Multi-Image Advantage in Face Identity Matching Depends on Face Representation Formation

Authors: Feng Junye, Wang Zhe, Sun Yuhao, Sun Yuhao

Date: 2025-06-22T00:00:00+00:00

Abstract

Presenting multiple face images of the same person can significantly improve participants' performance in recognizing face identity. However, the cognitive mechanisms underlying the improvement in face recognition performance through multiple images remain unclear. This study comprises two experiments. Experiment 1A employed a face matching paradigm, presenting one, two, or three faces either simultaneously or sequentially, to measure participants' discriminability under different conditions. The results revealed: (1) Only under sequential presentation conditions did participants' discriminability improve with increasing number of images (demonstrating a multiple-image advantage); (2) When three face images were presented, participants' discriminability under sequential presentation conditions was higher than that under simultaneous matching conditions. Experiment 1B controlled face presentation time and replicated the above results. Experiment 2, building upon Experiment 1A, inverted faces to disrupt the integration process of face representation. The results showed that, (3) Regardless of whether the study images were single or multiple, participants' discriminability under sequential presentation conditions was lower than that under simultaneous presentation conditions, (4) No multiple-image advantage was found in either task. In summary, the experimental results suggest that the multiple-image advantage in face identity discrimination originates from the formation of face representations and that this process requires the involvement of memory.

Full Text

The Multiple-Image Advantage in Face Identity Matching Relies on Face Representation Formation

FENG Junye, WANG Zhe, SUN Yuhao P.

Department of Psychology, Zhejiang Sci-Tech University, Hangzhou 310018, China

Presenting multiple face images of the same individual significantly improves participants' performance in face identity recognition. However, the cognitive mechanisms underlying this improvement remain unclear. This study comprised two experiments. Experiment 1A employed a face-matching paradigm in which one, two, or three face images were presented either simultaneously or sequentially, measuring participants' discriminability across different conditions. The results revealed: (1) discriminability increased with the number of images only in the sequential presentation condition (demonstrating a multiple-image advantage); (2) when three face images were presented, discriminability in the sequential condition exceeded that in the simultaneous matching condition. Experiment 1B replicated these results while controlling for face presentation duration. Experiment 2 built upon Experiment 1A by inverting all faces to disrupt the integration process of face representation. The findings showed that (3) regardless of whether study images were single or multiple, discriminability in the sequential condition was lower than in the simultaneous condition, and (4) no multiple-image advantage emerged in either task. Taken together, these results suggest that the multiple-image advantage in face identity discrimination stems from face representation formation, a process that requires memory involvement.

Keywords: multiple-image advantage, face representation, face matching, face recognition
Classification Code: B842

Recognizing unfamiliar faces is notoriously difficult (Young & Burton, 2018). However, researchers have found that presenting multiple face images of the same identity can enhance face recognition performance (Andrews et al., 2015; Baker & Mondloch, 2019; Baker & Mondloch, 2023; Bindemann & Sandford, 2011; Dowsett et al., 2016; Longmore et al., 2017; Menon et al., 2015a; Murphy et al., 2015; Sandford & Ritchie, 2021), a phenomenon termed the multiple-image advantage.

Received: 2024-12-07
Funded by the Zhejiang Provincial Natural Science Foundation (LY19C090006, LY20C090010)
Corresponding author: SUN Yuhao, E-mail: sunyuhao@zstu.edu.cn

Researchers attribute the multiple-image advantage to the formation of face representation—that is, through exposure to multiple images of the same identity, individuals develop a representation of that identity, with more images yielding a more robust representation and consequently better recognition performance (Devue & de Sena, 2023; Ritchie et al., 2021). An alternative explanation posits that the advantage arises from the quantity of image information: more images provide more information, thereby improving recognition performance (Kramer et al., 2020; Menon et al., 2018; White et al., 2014). Thus, whether the multiple-image advantage truly originates from face identity representation formation remains questionable.

The multiple-image advantage phenomenon is not consistently observed. The presentation mode of study and target faces may be a critical factor influencing this effect (Sandford & Ritchie, 2021). Studies reporting a multiple-image advantage typically employ sequential presentation, where multiple images are shown first, followed by the target image (Andrews et al., 2015; Baker & Mondloch, 2019; Baker & Mondloch, 2023; Longmore et al., 2017; Menon et al., 2015a; Murphy et al., 2015; Sandford & Ritchie, 2021). For example, Baker and Mondloch (2023) used a new/old face task paradigm and found that discriminability for target faces improved as the number of images presented during the learning phase increased. Studies failing to find a multiple-image advantage typically use simultaneous presentation of multiple images and the target image (Kramer & Reynolds, 2018; Ritchie et al., 2020; Ritchie et al., 2021). For instance, Ritchie et al. (2020) employed a live face-matching task and found that presenting multiple face images did not improve accuracy compared to presenting a single image.

Diverging from these two lines of research, Sandford and Ritchie (2021) examined both tasks within a single study. In the simultaneous face-matching task, the study face image set (containing one, two, or three images of the same person) and a target face image were presented concurrently. In the sequential face-matching task, the study face image set was presented for five seconds, followed by the target face image alone. Participants judged whether the study set and target face depicted the same person. Results showed that in the simultaneous task, the number of images in the study set did not affect discriminability. In contrast, discriminability in the sequential task gradually improved as the number of study images increased, demonstrating a multiple-image advantage.

Some researchers propose that simultaneous versus sequential presentation of study and target faces influences the degree to which memory participates in face processing (Menon et al., 2015a; Sandford & Ritchie, 2021; Ritchie et al., 2021; Baker et al., 2023). It is widely believed that when study and target faces are presented simultaneously, participants perform a perceptual task because they can scan back and forth between faces (Megreya et al., 2011; Menon et al., 2015b; Davis et al., 2021; Sandford & Ritchie, 2021; Ritchie et al., 2021). In sequential presentation, participants must rely on memory to retain the study faces and compare them to the target (Honig et al., 2022; Dowsett et al., 2016; Matthews & Mondloch, 2022; Pitcher et al., 2023; Menon et al., 2015b; Mileva et al., 2021; Matthews et al., 2024). This speculation suggests that face representation formation requires memory involvement.

However, these results still fail to rule out the role of image information quantity. For example, research has found that images with greater variability are more likely to produce a multiple-image advantage than those with less variability (Baker et al., 2017; Menon et al., 2015a; Ritchie & Burton, 2017; Sandford & Ritchie, 2021), indicating that more information yields better matching performance. This raises the possibility that an interaction between image information quantity and task difficulty produces performance differences between sequential and simultaneous matching tasks. Specifically, because sequential matching tasks require participants to remember multiple images, they are more difficult, resulting in relatively lower performance that makes differences between single and multiple image conditions more apparent. Conversely, because simultaneous matching tasks are less difficult, performance for both single and multiple image conditions is high, with small differences approaching ceiling effects (see Kramer & Reynolds, 2018), making it difficult to detect differences between conditions. Additionally, it is possible that in simultaneous matching tasks, increased information load under multiple-image conditions creates higher cognitive load, eliminating performance differences between single and multiple images (see Mileva & Burton, 2019). In short, until the role of image information quantity is completely excluded, previous experimental results cannot fully support the hypothesis that the multiple-image advantage stems from face representation formation.

In brief, the two hypotheses– "face representation formation" and "increased image information quantity" –make different predictions about performance differences between simultaneous and sequential matching tasks under multiple-image conditions. The former predicts that multiple images in sequential conditions will facilitate the formation of a memorial representation of face identity, which helps participants perform better at face matching, thus sequential performance will exceed simultaneous performance. Conversely, the latter predicts that sequential matching performance will be lower than simultaneous performance. To test the first hypothesis and exclude the second, we conducted two experiments. Experiment 1, similar to Sandford and Ritchie (2021), required participants to judge whether a set of study face images (containing one, two, or three images of the same person) matched a target face. Two tasks were employed: a simultaneous matching task where study and target faces were presented together, and a sequential matching task where study faces were presented first, followed by the target face. Experiment 2 repeated the two tasks from Experiment 1 but with all faces inverted.

