Research Article

The Visual Mandela Effect as Evidence for Shared and Specific False Memories Across People

Deepasri Prasad and Wilma A. Bainbridge
Department of Psychology, The University of Chicago

Abstract

The Mandela effect is an Internet phenomenon describing shared and consistent false memories for specific icons in popular culture. The visual Mandela effect is a Mandela effect specific to visual icons (e.g., the Monopoly Man is falsely remembered as having a monocle) and has not yet been empirically quantified or tested. In Experiment 1 ($N = 100$ adults), we demonstrated that certain images from popular iconography elicit consistent, specific false memories. In Experiment 2 ($N = 60$ adults), using eye-tracking-like methods, we found no attentional or visual differences that drive this phenomenon. There is no clear difference in the natural visual experience of these images (Experiment 3), and these errors also occur spontaneously during recall (Experiment 4; $N = 50$ adults). These results demonstrate that there are certain images for which people consistently make the same false-memory error, despite the majority of visual experience being the canonical image.

Popular icons, such as characters or logos, are intentionally designed to be eye catching and memorable. Unless the design changes, people are repeatedly exposed to the same canonical icon (i.e., the official, in-use design), which creates a strong sense of familiarity. However, there are some icons for which many people report strongly remembering a different version, one that is not the canonical icon. Interestingly, this incorrect version is the same across those individuals with this false memory. This phenomenon of specific and consistent visual false memories for certain images in popular culture is called the visual Mandela effect (VME). If such specific shared false memory exists, it suggests commonalities across our experiences of these images or a role of properties intrinsic to these images on false memory.

The term "Mandela effect" was coined by Fiona Broome, a paranormal researcher, to describe her false memory of Nelson Mandela dying in prison in the 1980s (Broome, 2010). She claimed that other people also had the same false memory. The term has since propagated on the Internet to describe instances in which many

people share highly specific false memories for names, events, or images. For example, people report having a strong recollection that the Monopoly Man, mascot for the board game Monopoly, wears a monocle. To their surprise, he does not nor has he ever (Fig. 1). Broome and other individuals on the Internet have interpreted this shared false experience as evidence of alternate dimensions. Thus, most of the existing literature on the Mandela effect uses it as an example of a conspiracy theory (Maswood & Rajaram, 2019). However, examining the Mandela effect as a psychological phenomenon could shed light on the nature of memory representations and how false memories form. In this study, we provide a comprehensive examination of the VME, a specific type of Mandela effect in the visual modality, and investigate its consistency across people.

Corresponding Author:
Wilma A. Bainbridge, The University of Chicago, Department of Psychology
Email: wilma@uchicago.edu

Although visual memory is robust, with a high capacity for detail (Brady et al., 2008; Hollingworth & Henderson, 2002; Nickerson, 1965), it is still susceptible to errors. Specifically, schema-consistent details (i.e., aligning with expectations of the image) are more likely to be falsely remembered (Koutstaal & Schacter, 1997; Miller & Gazzaniga, 1998; Roediger & McDermott, 1995; Seamon et al., 2000). These schema-consistent errors even occur for highly familiar stimuli. Studies with specific coins (e.g., the U.S. penny) and stamps showed that people had remarkably poor memory for their features (Jones, 1990; Nickerson & Adams, 1979). Wong et al. (2018) showed that people misrecall what letter-form of "g" they read in print (reporting g rather than g). People falsely remember features on even the iconic Apple logo, adding stems and extra leaves (Blake et al., 2015). These results show that even highly familiar stimuli are susceptible to schema-consistent errors.
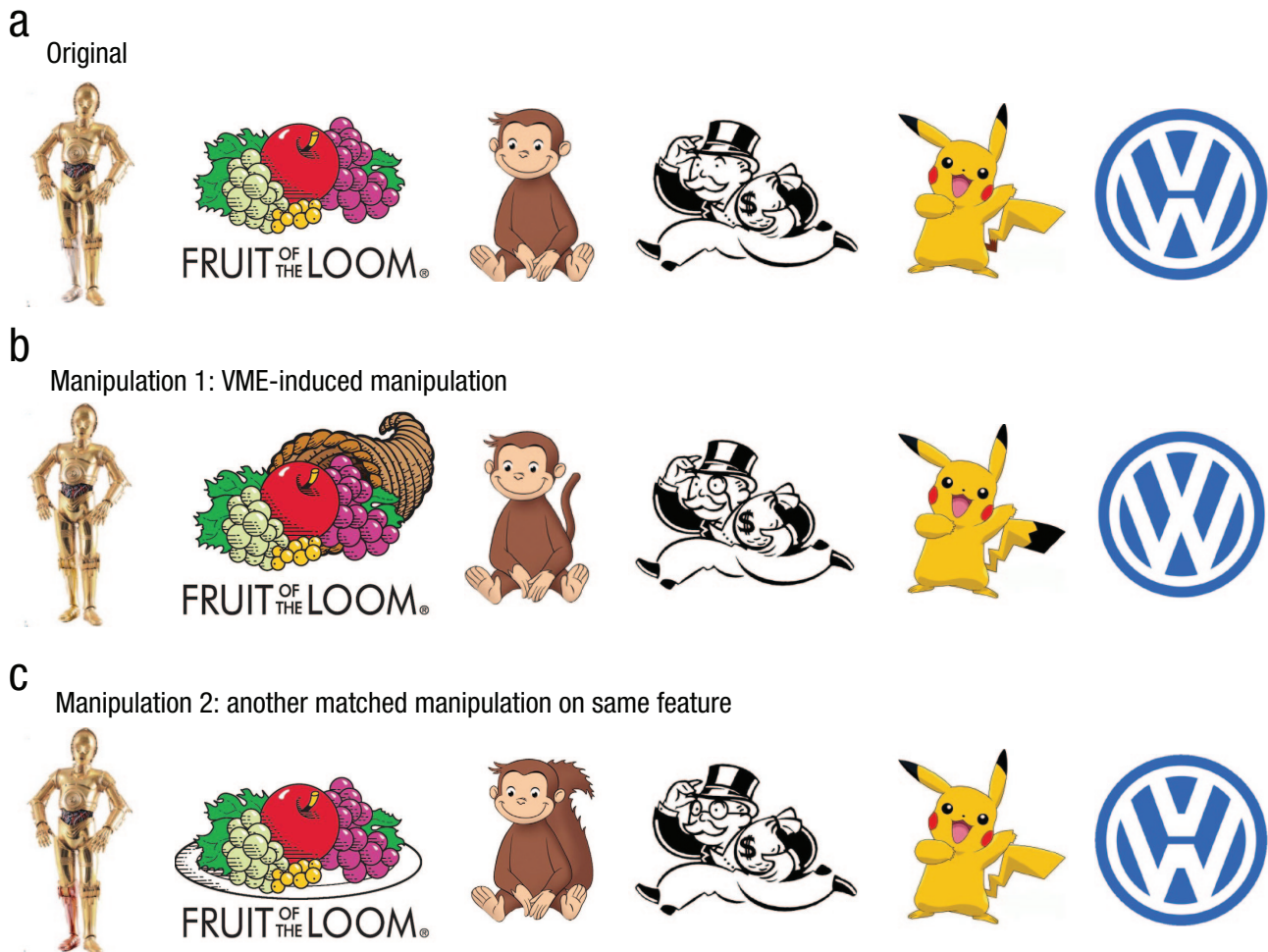
In casual popular science discussions, schema theory has been mentioned as a possible explanation for the VME (Dagnall & Drinkwater, 2018), although it has not been formally investigated. Indeed, in many reported VMEs, the erroneous feature is schema consistent. The Monopoly Man, for example, is the quintessential older rich man, and a monocle is stereotypical of this schema. However, there are defining aspects of the VME that this theory cannot fully account for. In previous studies of highly familiar stimuli, although falsely remembered features were schema consistent, they varied across people. Conversely, the false features reported from the VME are specific and consistent across people. Furthermore, the VME has been reported for only a select few visual icons; if it were simply a matter of schema-consistent errors, we would expect to see more images affected by the VME phenomenon. Given the high specificity and consistency reported with the VME, the schema explanation does not suffice by itself.

Consistent memory performance is not unique to the VME; people are highly consistent in their accurate memory for images, universally remembering some over others (Bainbridge, 2019; Bainbridge et al., 2013, 2017). These results suggest that a proportion of what dictates memory performance is intrinsic to the stimulus and independent of individual experience. These studies investigating the intrinsic memorability of images also show that images, despite being of a similar schema (e.g., faces), can elicit highly consistent memory behaviors across participants (Bainbridge et al., 2013). Whereas memorability work has mainly focused on successful recognition of images, some work also suggests high consistency in false recognition (Bainbridge et al., 2013). Applying this to the VME, perhaps something about the images themselves is what drives the effect.

Statement of Relevance

There are widespread informal reports on the Internet of many people having the same specific false memory for certain items, usually from popular culture, called the Mandela effect. Surprisingly, there are even images that people falsely remember, despite almost never seeing this false memory in the world (the visual Mandela effect [VME]). However, there has been no experimental confirmation of the VME or any attempt to characterize these false memories. Using a recognition task, we demonstrated that people have specific and consistent false memories for certain popular culture images, showing that the VME does exist. Furthermore, our study suggests that these false memories are not driven by low-level feature differences or attentional differences and that they can occur during recall. These data add to a growing body of surprising results showing consistency in what people remember by demonstrating new evidence that there is also consistency in what people misremember.

