two, perhaps by alleviating the cognitive load of the task, however, we failed to see a rescuing of the name–face memory for bizarre trials compared to normal trials as predicted by the dual task cost theory. Since we made the task less cognitively demanding but still observed a bizarre memory cost, we can conclude that a dual task cost is not driving our memory cost. However, it is worth noting that study two tested different participants than study one. It would be naïve to not acknowledge the possibility that any differences in performance across experiments might be due to between subjects differences. Although, it is unlikely that between subjects variation was the sole contributor to our main effects of experiment. Independent of this possibility, additional research should be conducted to determine why bizarre production and even hearing a name in a bizarre voice is producing a name–face memory cost. In addition, further research must be conducted to investigate the impacts of bizarreness on name and face memory alone.

Our experimental design assumed that each memory task (e.g., cued recall, free recall, and recognition tasks) would test independent pieces of the name–face associations. It was assumed that the cued recall task would test for name–face memory associations, the free recall task would test for name memory alone, and the recognition task would test for face memory alone. While it is true that cued recall performance reliably tests associative memory, it is not guaranteed that the free recall and recognition task performances represented name and face memory performance alone. Participants may have used associations between related names in a list, associations to familiar names, or even name–face associations to recall each name. For example, participants my have created a mental list of names, relating each sequential name to the next. In this way, simple position in this mental list would predict the probability of recall, not the variable of voice type. Similarly, participants may have been using associations between known faces or familiar names to determine familiar from unfamiliar faces. For example, it is possible participants correctly judged an old face simply because the face looked like someone they knew. To state that producing a name in a bizarre voice does not influence name or face memory alone would be an over statement. Similarly, to state that hearing a name produced in a bizarre voice significantly hinders name and face memory alone would be an irresponsible conclusion. However, our experimental design was justified given our main applied question: how does bizarreness and production influence name–face memory associations? Further research dedicated to the impacts of bizarreness and production on name and face memory alone must be conducted before drawing any definitive conclusions from our free recall and recognition task results.

Producing a name in a bizarre voice and hearing a name produced in a bizarre voice leads to a memory cost for name–face associations. The House Bunny mnemonic device would therefore appear to be a memory hinderance, rather than a memory aid. Producing a name–face mnemonic device would appear to be extremely beneficial for the majority of the population. However, perhaps our most potent name–face memory aid already exists in the form of social media. It is a relatively simple task to find someone's name on their Facebook account, LinkedIn profile, or conference bibliography. Perhaps further research should be more focused towards improving memory of alternative name and item pairings. For example, medical education includes the memorization of the human anatomy. Often these name and anatomy pairings are not instinctually related. The medical field demands a name–anatomy mnemonic that aids in the training of the future medical personnel. The results of this study may be used to produce such a mnemonic device.

References

Bredart, S. (2019). Strategies to improve name learning. *European Psychologist*, *24*(4), 349–358.

Burton, M., & Bruce, V. (1993). Naming faces and naming names: Exploring an interactive activation model of person recognition. *Memory*, *1*(4), 457–480.

Cornoldi, C., Cavedon, A., De Beni, R., & Pra Baldi, A. (1988). The influence of the nature of material and of mental operations on the occurrence of the bizarreness effect. *The Quarterly Journal of Experimental Psychology*, *40A*(1), 73–85.

Cox, S. D., & Wollen, K. A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society, 18*(5), 244–245.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236–243.

Laming, D. (2010). Serial position curves in free recall. Psychological Review, 117(1), 93–133.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, R. E. Morris, A R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London, UK: Academic Press.

Logie, R. H., Sala, S. D., MacPherson, S. E., Cooper, J. (2007). Dual task demands on encoding and retrieval processes: Evidence from healthy adult ageing. *Cortex, 43*(1), 159–169.

McCabe, J. A. (2015). Location, location, location! Demonstrating the mnemonic benefit of the method of loci. *Teaching of Psychology*, *42*(2), 169–173.

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*, 145–155.

McDaniel, M. A. & Pressley, M. (1984). Putting the keyword method in context. *Journal of Educational Psychology, 76*(4), 598–609.

Morris, E. P., & Fritz, O. C. (2000). The Name Game: Using retrieval practice to improve the learning of names. *Journal of Experimental Psychology: Applied, 6*(2), 124–129.

O'Mahony, C., & Newell, N. F. (2012). Integration of faces and voices, but not faces and names, in person recognition. *British Journal of Psychology, 103*, 73–82.

Patel, M. (2020). Producing names with a bizarre voice does not improve memory for face–name pairs. [Unpublished Master's Thesis]. McMaster University.

Patton, G. W. R. (1994). Testing the efficacy of name mnemonics used during conversation. *Psychological Reports, 75*, 131–142.

Peirce, J. W., Gray, J. R., Simpsons, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behaviour made easy. *Behaviour Research Methods.*