The theoretical rationale for using sequential matching (memory-demanding) and simultaneous matching (perceptual matching) tasks to distinguish representation formation is supported by two lines of evidence. First, in previous research investigating the mechanisms of the multiple-image advantage, sequential matching tasks (e.g., study-test paradigms) have been widely used as classic paradigms for studying representation formation. Studies by Jones et al. (2017), Menon et al. (2015a, 2018), and Baker et al. (2023) all found that sequentially presenting multiple images significantly improved identity discrimination performance, an advantage generally attributed to stable representation formation. These studies support the use of sequential tasks as an experimental manipulation for representation formation. Second, in studies directly comparing these two task paradigms, this manipulation has been adopted because it establishes representations in memory (see Sandford & Ritchie, 2021; Ritchie et al., 2021;

Baker et al., 2023).

Based on the aforementioned hypotheses, four predictions can be made regarding the experimental results. (1) If the multiple-image advantage arises from memory facilitating face representation formation, Experiment 1 should find that discriminability improves with increasing image numbers only in the sequential face-matching task (i.e., a multiple-image advantage emerges). (2) More importantly, if the multiple-image advantage results from the formation of more robust face identity representations through memory, Experiment 1 should also find that when multiple face images are presented, discriminability in the sequential task will be higher than in the simultaneous task. (3) Conversely, if task difficulty in sequential presentation is the key factor, both Experiment 1 and Experiment 2 should find that when multiple face images are presented, discriminability in the sequential task will be lower than in the simultaneous task due to higher difficulty. (4) Given that face inversion disrupts face representation formation (Leder & Bruce, 2000; Xu & Tanaka, 2013) and substantially increases the difficulty of face recognition (Freire et al., 2000; Yin, 1969), if the multiple-image advantage in sequential tasks stems from memory-based face representation formation, then Experiment 2 should find no multiple-image advantage in either task when faces are inverted because representation formation is disrupted. If the multiple-image advantage arises from an interaction between image information quantity and task difficulty, then participants should show a multiple-image advantage in the simultaneous task when faces are inverted.

2.1.1 Participants

Eighty-one university students (30 males) with a mean age of 19.7 years (range 17–24, SD = 1.7) participated. All had normal or corrected-to-normal vision and were right-handed. Based on Sandford and Ritchie' s (2021) results, which found a significant interaction between presentation mode (between-subjects) and image number (within-subjects) with a PES of 0.093 (converted effect size f 0.32), we used G*Power to calculate the minimum sample size for a repeated-measures ANOVA with power = 0.8 and effect size f = 0.32, which yielded N = 52. Considering that Sandford and Ritchie (2021) used 79 participants, we determined our sample size to be 81. The experiment was conducted under the guidance of the Ethics Committee of the School of Science at Zhejiang Sci-Tech University (Ethics No.: 202309P002). Participants provided informed consent before the experiment and received a small cash payment afterward.

2.1.2 Stimuli and Apparatus

We used 144 photographs of 36 Hong Kong celebrities (half male, half female) with low mainland China recognition as experimental materials (four photos per person). Each celebrity was also matched with a foil image from another identity with similar appearance description, totaling 180 images. All materials were sourced from photos posted on social media platforms such as Weibo (with posting times across different photos controlled within three years). Photos varied in facial expression, head angle, and environmental conditions (e.g., lighting, camera characteristics). The 36 celebrities were divided into two sets, with each set randomly assigned to either the simultaneous or sequential task. Materials used in the simultaneous task did not appear in the sequential task, and material assignment across tasks was counterbalanced between participants.

All face photos were 200$\times 280 pixelsandpresentedonablackbackgroundof 800\times$710 pixels. During the experiment, the target image always appeared at the top center, while study images appeared at the bottom. In 1-to-1 matching, the study image appeared at the bottom left; in 1-to-2 matching, comparison images appeared at the bottom left and center; in 1-to-3 matching, comparison images appeared at the left, center, and right positions. The gap between the left and right contours of study images was 100 pixels, and the gap between the target and the bottom-center study image was 150 pixels (see Figure 1 [Figure 1: see original paper]).

Stimuli were presented on a 15.4-inch LCD monitor with a 60Hz refresh rate and 1366$\times$768 pixel resolution. We used E-Prime 2.0 software to control stimulus presentation timing and collect response data.

Figure 1. Example of face material layout (from left to right: Block1 to Block3)

The experiment employed a two-factor mixed design. The between-subjects variable was face presentation mode (simultaneous, sequential), and the within-subjects variable was number of study images (1, 2, or 3).

The experimental procedure followed Sandford and Ritchie (2021). Each task comprised three blocks, with each block containing 36 trials (half match, half non-match trials), totaling 108 trials. Participants first completed a block with one-to-one matching between a target image and a single study image, followed by a one-to-two matching block, and finally a one-to-three matching block. Participants could take unlimited rest time between blocks, pressing a key to proceed to the next block after resting.

Before each block began, participants were informed that the several study images appearing at the bottom of the screen were all of the same person, while the image at the top center was the target image.

In the simultaneous face-matching task (see Figure 2 [Figure 2: see original paper]), each trial began with a white fixation cross "+" at the screen center (500 ms). Then, 1, 2, or 3 study images were presented at the bottom of the screen (the exact number determined by the block), while the target image appeared at the top simultaneously. Participants judged whether the target face and study faces depicted the same person. After making a key response, the next trial began until all trials in the current block were completed.

In the sequential face-matching task (see Figure 3 [Figure 3: see original paper]), each trial began with a white fixation cross "+" at the screen center (500 ms). Then, 1, 2, or 3 study images were presented at the bottom left of the screen

(5,000 ms; the exact number determined by the block). After a brief blank screen (500 ms), the target image appeared at the top center. Participants judged whether the target face and study faces depicted the same person. After making a key response, the next trial began until all trials in the block were completed.

Figure 2. Single-trial flowchart for simultaneous matching task (using one-to-three matching as example)
Figure 3. Single-trial flowchart for sequential matching task (using three-to-one matching as example)

After completing the identity-matching task, participants underwent a familiarity screening test to eliminate any faces they might recognize. Specifically, the names of the 36 Hong Kong celebrities used as experimental materials were presented. If participants could recall the corresponding celebrity' s face based on the name, the face was deemed familiar; otherwise, it was deemed unfamiliar. The current study only analyzed data for unfamiliar faces in the matching task.

2.2 Results and Discussion

First, data for faces identified as familiar during the familiarity screening were excluded, resulting in the removal of 6 trials (0.06% of total trials). Second, data from three participants with negative discriminability values were excluded. Thus, 78 participants were included in the final data analysis. We then conducted repeated-measures ANOVAs on discriminability (d'), response criterion (c), and reaction time. Participants' discriminability, response criterion, and reaction time results are shown in Table 1 .

Table 1. Descriptive statistical results for Experiment 1A (M±SE)

A 2 (face presentation mode: simultaneous, sequential) × 3 (number of study images: 1, 2, 3) repeated-measures ANOVA on discriminability revealed a significant main effect of image number, $F(2, 127) = 13.41$, $p < 0.001$, $p^2 = 0.15$. LSD post-hoc tests showed that discriminability for 3 study images ($M_3 = 1.59$, $SE_3 = 0.07$) was higher than for 1 study image ($M_1 = 1.25$, $SE_1 = 0.05$), $p < 0.001$, but did not differ significantly from 2 study images ($M_2 = 1.49$, $SE_2 = 0.06$), $p = 0.100$. