**Supplemental Material:**

**The Visual Mandela Effect as evidence for shared and specific false memories across people**

**BubbleView Experiment**

We conducted a supplementary experiment to compare fixation behavior between different versions of these image concepts; do VME manipulations capture more fixations than their original counterparts?

**Methods**

*Stimuli*

Stimuli consisted of the four image subsets tested in Experiment 4. Thirty-two test sets were presented to the participants, but only twenty-eight were utilized in the final analysis, due to the Pringles logo changing during the course of this study.

*Participants*

Sixty-one participants were recruited for the study (females=26, M age=38.8 years, SD age=11.0). Participants were excluded if more than six responses had a response time less than or equal to 100 ms or if they failed a task summary attentional check.

*Procedure*

We used BubbleView, a mouse-tracking method analogous to eye-tracking (Kim et al., 2017, de Leeuw, 2015; Choe, 2019/2021) to present a blurred image for participants to click on with their mouse and unblur small sections (‘bubbles’) of the image at a time, imitating foveation. BubbleView images were presented at a size of 500 x 500 pixels with a 10 sigma Gaussian blur, with a 30-pixel bubble radius. Participants saw one of the three possible versions from each image set.

First, participants were presented with a BubbleView image and were instructed to click around on the image to determine whether it was real (original) or fake (manipulated). They had a minimum of five clicks and a maximum of 100 clicks, with no time limit. They then indicated whether the image was real or fake. Next, they were presented with a BubbleView of the same image and were asked to click on the area that was most informative for their decision. Only one click was allowed. Finally, participants rated their confidence in their choice and their familiarity with the image concept on a 1 to 5 Likert scale.

Participants were informed that the fake images were very similar to the real version and were asked to pay attention to small details. They had a practice run with three trials with different images. After making their judgement, participants would see a side-by-side comparison of the presented image and the real image to get a sense of how small the changes between real and fake images could be. After completing the practice run, participants would move on to the main run.

*BubbleView Click Density Map Generation*

Inspection density maps were constructed using the methods in Experiment 2.

*Results*

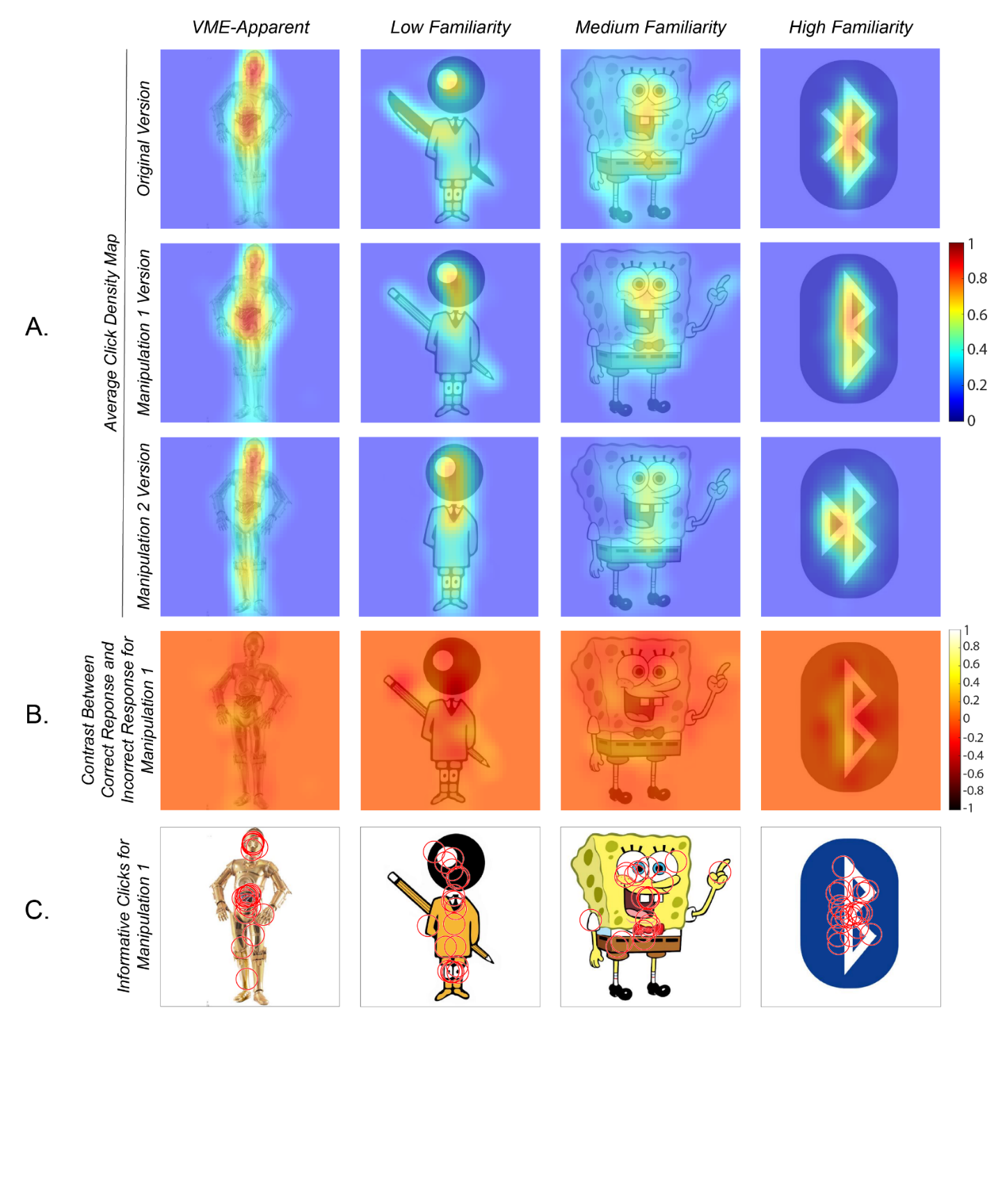
The average accuracy was 0.64 (SD = 0.09) for the high familiarity subset, 0.61 (SD = 0.05) for the medium familiarity subset, 0.55 (SD = 0.05) for the low familiarity subset, and 0.54 (SD = 0.09) for the VME-apparent images. VME-apparent images tended to have lower accuracy than their matched control, although the difference did not reach significance (independent t-test, t(12) = 1.78, p = 0.100; BF10 = 1.17, anecdotal evidence for the alternate hypothesis). There was no significant difference in familiarity between VME-apparent images and the medium familiarity matched group (independent t-test, t(12) = 1.18, p = 0.262; BF01 = 1.43, anecdotal evidence for the null hypothesis).

We examined the effects of three factors on the average click density value and the maximum click density value of Manipulation 1 images: image category (VME/low/medium/high), image area (unaltered/manipulated), and response type (correct/incorrect). A 3-way ANOVA showed a significant main effect of image area on the average click density value (F(1,96) = 105.40, p = 3.99×10-17, ηp2 = 0.50; BF10 = 1.53×1016, extreme evidence for the alternate hypothesis), with the manipulated area viewed significantly more than the unaltered area (post-hoc, t(110) =10.46, p = 3.37×10-18). There was no main effect of image subset or response type, and no significant interactions between factors (all p ≥ 0.217; all BF01 ≥ 3.21, moderate to strong evidence for the null hypothesis). A 3-way ANOVA on maximum click density showed the same statistical patterns.

For VME-apparent images only, we examined the effects of manipulation type (Manipulation 1/Manipulation 2), image area (manipulated/unaltered), and response type (correct/incorrect) on the average click value with a 3-way ANOVA. Image area had a significant main effect (F(1,48) = 49.27, p = 6.87×10-9, ηp2 = 0.49; BF10 = 1.19×107, extreme evidence for the alternate hypothesis). There was no main effect of manipulation type or response type and no significant interactions between the factors (all p > 0.191; all BF01 ≥ 1.77, anecdotal to moderate evidence for the null hypothesis). An ANOVA on maximum click density showed the same statistical patterns.

We also examined what area of images participants chose to be most the informative for their decisions (Figure SR1C). There was no significant difference in click behavior between VME-apparent images and the familiarity-matched medium subset (independent t-test, t(12) = 0.82, p = 0.430; BF01 = 1.80, anecdotal evidence for the null hypothesis). Additionally, there was no significant difference in informative click percentage between Manipulation 1 of the VME-apparent images and Manipulation 2 (independent t-test, t(12) = 0.08, p = 0.940; BF01 = 2.24, anecdotal evidence for the null hypothesis).

In sum, these results suggest no differences in inspection behavior between VME and non-VME images.



**Fig. S1. Attentional differences do not drive VME.** On the average click density maps (a), the higher values indicated in red represent areas of the image where participants clicked more. There were no interaction effects between image subset, response type, and image area on the average and maximum click density of the area that explained the false memory effects of VME (all p ≥ 0.121). (b) For these contrast images, brighter yellow regions indicate areas on the image viewed more by participants who correctly responded ‘fake’. Darker red regions on the images indicate areas viewed more by those who incorrectly responded ‘real’. There were no interaction effects between image subset and image area on the average and maximum click density of these contrasts that explained the false memory effects of VME (all p ≥ 0.38) (c) Each red circle is an informative click from a participant. There was no significant difference in the frequency of informative clicks in the manipulated area across the image subsets (t(12) = 0.82, p = 0.430).

**Supplemental References**

Choe, K. W. (2021). *Kywch/BubbleView\_jsPsych* [JavaScript]. https://github.com/kywch/BubbleView\_jsPsych (Original work published 2019)

Kim, N. W., Bylinskii, Z., Borkin, M. A., Gajos, K. Z., Oliva, A., Durand, F., & Pfister, H. (2017). BubbleView: An Interface for Crowdsourcing Image Importance Maps and Tracking Visual Attention. *ACM Transactions on Computer-Human Interaction*, *24*(5), 1–40. https://doi.org/10.1145/3131275

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Average | | | | |  | Max | | | | |  |
|  |  | Sum Sq | df | Mean Sq | F | p | BF01 | Sum Sq | df | Mean Sq | F | p | BF01 |
| C3PO | Response | 0.006 | 1 | 0.006 | 1.580 | 0.211 | 3.29 | 0.017 | 1 | 0.017 | 1.181 | 0.279 | 4.12 |
| Area | 0.157 | 1 | 0.157 | 40.239 | 4.49E-09 | 2.57E-07 | 0.119 | 1 | 0.119 | 8.467 | 0.004 | 0.15 |
| Response\*Area | 0.002 | 1 | 0.002 | 0.573 | 0.451 | 4.15 | 0.029 | 1 | 0.029 | 2.035 | 0.156 | 2.22 |
| Errors | 0.452 | 116 | 0.004 |  |  |  | 1.629 | 116 | 0.014 |  |  |  |
| Total | 0.653 | 119 |  |  |  |  | 1.765 | 119 |  |  |  |  |
| Fruit of the Loom | Response | 8.19E-04 | 1 | 8.19E-04 | 0.170 | 0.681 | 7.15 | 1.18E-04 | 1 | 1.18E-04 | 0.022 | 0.882 | 7.50 |
| Area | 0.023 | 1 | 0.023 | 4.825 | 0.030 | 0.77 | 0.049 | 1 | 0.049 | 9.051 | 0.003 | 0.10 |
| Response\*Area | 0.000 | 1 | 0.000 | 0.000 | 0.999 | 5.71 | 0.000 | 1 | 0.000 | 9.95E-05 | 0.992 | 5.50 |
| Errors | 0.559 | 116 | 0.005 |  |  |  | 0.623 | 116 | 0.005 |  |  |  |
| Total | 0.587 | 119 |  |  |  |  | 0.680 | 119 |  |  |  |  |
| Curious George | Response | 3.35E-04 | 1 | 3.35E-04 | 0.061 | 0.806 | 8.02 | 0.036 | 1 | 0.036 | 1.642 | 0.203 | 3.82 |
| Area | 0.072 | 1 | 0.072 | 13.043 | 4.51E-04 | 1.90E-02 | 0.115 | 1 | 0.115 | 5.308 | 0.023 | 0.64 |
| Response\*Area | 6.50E-05 | 1 | 6.50E-05 | 0.012 | 0.914 | 5.91 | 0.010 | 1 | 0.010 | 0.461 | 0.499 | 4.76 |
| Errors | 0.640 | 116 | 0.006 |  |  |  | 2.521 | 116 | 0.022 |  |  |  |
| Total | 0.712 | 119 |  |  |  |  | 2.680 | 119 |  |  |  |  |
| Monopoly | Response | 0.023 | 1 | 0.023 | 4.096 | 0.045 | 1.16 | 0.004 | 1 | 0.004 | 0.319 | 0.573 | 6.85 |
| Area | 0.122 | 1 | 0.122 | 21.258 | 1.04E-05 | 4.88E-04 | 0.109 | 1 | 0.109 | 8.927 | 0.003 | 0.12 |
| Response\*Area | 0.010 | 1 | 0.010 | 1.670 | 0.199 | 2.59 | 0.007 | 1 | 0.007 | 0.544 | 0.462 | 4.63 |
| Errors | 0.665 | 116 | 0.006 |  |  |  | 1.413 | 116 | 0.012 |  |  |  |
| Total | 0.849 | 119 |  |  |  |  | 1.526 | 119 |  |  |  |  |
| Pikachu | Response | 0.002 | 1 | 0.002 | 0.434 | 0.511 | 6.81 | 0.026 | 1 | 0.026 | 0.863 | 0.355 | 5.43 |
| Area | 0.001 | 1 | 0.001 | 0.371 | 0.544 | 7.02 | 12.436 | 1 | 12.436 | 420.803 | 2.14E-40 | 2.36E-38 |
| Response\*Area | 1.32E-04 | 1 | 1.32E-04 | 0.033 | 0.857 | 6.04 | 0.026 | 1 | 0.026 | 0.863 | 0.355 | 3.97 |
| Errors | 0.466 | 116 | 0.004 |  |  |  | 3.428 | 116 | 0.030 |  |  |  |
| Total | 0.469 | 119 |  |  |  |  | 15.966 | 119 |  |  |  |  |
| Volkswagen | Response | 0.021 | 1 | 0.021 | 3.138 | 0.079 | 1.76 | 0.003 | 1 | 0.003 | 0.274 | 0.602 | 7.16 |
| Area | 1.55E-04 | 1 | 1.55E-04 | 0.023 | 0.880 | 7.90 | 0.009 | 1 | 0.009 | 0.801 | 0.373 | 5.45 |
| Response\*Area | 0.000 | 1 | 0.000 | 1.49E-04 | 0.990 | 5.92 | 0.001 | 1 | 0.001 | 0.107 | 0.744 | 5.53 |
| Errors | 0.781 | 116 | 0.007 |  |  |  | 1.367 | 116 | 0.012 |  |  |  |
| Total | 0.802 | 119 |  |  |  |  | 1.379 | 119 |  |  |  |  |
| Waldo | Response | 0.009 | 1 | 0.009 | 1.309 | 0.255 | 4.34 | 8.10E-04 | 1 | 8.10E-04 | 0.019 | 0.892 | 8.28 |
| Area | 0.120 | 1 | 0.120 | 17.769 | 4.96E-05 | 2.40E-03 | 7.008 | 1 | 7.008 | 160.888 | 1.17E-23 | 1.29E-21 |
| Response\*Area | 0.002 | 1 | 0.002 | 0.312 | 0.577 | 5.13 | 0.000 | 1 | 0.000 | 9.35E-05 | 0.992 | 6.07 |
| Errors | 0.786 | 116 | 0.007 |  |  |  | 5.053 | 116 | 0.044 |  |  |  |
| Total | 0.918 | 119 |  |  |  |  | 12.070 | 119 |  |  |  |  |

**Table. S1. MouseView 2-way ANOVA results for VME images.** For each VME image, we tested the effects of area of the image (manipulated area / unmanipulated area) and response type (correct / incorrect) on inspection density, as measured both by average inspection and maximum inspection density. No images showed a significant interaction of inspection area by response type, implying that no images showed a difference in inspection behavior between when participants had a correct or false memory.

**Calendar

Description automatically generated**

**Fig. S2. MouseView subtraction heat maps.** The heatmaps showthe difference in inspection behavior between those who correctly versus incorrectly remembered all 14 images tested in Experiment 2. Crucially, there were no significant differences in inspection behavior related to the type of image that was being viewed (VME or matched), or whether they memory was correct or incorrect. Dark blue indicates areas viewed more by participants who answered incorrectly. Red indicates areas viewed more by participants who answered correctly.



**Fig. S3. Image scraping examples.** Results from image scraping show high variation in the natural viewing experience of these images. For some images, like C3PO, a substantial portion of scraped images contained the VME. For other images, like Pikachu, very few scraped images contained the VME. VME-features, when present, are circled in green.