Here, we provide empirical evidence demonstrating that there are certain images that elicit a specific false memory, despite high familiarity and confidence. In Experiment 1, we examined the VME through a forced-choice recognition task, in which participants are asked to pick the canonical version of an image among a set of manipulated versions. In Experiment 2, we characterized the perceptual and attentional forces behind this effect through a computer-based method analogous to eye tracking. In Experiment 3, we quantified the natural visual experience for these images by scraping real-world images from the Internet. Finally, in Experiment 4, we tested whether the VME occurs during free recall with an image-drawing paradigm. We reveal that the VME exists for certain images, showing that there are consistencies in people's false memories for both recognition and recall. We observed these memory errors both when participants report long-term knowledge (Experiments 1 and 4) and when they report short-term memory (Experiments 2 and 4). Furthermore, we found that differences in the visual inspection of an image do not account for these false memories (Experiment 2), and there is no universal facet of real-world viewing experience that explains the effect (Experiment 3). These results suggest that there may not be a unitary account driving the VME, despite the high consistency across people.

a
Original

b
Manipulation 1: VME-induced manipulation

c
Manipulation 2: another matched manipulation on same feature

Fig. 1. Which image is the real version? (a) The original, unaltered images that are reported to be affected by the visual Mandela effect (VME). People are exposed only to this version of the images in their daily lives. (b) The reported VME-induced representation for which people have a strong false memory, despite never seeing this version. Their reported VME representations are a golden leg for C-3PO, a cornucopia for the Fruit of the Loom logo, a tail for Curious George, a monocle for the Monopoly Man, a black-tipped tail for Pikachu, and no bar between the "V" and the "W" of the Volkswagen logo. (c) An alternative, matched manipulation on the same feature that has not been reported as a VME representation.

Experiment 1

Based on informal reports of the effect, there are five defining characteristics of images affected by the VME: low identification accuracy, a specific false representation, high consistency across participants, high familiarity among participants, and high confidence in the wrong image. In Experiment 1, we aimed to establish whether there were images that met these criteria and were thus VME apparent by using a forced-choice recognition task.

Method

Participants. Participants on Amazon Mechanical Turk (MTurk), an online crowdsourcing platform for tasks, were screened for location (United States) and English comprehension. Participants were excluded if more than 50% of the questions were not answered. Because Experiment 1 served as a first exploratory analysis of the VME, we aimed to recruit a high number of participants to make judgments on each image; ultimately, 100 MTurk participants (35 female; age: $M = 39.4$ years, $SD = 12.3$) successfully completed this task. No personally identifiable information was collected from any participants, and participants had to acknowledge participation in order to continue, following the guidelines approved by The University of Chicago Institutional Review Board (IRB19-1395).

Stimuli. Forty image concepts were used to generate image sets for Experiment 1. Throughout the article, we use the term *image concept* to refer to the icon represented in the image (e.g., Pikachu) as opposed to the

specific image used. Six of these image concepts were informally reported in the popular press as affected by the VME and had an associated false-memory representation (Bakkila, 2019). These images included four characters and two brand logos: C-3PO from the Star Wars franchise, the Fruit of the Loom logo, Curious George from the Curious George series, the Monopoly Man from the Monopoly board game, Pikachu from the Pokémon franchise, and the Volkswagen logo (Fig. 1a). Their reported VME representations were a golden leg for C-3PO, a cornucopia for the Fruit of the Loom logo, a tail for Curious George, a monocle for the Monopoly Man, a black-tipped tail for Pikachu, and no gap between the "V" and the "W" of the Volkswagen logo (Fig. 1b). The remaining 34 images in this study were pulled from similar areas of popular iconography and had no previously associated false-memory representations. These included such icons as the Bluetooth symbol, Hello Kitty, and Bugs Bunny, for a total of 22 characters, 16 brand logos, and two symbols.

For each image concept, we created an associated *image set* made up of three versions of that image concept: an unaltered image and two manipulated images (Figs. 1a–1c). The unaltered version is the canonical version used as the official design for the icon. Manipulations were done in GNU Image Manipulation Program, an open-source image editor (The GIMP Team, 2020). Only one feature of the image was altered across manipulations. For the six images reported to cause the VME, the VME-induced representation was one of the manipulated versions (Fig. 1b). There were five possible manipulations: feature addition, where a new feature was added to the image; feature subtraction, where an existing feature was removed from the image; feature change, where an existing feature was altered but not removed; position/orientation change of a feature; and color change of a feature. The 80 manipulations included 26 feature subtractions, 21 feature additions, 24 feature changes, three position/orientation changes, and six color changes. Manipulations were intended to be relatively schema consistent and to fit the original design of the image. The manipulation type was not predictive of the correct answer, and the absence or presence of a feature on an image did not always mean that the image was the original. Thus, participants could not determine the correct image just from reasoning about the types of manipulations in the three versions of the image. All stimuli are accessible through our publicly available repository on the Open Science Framework, OSF (https://osf.io/7cmwf/).

Procedure. Each participant saw all 40 image sets with all three versions. Presentation order of both the image sets and the versions within the set was randomized. Participants were asked to choose the correct version (i.e.,

the canonical version) of the image from the three versions and rate their confidence in their choice on a Likert-type scale (1 = *not at all confident*, 5 = *extremely confident*). They also rated their familiarity with the image concept on a Likert-type scale (1 = *not at all familiar*, 5 = *extremely familiar*) and how many times they had previously seen the image concept (0, 1–10, 11–50, 51–100, 101–1,000, 1,000+). See Fig. 2 for an example question.

Results

Each image had an average of 98.1 responses. If a participant skipped the image-choice question, follow-up questions for that image were not included in the analysis. In every experiment of this study, an α of .05 was used for all tests.

For an image to show a VME, five criteria had to be met: (a) The image must have low identification accuracy, (b) there must be a specific incorrect version of the image falsely recognized, (c) these incorrect responses have to be highly consistent across people, (d) the image shows low accuracy even when it is rated as being familiar, and (e) the responses on the image are given with high confidence even though they are incorrect. Toward the first criterion, we calculated the percentage accuracy for each image (Fig. 3a). Five images had an identification accuracy below chance (< 33%), whereas two others were close to chance (34% and 35%). To determine whether a specific incorrect manipulation was chosen for these images, we first labeled the two different manipulations: For each image concept, Manipulation 1 refers to the manipulation that had the highest proportion of incorrect responses and Manipulation 2 refers to the remaining manipulation. Then, to determine whether the proportion of correct responses was significantly different from the proportion of Manipulation 1 responses, we performed a χ^2 test of independence. Of the 40 image concepts, 39 showed independence (all χ^2s ≥ 6.089; all ps < .014), indicating that the response rate significantly differed from the correct image and from the Manipulation 1 image. Of these, seven had a significantly higher proportion of Manipulation 1 than correct responses, whereas the remaining 32 had a significantly higher proportion of correct responses than Manipulation 1 responses. These results indicate that these seven images (C-3PO, Curious George, the Fruit of the Loom logo, the Monopoly Man, Pikachu, the Volkswagen logo, and Waldo from Where's Waldo?) had not only accuracies below chance but also specific incorrect versions that were falsely recognized as the original. Thus, these seven images were labeled "VME apparent."

VME is also defined by its consistency; it is a *shared* specific false memory. We used a split-half consistency

Question 1 out of 40

a. Choose the correct image for **the Apple logo**:

○   ○   ○

b. How **confident** are you in your choice? *(1 is not at all confident, 5 is extremely confident)*

not at all confident   ○ 1   ○ 2   ○ 3   ○ 4   ○ 5   extremely confident

c. How **familiar** are you with the Apple logo? *(1 is not at all familiar, 5 is extremely familiar)*

not at all familiar   ○ 1   ○ 2   ○ 3   ○ 4   ○ 5   extremely familiar

d. How **many times** have you seen the Apple logo before?

○ 0   ○ 1-10   ○ 11-50   ○ 51-100   ○ 101-1000   ○ 1000+

Fig. 2. An example question from the forced-choice recognition task. For 40 icons, participants were asked to choose the correct version from a set of three and to rate their confidence in their choice, their familiarity with the image concept, and the approximate number of times they have seen the image concept.

analysis to determine whether people were consistent in the image choices they made. Participants were randomly split into two halves; for each half, the proportion of correct responses and Manipulation 1 responses was calculated for each image. We then calculated the Spearman rank correlation for each response type between the participant halves, across 10,000 random shuffles of participant halves. The mean Spearman rank correlation across the iterations for the proportion of correct responses was 0.92 ($p < .0001$; Fig. 3b). The mean correlation across the iterations for the proportion of Manipulation 1 responses was 0.88 ($p < .0001$; Fig. 3b). This suggests that people are highly consistent in what images they respond correctly and incorrectly to. In other words, just as people correctly remember the same images as each other, they also have false memories of the same images as each other.

To analyze how familiarity and confidence compare between VME-apparent images and the other images, we compared the mean confidence, mean familiarity, and accuracy for each image (Fig. 4). As expected, there was a significant correlation between mean confidence and mean familiarity across all images ($r = .90$, $p = 2.64 \times 10^{-15}$), indicating that people made more confident responses for the items that were familiar. However,

there was a surprisingly much lower correlation between accuracy and confidence ($r = .59$, $p = 5.61 \times 10^{-5}$) and between accuracy and familiarity ($r = .42$, $p = .007$), although still significant. Visual inspection of these plots showed that eight images fell below the distribution (Figs. 4b and 4c), where these images had low accuracy but high familiarity and confidence ratings. Seven of these eight images were the seven VME-apparent images identified by the χ^2 test. To determine whether VME-apparent images were driving these decreased correlations between memory accuracy with familiarity and confidence, we ran a permutation test (Fig. 4). We randomly dropped seven images from the sample and calculated the new correlation coefficient 1,000 times. We compared the distribution of these correlation coefficients for these randomly removed samples with the correlation coefficient for the sample when only the seven VME-apparent images were removed. If the correlation increased when VME-apparent images were removed from the data compared with a random set of seven, it would suggest that these specific images were negatively impacting the correlation. There was a significant effect of VME-apparent images on the correlation between accuracy and familiarity ($p = .044$) and between accuracy and confidence ($p = .001$) but not
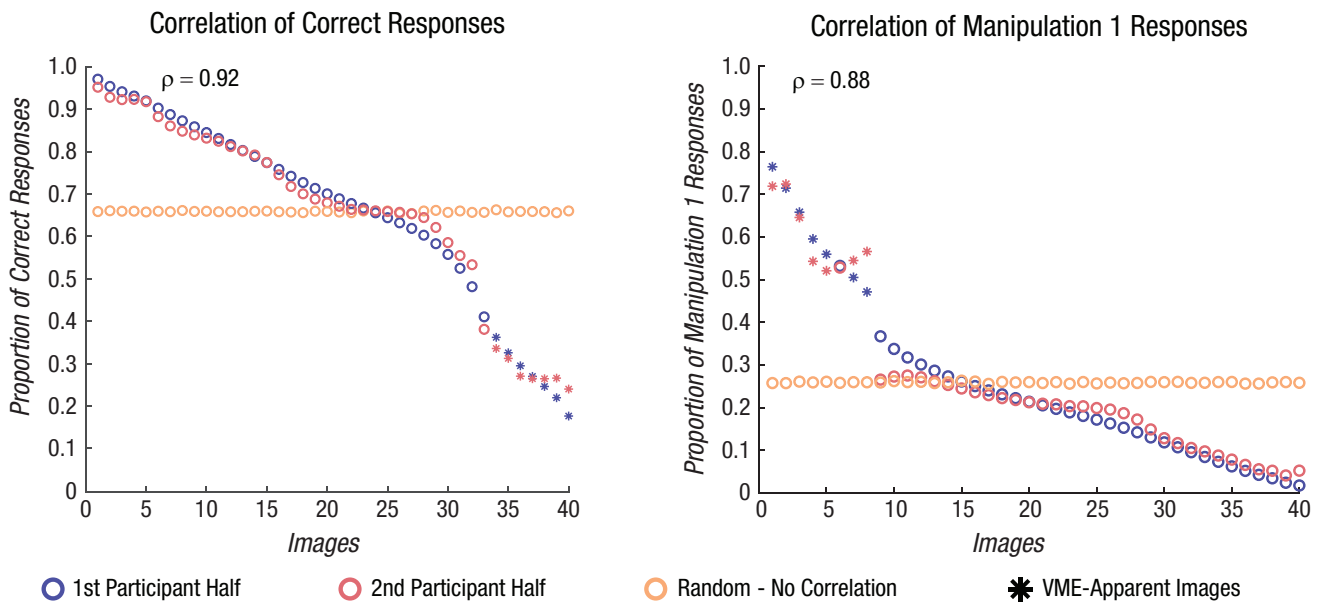
a

Ratio of Response Choices by Images

b

Split-Half Consistency Analysis

Correlation of Correct Responses
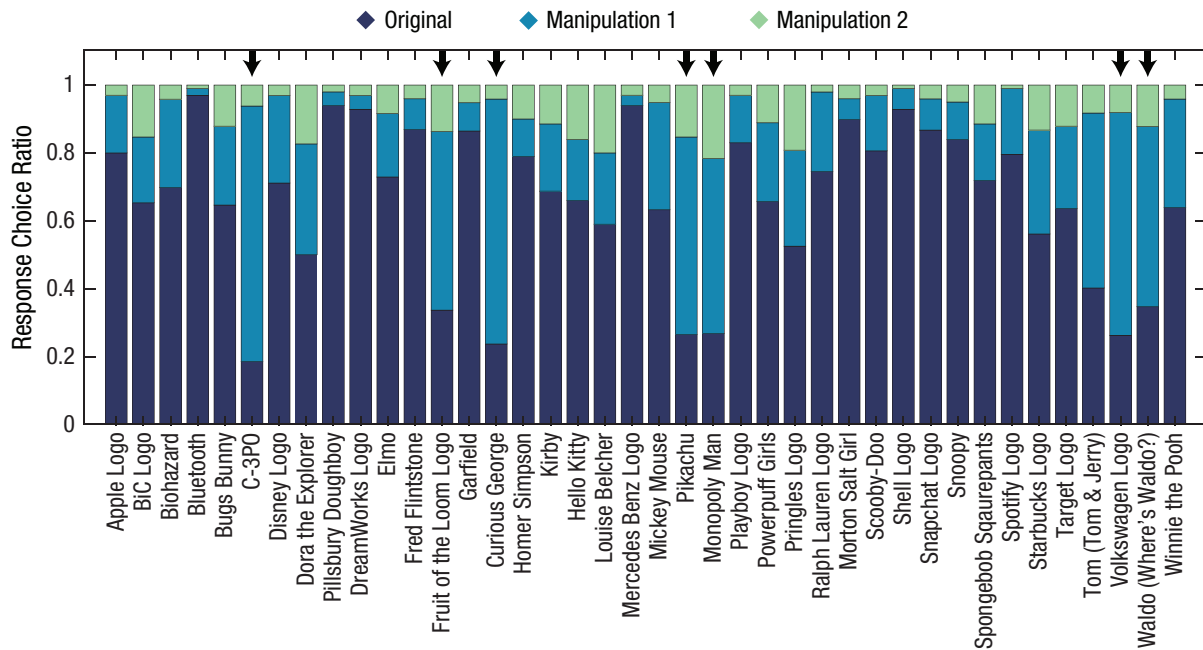
Correlation of Manipulation 1 Responses

Fig. 3. Seven images show shared and specific incorrect responses. (a) Seven images had identification accuracy below or near chance (33%). For these seven images, denoted by the arrows, the proportions of correct responses and Manipulation 1 responses were independent (all χ^2s ≥ 6.089; all ps < .01), with a significantly high proportion of Manipulation 1 responses, indicating that a specific incorrect version was chosen. (b) The Spearman rank correlation across the 10,000 shuffled participant halves showed high consistency across participants for both correct responses ($\rho = 0.922$, $p < .0001$) and Manipulation 1 responses ($\rho = 0.876$, $p < .0001$). Visual Mandela effect (VME)-apparent images, indicated by the colored asterisks, fall at the extremes of the proportions.
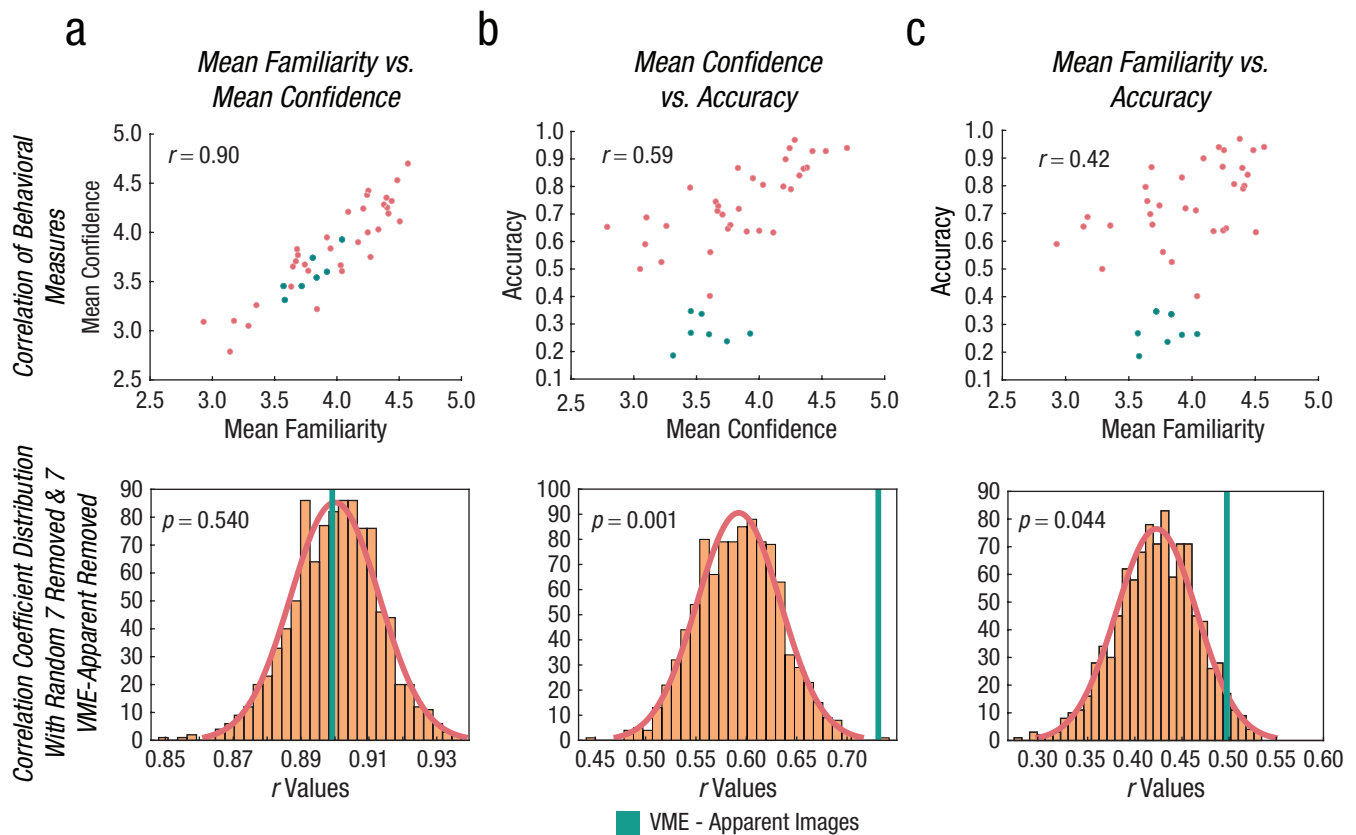
Fig. 4. Visual Mandela effect (VME)-apparent images show poor accuracy despite high familiarity and confidence. Below each correlation is the distribution of the correlation coefficients when seven random images were removed from the sample. The correlation coefficient when the seven VME-apparent images were removed from the sample is indicated by the green line. (a) Mean image familiarity ratings were strongly correlated with confidence ($r = .90$, $p = 2.64 \times 10^{-15}$). The seven VME images did not negatively impact this correlation ($p = .540$). (b) The correlation of confidence and accuracy was low ($r = .59$, $p = 5.61 \times 10^{-5}$); when the seven VME-apparent images are removed, the correlation significantly increases ($p = .001$). (c) This is similarly the case for the correlation between familiarity and accuracy ($r = .42$, $p = .007$); removing the seven VME-apparent images significantly increases the correlation ($p = .044$).

for the correlation between confidence and familiarity ($p = .540$). This suggests that the seven VME-apparent images did not fit into the overall trend; for these images, accuracy did not increase with increasing familiarity and confidence. Furthermore, their accuracy is surprisingly low given the reported familiarity and confidence that people had with these images. There was also no significant difference between the number of times that participants had seen VME-apparent images and the number of times they had seen the images that were correctly identified (Wilcoxon rank sum; $z = 0.64$, $p = .523$), supporting the idea that there is no difference in prior exposure between VME-apparent images that induce false memory and images that do not.

Together, the analyses show that seven images meet the five criteria of VME that we previously defined. First, the response choice analysis shows that these seven images have low identification accuracy. Second,

the results of the χ^2 test show that a specific representation was falsely remembered for these images. Furthermore, the split-half consistency analysis provides evidence that this incorrect choice is highly consistent across participants. Finally, the permutation tests suggest that these seven images have high familiarity and confidence, despite their poor accuracy. The combination of these results demonstrates a true VME, showing that certain images elicit shared and specific false memories.

Experiment 2

We next sought to understand why these seven images induce a specific shared false memory. Specifically, we wanted to test how attention and perceptual processing of the images would influence these false memories—are the features that elicit false memories viewed differently from those that are correctly remembered? For
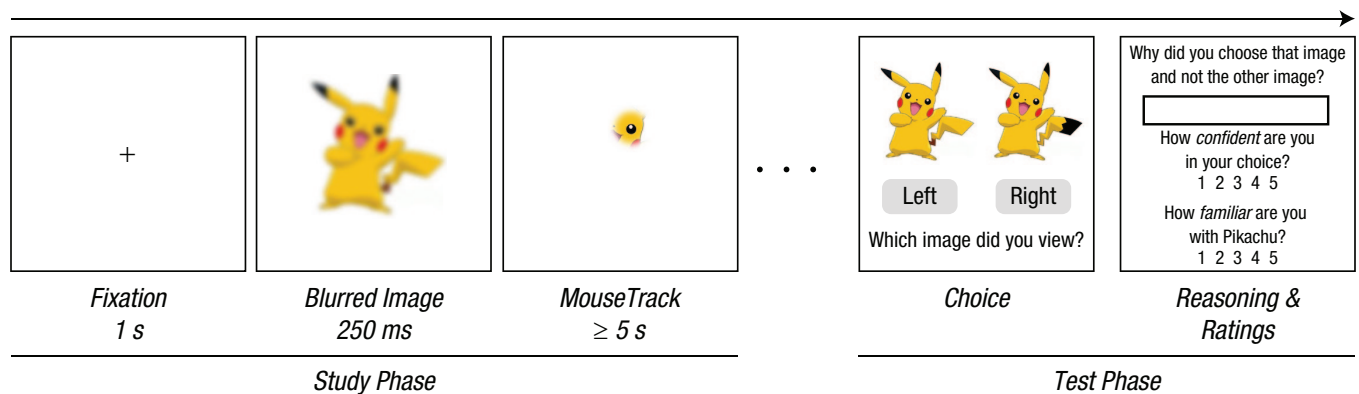
Fig. 5. Experiment 2 method. During the study phase, participants inspected the correct version of seven visual Mandela effect (VME) images and seven matched non-VME images. A blurred version of the image was shown for 250 ms, and then participants had to move a circular aperture to reveal the image underneath, approximating fixation behavior. Participants were told to examine each image and then were allowed to move on to the next trial after at least 5 s of inspection. During the test phase, participants had to indicate their memory for the images they inspected, choosing between the correct original image and an incorrect manipulated version. Finally, participants answered a series of questions about why they chose that image, their confidence in their decision, and their level of familiarity with the image concept.

example, perhaps the VME occurs because the feature is not fixated on during normal perception, and thus people "fill in" that region with prior knowledge. Another possibility is that the VME could result from a source memory error, where participants have actually seen a noncanonical version of the icon prior to the experiment and think that is the correct version. To address these possible explanations, we conducted an experiment using a short-term memory task with a mouse-tracking technique similar to eye tracking. By using a short-term memory task, we could test whether these errors still occur shortly after the perception of the canonical image. In addition, the mouse-tracking technique allowed us to see how participants observe these images and how inspection patterns may relate to later false memories. We also examined the differences between images using low-level visual feature analysis and polled participants' intuitions for their decisions.

Method

Participants. Participants on Prolific, an online crowdsourcing platform for tasks, were screened for location (United States) and English comprehension. Participants were excluded if they failed the attentional check or if they did not follow task instructions; one participant was excluded for not moving their mouse during the task. Sixty participants from Prolific.co (32 female; age: $M =$ 34.7 years, $SD = 13.5$) successfully completed this task. The participant sample size was selected to match the sample size of a related experiment (see BubbleView Experiment in the Supplemental Material available online). No personally identifiable information was collected from any participants, and participants had to

acknowledge participation in order to continue, following the guidelines approved by The University of Chicago Institutional Review Board (IRB19-1395).

Stimuli. Because only seven of the 40 image sets used in Experiment 1 were VME apparent, we tested the VME-apparent image concepts and a matched subset of seven non-VME control image concepts, selected as those that were high in accuracy but matched in familiarity (independent-samples t test), $t(12) = 0.11$, $p = .911$ (VME apparent: $M = 3.78$, $SD = 0.17$; matched subset: $M = 3.77$, $SD = 0.15$).

Procedure. To determine how people viewed these images, we used MouseView, a mouse-tracking method analogous to eye tracking (Anwyl-Irvine et al., 2021). Generally, a target image is obscured by a white overlay, and MouseView creates a circular aperture that moves with the computer mouse to reveal the image underneath it (Fig. 5). By moving their cursor around, participants can continuously uncover small sections of the target image at a time, imitating foveation. MouseView measures have been shown to significantly predict real human fixation measures (Anwyl-Irvine et al., 2021). In the current experiment, all MouseView images were presented at a size of 500 × 500 pixels with an aperture of 5% of the viewing window. We used the jsPsych implementation of MouseView (de Leeuw, 2015) and hosted the experiment on Cognition, an online jsPsych experiment platform (https://www.cognition.run).

This experiment consisted of two main phases: the study phase and the test phase. Each phase had 14 trials, one for each VME and matched non-VME image. For the study phase, participants were directed to

examine each image using their mouse cursor to move around and reveal the image underneath a white overlay. Participants saw only the correct version of each image concept (e.g., the canonical Monopoly Man without a monocle). Each image was preceded by a rapidly presented (250 ms) blurred version of the image (10 sigma Gaussian), in order to approximate the first gist-like view on an image that people use to guide their inspections (Henderson, 2007). Participants were required to inspect the image for a minimum of 5 s before continuing to the next trial, which was preceded by a 1-s fixation cross. Importantly, participants were not explicitly told that they would be tested later on their memory of the images, in order to promote more naturalistic inspection. In the test phase, for each image set, participants were shown the correct image (that they studied) and the Manipulation 1 version of the image concept and chose which they had seen during the study phase. Location of the choices (left or right) was randomly determined. After making their selection, participants then reported why they chose that version over the other version. Participants also rated their confidence in their choice and their familiarity with the image concept, both on Likert-type scales ranging from 1 to 5. Finally, at the end of the experiment, participants reported what they thought the experiment was about and whether they had heard of the Mandela effect prior to the experiment.

MouseView inspection-density-map generation. For each participant response, we created a map of the two-dimensional points in which people fixated during the study phase. We converted this into an inspection density map for the image using a thin plate smoothing spline interpolation with a 50-pixel blur radius around each point, akin to prior work generating semantic heat maps for images (Henderson & Hayes, 2017). Higher inspection density values indicate that there was more inspection in that area, suggesting that participants examined that area of the image more.

Results

First, we confirmed whether we replicated our results from Experiment 1. If the VME occurred in Experiment 2, we would expect to see low accuracy for the VME subset compared with the matched subset, despite equivalent familiarity and confidence ratings. Indeed, between our image subsets, we observed no significant difference in familiarity (independent-samples t test), $t(12) = -0.98$, $p = .347$, Bayes factor (BF_{01}) = 1.64 (anecdotal evidence for the null); VME apparent: $M = 3.65$, $SD = 0.15$; matched subset: $M = 3.82$, $SD = 0.43$, or

confidence (independent-samples t test), $t(12) = -0.93$, $p = .369$; $BF_{01} = 1.69$ (anecdotal evidence for the null); VME apparent: $M = 3.24$, $SD = 0.38$; matched subset: $M = 3.48$, $SD = 0.55$. However, VME-apparent images had a significantly lower accuracy than matched non-VME images (independent-samples t test), $t(12) = -4.82$, $p = 4.19 \times 10^{-4}$, $BF_{10} = 57.28$ (very strong evidence for the alternate hypothesis); VME apparent: $M = 52.86\%$, $SD = 15.02\%$; matched subset: $M = 82.38\%$, $SD = 6.07\%$. This low accuracy for the VME image set is remarkable, given that participants had just seen the correct image minutes prior during the study phase, yet they still chose the false version to indicate their memory.

We asked participants to report their reasoning behind choosing one image version over the other during the test phase. For each image, we split the responses by participants who responded correctly and those who responded incorrectly. We determined the proportion of responses that mentioned the manipulated feature in their reasoning and the proportion of responses that were guesses or had unclear reasoning. For both image types (VME and matched), when participants answered correctly, they most often attributed this to memory for a specific feature (for 78.32% of the VME responses and 71.57% of the matched responses; no significant difference, $\chi^2 = 3.446$, $p = .063$)—for example, they "only saw the fruit, not the cornucopia" when inspecting the Fruit of the Loom logo. However, participants who chose the incorrect variation of the VME-apparent image concepts also reported remembering seeing the manipulated feature, even though they did not. For example, for the Fruit of the Loom logo, participants reported that "I'm pretty sure it had a basket" and "there was a cornucopia in the image I saw" (for all responses, see https://osf.io/7cmwf/). In fact, incorrect responses to VME-apparent images were more often attributed to memory of the manipulated feature (66.54%) than those to matched non-VME images (44.92%), which instead tended to be more guess based (χ^2 test of independence; $\chi^2 = 10.466$, $p = .001$).

In this experiment, we collected MouseView cursor data, a measure comparable with eye tracking, to determine whether there were inspection differences that may drive VME. We analyzed these data on two key levels: by the condition and by the image. For the condition-level analysis, we examined the effects of three factors on the inspection density: condition (VME, matched), response type (correct, incorrect), and area of image (manipulated area, unchanged area). We found a significant effect for area, where average inspection density was higher in the area that was manipulated than outside that area, $F(1, 48) = 35.91$, $p = 2.58 \times 10^{-7}$, $\eta_p^2 = .428$, $BF_{10} = 2.69 \times 10^5$, extreme

evidence for the alternate hypothesis. This could be due to the low inspection outside the image borders; indeed, when we look at an alternate measure of maximum inspection density inside or outside of the manipulated area, there is no longer a significant effect ($p = .091$; $BF_{01} = 1.33$, anecdotal evidence for the null hypothesis). However, importantly, even for average inspection density, there was no difference in inspection behavior related to whether the image caused VME or not ($p = .143$; $BF_{01} = 1.90$, anecdotal evidence for the null hypothesis) or whether participants responded correctly or not ($p = .952$; $BF_{01} = 6.06$, moderate evidence for the null hypothesis). There were also no significant interactions across factors (all $ps > .50$; all $BF_{01}s > 3.70$, moderate evidence for the null hypothesis). This indicates that inspection behavior did not differ between images that caused false memories and those that did not.

We ran a similar analysis at the individual image level for all seven VME-apparent images (Fig. 6; see Fig. S2 in the Supplemental Material). Specifically, we tested two-way analyses of variance (ANOVAs) with factors of the area of the image (manipulated area, unmanipulated area) and response type (correct, incorrect). No images showed a significant interaction of inspection area by response type (all $ps > .10$; all $BF_{01}s > 2.20$, anecdotal to moderate evidence for the null hypothesis; see Table S1 in the Supplemental Material). This implies that no images showed a difference in inspecting behavior between when participants had a correct memory and when they had a false memory. We conducted an additional experiment to compare allocation of attention across the original versions of these image concepts and their manipulations and similarly found no differences (see BubbleView Experiment in the Supplemental Material). Thus, across two experiments, we found no evidence for differences in inspecting behavior for images that later elicit false memories.

Whereas the VME does not appear to be caused by differences in attention, we also tested whether the VME could be explained by differences in low-level visual features. If so, we would expect to identify visual differences between the image versions that do cause VME and those that do not. A comparison of spectra energy (i.e., the spatial frequency of an image) across the image versions found no significant difference (all $ps > .660$). A comparison of the proportion of nonwhite, foreground pixels also found no significant difference across the image versions (all $ps > .783$). Finally, we compared the color composition of these image versions, which again showed no significant difference across the red, green, and blue (RGB) channels (all $ps > .158$). These results show that there are no differences in low-level visual features between the images

that induce VME and the images that do not induce VME that could explain the phenomenon.

Finally, at the end of the experiment, we asked participants questions to ascertain whether they knew of the VME. Of the 58 participants who responded to the question "What do you think this study was about?" only five responses (8.62%) mentioned the Mandela effect or false memory; 67.92% of the participants who responded reported having heard of the Mandela effect prior to the experiment.

From this experiment, we replicated the VME, showing that these specific images elicit false memories, even when the canonical images are shown right before memory is tested. This suggests that source memory error cannot fully explain the VME because it is unlikely that the noncanonical version from prior experience is overriding their recent experience of the canonical image. We also found that participants' reasons for their choices did not differ by image type—whether correct or incorrect, they mentioned strong memories of the manipulated feature in guiding their choices. By analyzing inspection behavior through mouse tracking, we found no differences in inspection patterns related to the type of image or people's memory, showing that images do not cause the VME because of differences in attention or perception. Although this experiment does not capture all facets of perceptual encoding, it does suggest that participants are not misremembering these VME features because they fail to look at them. We also found that the VME is not due to differences in low-level visual features. Finally, although few people reported thinking that the experiment was about VME, more than two thirds reported having heard of the Mandela effect prior to the experiment.

Experiment 3

Although we observed no link in Experiment 2 between inspecting behavior and false memories, it is possible that these false memories were caused by differences in the accumulated viewing experience of the cultural icons over time. For example, perhaps people incorrectly remember the color of C-3PO's leg because his legs are rarely shown in the Star Wars movies. Or perhaps they have even seen the VME version of C-3PO, given that the Mandela effect has been covered in the popular media. Indeed, although only a few participants guessed that Experiment 2 was about VME, many people reported having heard of the phenomenon previously. To quantify the real-world visual experience of these icons, we conducted an experiment in which we automatically scraped images of these icons from Google Images and then quantified the presence of VME features.
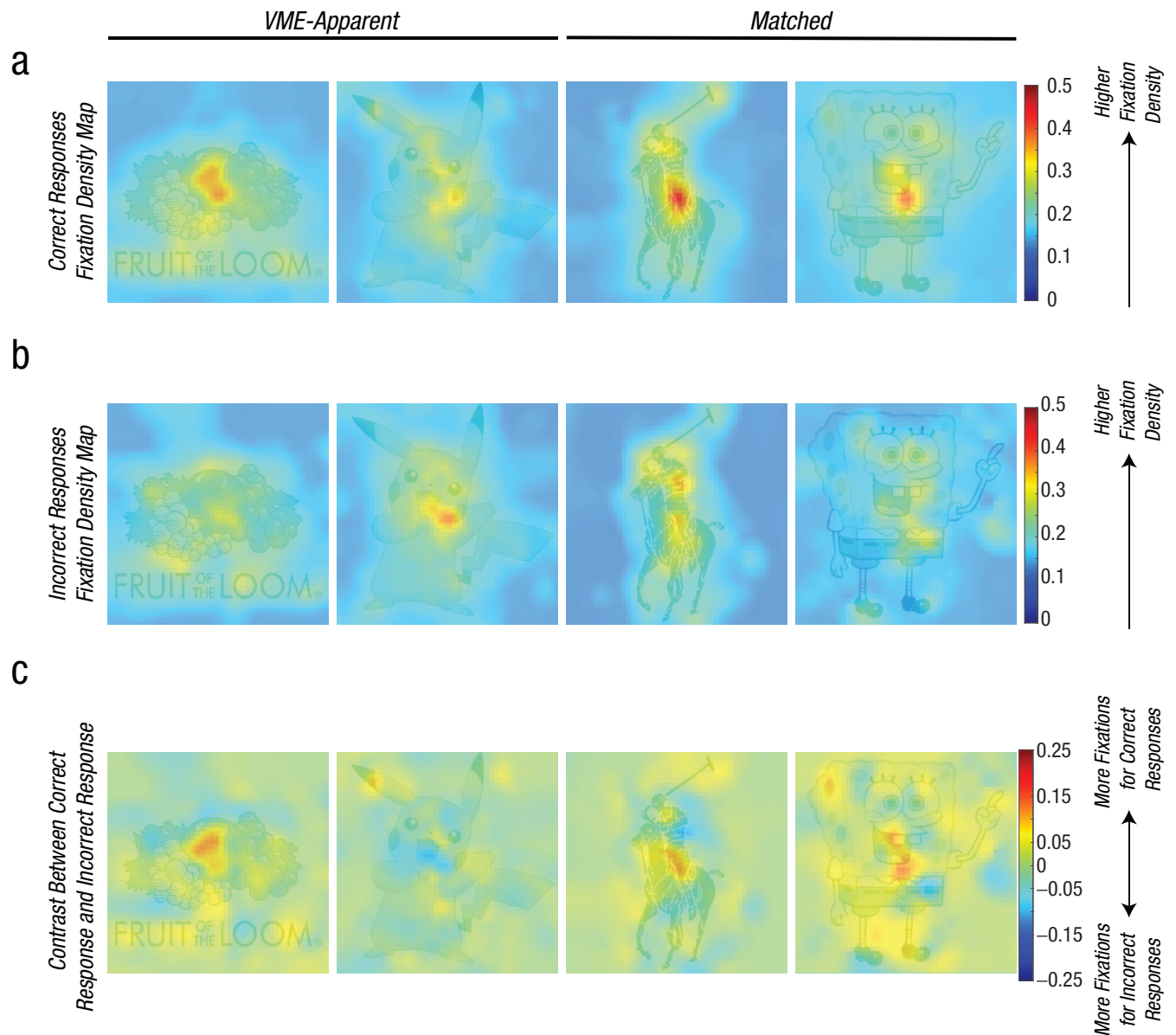
Fig. 6. Inspection differences do not drive the visual Mandela effect (VME). Inspection density heat maps using the MouseView interface: for two example VME-apparent images and two example matched images. (a) The inspection density map of inspecting behavior for participants who correctly remembered the images. (b) The inspection density map of inspecting behavior for participants who had a false memory for the images. (c) The difference in inspection behavior between participants who correctly versus incorrectly remembered the images. Crucially, there were no significant differences in inspecting behavior related to the type of image that was being viewed (VME or matched) or whether their memory was correct or incorrect.

Method

To approximate the natural viewing experience of the VME-apparent icons, we automatically scraped the top 100 Google Image results for each icon (Clinton, 2021), generated by queries of the name of each icon (e.g., "Volkswagen," "Pikachu," "Monopoly"). We then categorized these images into three different groups: (a) those that are unable to show the VME because the feature is not present (e.g., a headshot of C-3PO without his legs visible), (b) those that show the full feature and have no VME in the image (e.g., a full-body photo including C-3PO's silver leg), and (c) those that show the full feature and do have the VME in the image (e.g., a full-body photo of C-3PO with two golden legs). For images that do show the VME, we also categorized how many of these images originated from sources specifically describing the Mandela effect.
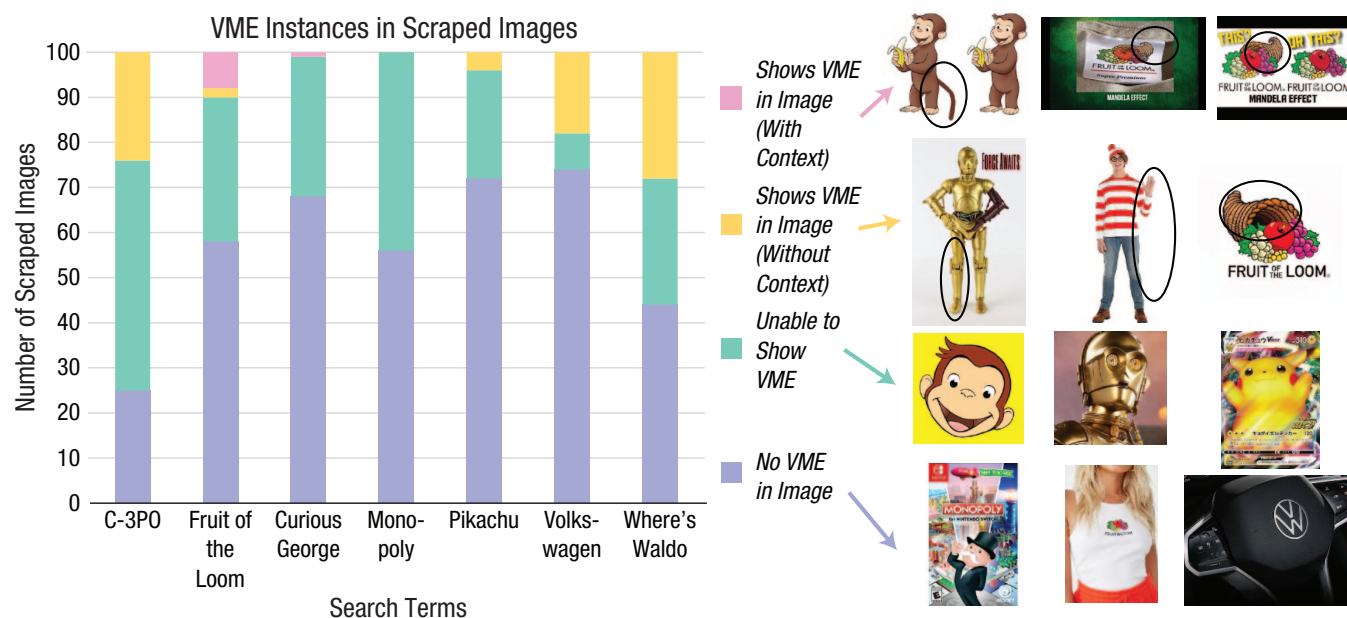
Fig. 7. Quantifying the natural viewing experience of visual Mandela effect (VME) images. For each of the seven VME-apparent icons, we scraped the top 100 images from the Internet and categorized them as not showing the VME in the image (i.e., having the correct feature), unable to show the VME, showing the VME in the image without context, or showing the VME in the image within a context related to the Mandela effect. There was no unifying account across images; for some images, there were examples of the VME (e.g., C-3PO, Where's Waldo), but for others, the correct version (i.e., the Monopoly Man without a monocle) made up the majority of the images. Examples of scraped images from each category are shown to the right, and examples containing the VME feature are circled in black.

Results

The distribution of scraped images across these four categories (unable to show VME, no VME in the image, VME in the image, and VME in the image from a source about the VME) is shown for each of the seven VME-apparent images in Fig. 7 (also see Fig. S3 in the Supplemental Material; for all scraped images, see https://osf.io/7cmwf/).

We observed high variation in the natural experience with these image concepts. For C-3PO, a majority of his images do not contain his legs (51%) or do contain the VME (24%; golden legs), suggesting that people's visual experiences of his legs may be incomplete or inaccurate. For other examples (Where's Waldo and Volkswagen), a substantial portion of the images show the VME (28% and 18%, respectively), although there are still more examples of the canonical feature (44% and 74%, respectively). For the remainder of the image concepts, the majority of the scraped images do show the correct feature of importance without a VME. For these examples (Pikachu, Fruit of the Loom, the Monopoly Man, and Curious George), people will rarely, if ever, encounter the VME version in the real world. For some (Fruit of the Loom, Curious George), when the VME version is encountered, it is in the context of a source about the Mandela effect, suggesting that the

existence of the memory error precedes visual examples of it. Therefore, the VME can also occur in spite of extensive experience with the correct version of the image.

Although this experiment does not fully address the issue of how common the VME features are during natural viewing, these results do suggest that it is possible for people to have experience with the noncanonical, VME versions of the images. Specifically, whereas most of the images that people are exposed to in their day-to-day lives are the canonical version, for some VME-apparent images (e.g., C-3PO, Where's Waldo), people may be exposed to the noncanonical VME version, which could distort their memory for that icon and thus cause them to remember the different version. However, for other images (e.g., Pikachu, Monopoly Man), the results suggest that people are not often exposed to the incorrect version in the real world. This could mean that something else may drive the VME for these specific images or that there is something particularly memorable about these noncanonical versions so that minimal exposure is enough to overcome memory for the more frequent canonical versions. Importantly, these results suggest that there is no single unifying account for how prior perceptual experiences could cause these visual false memories to occur.

Long-Term Task: Familiarity Score ≥ 3
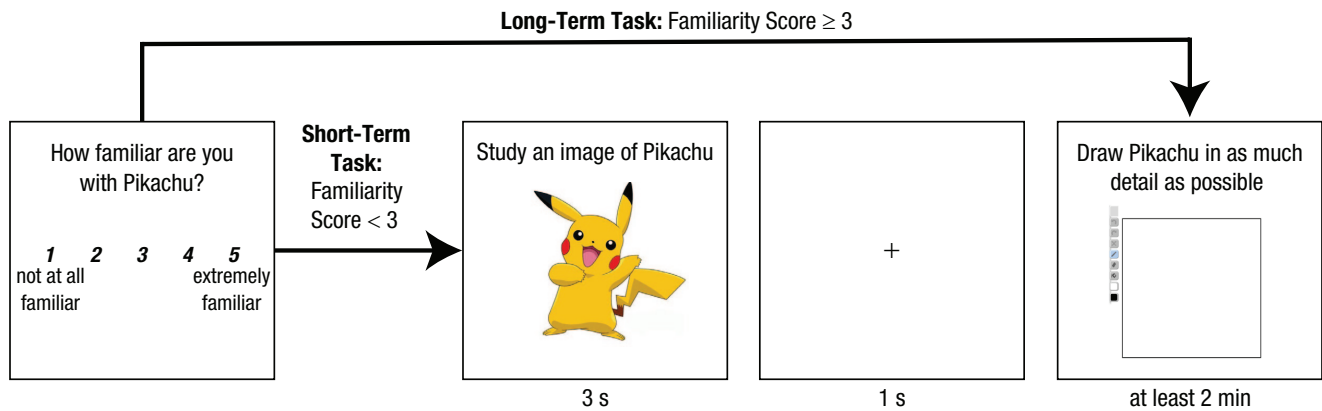
Fig. 8. Familiarity-based drawing recall paradigm. Participants first indicated their level of familiarity with a given image concept (e.g., "Pikachu"). If participants rated their familiarity with the image concept as low (< 3), they did a short-term recall task in which they studied the associated image for 3 s and, after a brief delay, drew the image from memory with as much detail as possible. For participants who were highly familiar with the image concept (≥ 3), they drew the image from their long-term memory in as much detail as possible.

Experiment 4

Finally, we tested the strength of the VME across memory paradigms; in Experiments 1 and 2, we used visual recognition paradigms to elicit VME. However, a stronger test of its effect is whether people spontaneously produce these VMEs unprompted. Here in Experiment 4, we used an online drawing task to assess whether VME occurs during free recall of an image (Bainbridge et al., 2019). By using different task paradigms depending on icon familiarity, this experiment also allowed us to reproduce whether the VME occurs from highly familiar, long-term knowledge of the icon (as in Experiment 1) as well as at shorter time scales (as in Experiment 2).

Method

Participants. Participants on MTurk were screened for location (United States) and English comprehension. Participant responses were also excluded if they did not answer the required questions or if their submitted drawings did not depict the prompted image concept. Prior studies of memory drawings generally collected drawings from 15 to 30 participants per image (Bainbridge et al., 2019, 2021). To ensure that we had high enough power to detect effects, we decided to collect 50 participants for each image, resulting in 459 total MTurk participants (277 female; age: $M = 38.1$ years, $SD = 11.1$) who successfully completed the task. No personally identifiable information was collected from any participants, and participants had to acknowledge participation in order to continue, following the guidelines approved by The University of Chicago Institutional Review Board (IRB19-1395).

Stimuli. We tested participants on a subset of 32 of the images from Experiment 1, including all VME-apparent images, one other low-accuracy image, and eight images each matched for low, medium, and high familiarity. Here, we report only comparisons of the VME-apparent images and their familiarity-matched controls, but all data are available on OSF (https://osf.io/7cmwf/).

Procedure. Participants were first presented with the name of an image concept (e.g., "Pikachu") and rated their familiarity with the image concept on a 5-point Likert-type scale (1 = *not at all familiar*, 5 = *highly familiar*; Fig. 8). The subsequent task differed on the basis of their response type. Participants who responded with low familiarity (scores < 3) did a short-term memory recall task. In that short-term recall task, participants would see a 500- × 500-pixel display of the image associated with the image concept for 3 s and, after a 1-s delay with a fixation cross, were asked to draw the image concept from memory for at least 2 min, but with no time limit. In contrast, participants who responded with high familiarity (scores ≥ 3) did a long-term memory recall task. They were asked to draw the familiar image concept from their memory for at least 2 min, with no time limit and without seeing the image. The differential tasks based on familiarity response allowed participants to submit a drawing of the image concept regardless of their familiarity. For both tasks, there was no time limit for the drawings.

The online drawing apparatus used in the procedure was adapted from the work by Bainbridge et al. (2019). Participants were presented with a 500 × 500 pixel blank square when asked to draw the image concept from memory. The drawing interface operated like a simple paint program: Participants could draw with a pen in multiple colors, fill sections in multiple colors,

erase lines, and undo or redo actions. Participants were not limited in what devices they could use to draw but were highly discouraged from using their fingers on a touch-screen device because that made the drawing interface more difficult to use.

Three independent scorers rated each image for errors on the basis of missing, added, or changed features but not drawing quality. The scorers listed and counted the errors in each drawing; if the drawing was of a VME-apparent image, they first noted whether it was possible for the VME error to exist (e.g., drawing only C-3PO's face and excluding his legs) and, if so, whether the VME error was made. For short-term memory drawings, errors were determined by comparing the drawing directly with the image that was presented during the task. Because no image was presented during the long-term memory task, the scorers were instructed to determine drawing errors not on the basis of a specific image but, rather, on general portrayals of the image concept. Using only the errors identified by all three scorers, we calculated the most frequent long-term memory and short-term memory drawing error for each image.

Results

Each image had an average of 50.1 responses, with an average of 13.2 short-term recall responses and 36.9 long-term recall responses.

To see whether recall induces the same VME errors that recognition does, we calculated the frequency of those errors for VME-apparent images in both the long-term and short-term tasks. Because there was no image presented to the participant in the long-term task, it was possible for the participants' drawings to not include the area necessary for the VME feature to exist; therefore, the frequency of VME error was calculated only from drawings that included the area where the VME feature could occur (Fig. 9). For long-term task drawings, the total frequency of VME errors was 47.72%, which means that roughly half of all images drawn from long-term memory that could have the VME feature did have the VME feature. However, the frequency of the VME feature for each image varied (M = 51.27%, SD = 36.22%), suggesting that the occurrence of VME features during long-term recall is highly dependent on the individual image. C-3PO, Waldo, Curious George, and the Volkswagen logo had the highest frequency of VME error in long-term recall (94.44%, 92.86%, 57.69%, and 43.33%, respectively), whereas the Fruit of the Loom logo, Pikachu, and the Monopoly Man had the lowest (16.13%, 22.73%, and 31.71%, respectively). Some drawings from participants of the short-term memory version of the task, who were

unfamiliar with the image concept, still showed a VME error, although fewer than for the long-term memory task (total frequency = 20.97%, M = 27.49%, SD = 37.23%). On a by-image basis, 100% of C-3PO drawings, 9.09% of Curious George drawings, 50.00% of the Monopoly Man drawings, and 33.33% of Waldo drawings showed their associated VME error during short-term recall. No other VME-apparent images displayed VME error in their short-term memory drawings.

We then compared how the frequency of VME drawing errors compared with the frequency of drawing errors for non-VME images. We used the seven images from the medium familiarity category established in Experiment 2 to serve as a familiarity matched, non-VME control. For each image in this category, we determined the drawing error with the highest frequency, as scored by the three drawing scorers. Using a two-way ANOVA, we tested the effects of image category (VME or non-VME) and memory type (long-term or short-term) on the frequency of the most frequent drawing error of each image. There was a significant main effect of image category, $F(1, 24) = 11.17$, $p = .003$, $\eta_p^2 = .292$, but no significant effect of memory type, $F(1, 24) = 1.80$, $p = .193$, $\eta_p^2 = .047$, or any significant interaction between image category and memory type, $F(1, 24) = 1.25$, $p = .275$, $\eta_p^2 = .033$. VME-apparent images had a significantly higher frequency of drawing error (M = 39.39%, SD = 35.64%) than non-VME apparent images (M = 7.04%, SD = 9.67%), regardless of whether the drawings were from short-term or long-term memory, implying that people were more likely to make VME errors in their drawings for these images.

The occurrence of VME errors during both short-term and long-term recall suggests that people do spontaneously generate these specific errors; VME is not a recognition-only phenomenon. Given the variability of the error frequency, it also suggests that the ease of spontaneously generating these errors may depend on the specific VME-apparent image. Furthermore, the fact that VME errors can occur during short-term recall, despite limited familiarity with the image, could suggest that there is something intrinsic to these stimuli that encourages these errors.

Discussion

Here, we characterized the VME, empirically demonstrating that seven familiar images from popular iconography have low memory accuracy, with a specific incorrect version consistently remembered across people. These VME errors occur during recognition and recall and when using short-term or long-term memory. Furthermore, there might not be a universal explanation
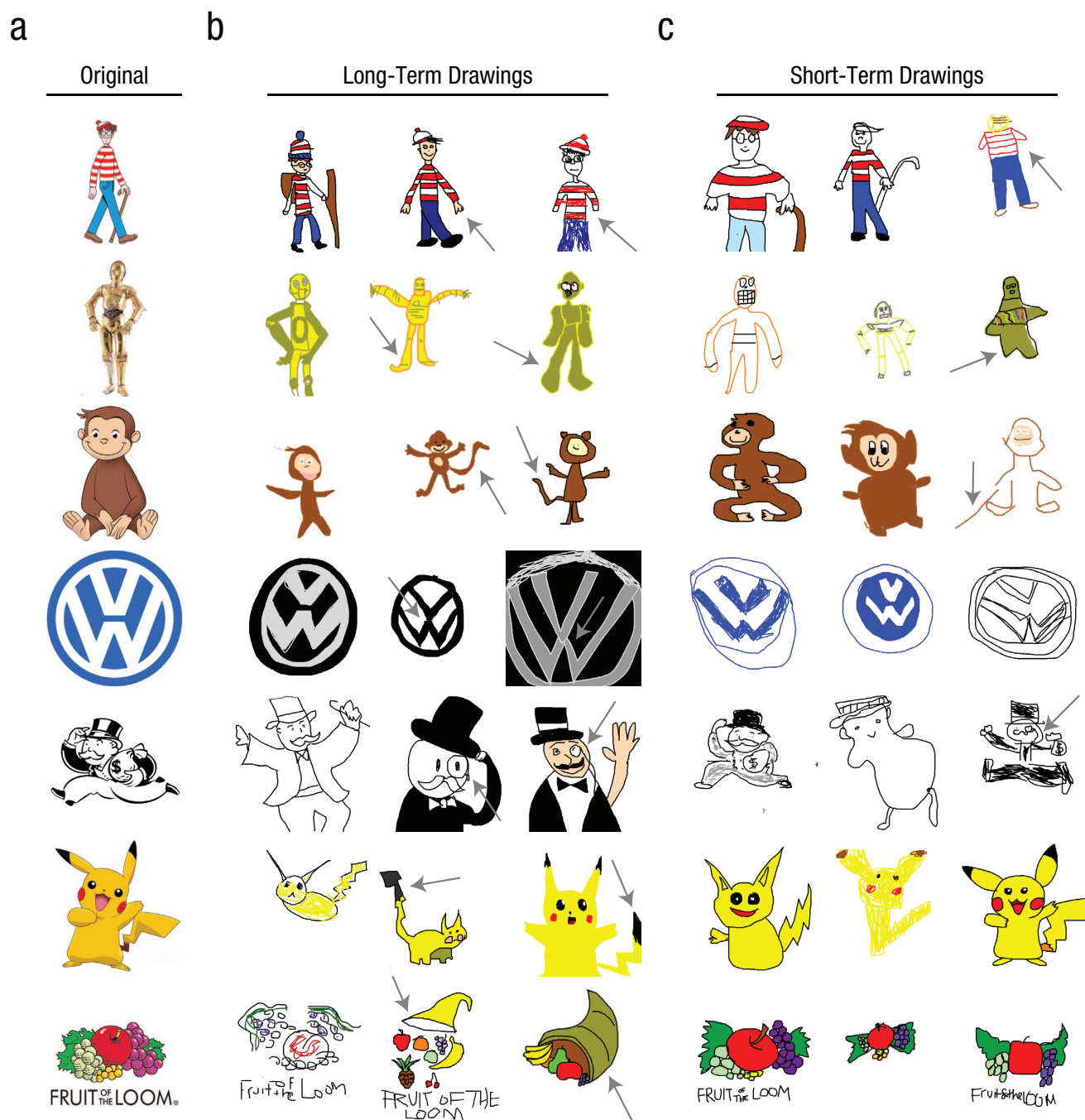
Fig. 9. Drawings from both short-term and long-term recall include spontaneous visual Mandela effect (VME) errors. The original version of the image concept (a) was seen only by participants of the short-term recall task who had a low familiarity with it. The gray arrows for the long-term drawings (b) and short-term drawings (c) point to the respective VME error on the drawing. Participants spontaneously generated the VME error in a total of 47.72% of long-term drawings and 20.97% of short-term drawings, despite not having seen that version of the image. These results suggest that the VME is not a recognition-only error and can be generated during recall as well.

for why the VME occurs. While we showed that these errors are unlikely to be explained by attentional or low-level visual differences, the VME may be driven by schema-based perceptual knowledge for some icons and driven by visual experience with the noncanonical version for others.

The untested explanation for the VME relied solely on the schema theory of false memory: People are more likely to misremember details when they align with expectations of the image. Although this explanation is limited because it fails to fully explain the consistency and specificity of the VME, it may play a role in driving

VME errors resulting from an incomplete perceptual experience. In Experiment 3, we examined how these icons are represented in the real world and found that the frequency of the specific feature may be low for some images (e.g., C-3PO's legs are rarely shown). Additionally, whereas our results from Experiment 2 showed no significant differences in inspection patterns between VME and non-VME images, the analysis might not fully capture how perceptible these features are under naturalistic viewing conditions. Therefore, it is possible that for some VME images (e.g., C-3PO, Volkswagen, Where's Waldo), the error may be driven by filling in these perceptual gaps with schematic knowledge, creating a more schema-typical false memory (Koutstaal & Schacter, 1997; Miller & Gazzaniga, 1998; Seamon et al., 2000). However, this is unlikely to apply to all VME-apparent images. For some icons, the feature of interest is almost always shown (e.g., the Monopoly Man's face), and some VME features are intuitively atypical to the schema (e.g., a cornucopia for the Fruit of the Loom logo). Furthermore, it is unclear how such schemas are formed for new image classes; for example, as Pokémon's most popular character, Pikachu is likely the basis for that schema, yet people falsely remember a black tip on its tail. Future research into VME should examine the schema consistency of VME features compared with other manipulations.

Another possible account for the VME could be source confusion of the memory. Our image scraping results suggest that, although uncommon, it is possible for people to have previously experienced the VME version (e.g., C-3PO, Fruit of the Loom). In this case, the VME error would not be a false memory but rather source confusion: remembering the noncanonical version they saw instead of the canonical design. As the Mandela effect becomes more known in popular culture, these noncanonical memories may become amplified, which could contribute to this account. Indeed, 67.92% of participants had heard of the effect, and some scraped images with a VME were from sites discussing the VME (e.g., Fruit of the Loom in Fig. 7). It will be important to characterize this phenomenon while it is still relatively new.

Although these two accounts may explain some examples of the VME, there are results that neither can fully account for. First, participants who made VME errors reported high familiarity, implying that they had extensive experience with the canonical version that may not be reflected by the image scraping analysis (e.g., watching the Pokémon TV show). Second, we found that for most of the VME icons (five of seven), the majority of the scraped images were the canonical version. This suggests that the VME versions are rare

and may not be experienced by everyone. Third, a subset of the VME-apparent images from Experiment 3 originated from websites and articles about the VME (Fig. 7), implying that the widespread memory error preceded reporting of the phenomenon. Additionally, Experiments 2 and 4 demonstrated that the VME can occur even on short-term memory scales: Participants made VME errors, despite perceiving the canonical images a few moments prior. It is unlikely that this occurrence of the VME is due to a rare variant from prior experience overriding the recent experience of the more frequent canonical image. This converging evidence suggests that some people may be making consistent memory errors, even with extensive visual experience with the icon and without having experienced variants before this.

Another potential account that could partially explain the VME is that these errors may occur because of intrinsic properties of the VME image. Many studies have shown consistency in what we remember (Bainbridge, 2019; Bainbridge et al., 2013, 2017) due to an intrinsic pull of the stimulus (its memorability). The memorability of an image is not driven by low-level features (Isola et al., 2014), bottom-up or top-down attention (Bainbridge, 2020), or semantic processing of the images (Lin et al., 2021) but can be predicted with models incorporating both visual and semantic information, to moderate success (Kramer et al., 2022; Needell & Bainbridge, 2022). These studies have also observed consistency across observers in false recognition for face stimuli (Bainbridge, 2017; Bainbridge et al., 2013), in which some novel images are consistently reported as old, across people. In the instance of source confusion, it could be that the VME versions are more memorable than the canonical versions, which is why they are remembered instead. However, there remains the same question of what features drive these memories. Studies examining consistency in true memory used hundreds or thousands of stimuli (Bainbridge et al., 2013, 2017; Isola et al., 2014), but here, we examined only 40 images from popular iconography. Future studies should make use of many more images and leverage methods designed to induce false memories for a stimulus (Koutstaal & Schacter, 1997; Miller & Gazzaniga, 1998; Roediger & McDermott, 1995) to induce consistent false memories across people.

Although these images cause consistent errors across individuals, our study tested only icons and adults on Amazon Mechanical Turk and Prolific within the United States. Because these images are culturally specific, we might expect differences outside of the United States. Indeed, prior work on word false memory shows that effects can be culturally specific (Roediger et al., 2001;

Stadler et al., 1999). However, our results do show that individuals will spontaneously create false memories even for icons with which they are unfamiliar (Experiment 4). Future research will need to investigate reports of VME-like memories in other cultures and whether one can create culturally agnostic VME stimuli. Given that we demonstrated the VME in long- and short-term memory tasks as well as in both recognition and recall tasks, we expect that this effect is robust and could be replicated with different task paradigms. Given that the effect was originally identified with these icons in popular culture, we also think these results would generalize beyond an online sample. One key distinction to make is that the false memories induced by the VME should not be considered the same as rich episodic autobiographical memories, and therefore our results regarding the prevalence of VME-induced false memories should not be generalized to episodic memories in the same way. What these results do suggest, though, is that part of what drives false memories may be dependent on the stimulus, and thus peoples' propensities for false memories could be better predicted with this in mind.

In sum, we revealed a set of images that cause consistent and shared false memories across people, spurring new questions on the nature of false memories. We showed that the VME cannot be universally explained by a single account. Instead, perhaps different images cause a VME for different reasons—some related to schema, some related to visual experience, and some related to something entirely different about the images themselves.

Transparency

Action Editor: Marc Buehner
Editor: Patricia J. Bauer
Author Contributions
 D. Prasad and W. A. Bainbridge developed the study concept and designed the study. D. Prasad programmed the experimental tasks, collected the data, and conducted the analyses. Both authors contributed to the final version of the manuscript. W. A. Bainbridge supervised the project. Both authors approved the final manuscript for submission.
Declaration of Conflicting Interests
 The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Open Practices
 All data and stimuli have been made publicly available via OSF and can be accessed at https://osf.io/7cmwf/. The design and analysis plans for the experiments were not preregistered. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologi

Acknowledgments

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976221108944

References

Anwyl-Irvine, A. L., Armstrong, T., & Dalmaijer, E. S. (2021). MouseView.js: Reliable and valid attention tracking in web-based experiments using a cursor-directed aperture. *Behavior Research Methods*. Advance online publication. https://doi.org/10.3758/s13428-021-01703-5

Bainbridge, W. A. (2017). The memorability of people: Intrinsic memorability across transformations of a person's face. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 706–716. https://doi.org/10.1037/xlm0000339

Bainbridge, W. A. (2019). Memorability: How what we see influences what we remember. In K. D. Federmeier & D. M. Beck (Eds.), *Psychology of learning and motivation* (Vol. 70, pp. 1–27). Elsevier. https://doi.org/10.1016/bs.plm.2019.02.001

Bainbridge, W. A. (2020). The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia*, *141*, Article 107408. https://doi.org/10.1016/j.neuropsychologia.2020.107408

Bainbridge, W. A., Dilks, D. D., & Oliva, A. (2017). Memorability: A stimulus-driven perceptual neural signature distinctive from memory. *NeuroImage*, *149*, 141–152. https://doi.org/10.1016/j.neuroimage.2017.01.063

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications*, *10*(1), Article 5. https://doi.org/10.1038/s41467-018-07830-6

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

Bainbridge, W. A., Kwok, W. Y., & Baker, C. I. (2021). Disrupted object-scene semantics boost scene recall but diminish object recall in drawings from memory. *Memory & Cognition*, *49*(8), 1568–1582. https://doi.org/10.3758/s13421-021-01180-3

Bakkila, B. (2019, August 6). *What is the Mandela effect?* Good Housekeeping. https://www.goodhousekeeping.com/life/entertainment/g28438966/mandela-effect-examples/

Blake, A. B., Nazarian, M., & Castel, A. D. (2015). Rapid communication: The apple of the mind's eye: Everyday attention, metamemory, and reconstructive memory for the Apple logo. *Quarterly Journal of Experimental Psychology*, *68*(5), 858–865. https://doi.org/10.1080/17470218.2014.1002798

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Broome, F. (2010, September 9). *Nelson Mandela died in prison?* Mandela Effect. https://mandelaeffect.com/nelson-mandela-died-in-prison/

Clinton, J. (2021). *Google Images download* [Python script]. GitHub. https://github.com/Joeclinton1/google-images-download

Dagnall, N., & Drinkwater, K. (2018, February 14). *The Mandela effect: Explaining the science behind false memories.* The Independent. https://www.independent.co.uk/news/science/mandela-effect-false-memories-explain-science-time-travel-parallel-universe-matrix-a8206746.html

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

The GIMP Team. (2020). *GNU image manipulation program (GIMP)* (Version 2.10.18) [Computer software]. https://www.gimp.org/

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16*(4), 219–222. https://doi.org/10.1111/j.1467-8721.2007.00507.x

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, *1*(10), 743–747. https://doi.org/10.1038/s41562-017-0208-0

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113–136. https://doi.org/10.1037/0096-1523.28.1.113

Isola, P., Xiao, J., Parikh, D., Torralba, A., & Oliva, A. (2014). What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(7), 1469–1482. https://doi.org/10.1109/TPAMI.2013.200

Jones, G. V. (1990). Misremembering a common object: When left is not right. *Memory & Cognition*, *18*(2), 174–182. https://doi.org/10.3758/BF03197093

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*(4), 555–583. https://doi.org/10.1006/jmla.1997.2529

Kramer, M. A., Hebart, M. N., Baker, C. I., & Bainbridge, W. A. (2022). The features underlying the memorability of Objects. bioRxiv. https://doi.org/10.1101/2022.04.29.490104

Lin, Q., Yousif, S. R., Chun, M. M., & Scholl, B. J. (2021). Visual memorability in the absence of semantic content. *Cognition*, 212, Article 104714. https://doi.org/10.1016/j.cognition.2021.104714

Maswood, R., & Rajaram, S. (2019). Social transmission of false memory in small groups and large networks. *Topics in Cognitive Science*, *11*(4), 687–709. https://doi.org/10.1111/tops.12348

Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36*(6), 513–520. https://doi.org/10.1016/S0028-3932(97)00148-6

Needell, C. D., & Bainbridge, W. A. (2022). Embracing new techniques in deep learning for estimating image memorability. *Computational Brain & Behavior*, *5*(2), 168–184. https://doi.org/10.1007/s42113-022-00126-5

Nickerson, R. S. (1965). Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *19*(2), 155–160. https://doi.org/10.1037/h0082899

Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, *11*(3), 287–307. https://doi.org/10.1016/0010-0285(79)90013-6

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*(3), 385–407. https://doi.org/10.3758/BF03196177

Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language*, *42*(1), 120–146. https://doi.org/10.1006/jmla.1999.2676

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*(3), 494–500. https://doi.org/10.3758/BF03211543

Wong, K., Wadee, F., Ellenblum, G., & McCloskey, M. (2018). The devil's in the g-tails: Deficient letter-shape knowledge and awareness despite massive visual experience. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1324–1335. https://doi.org/10.1037/xhp0000532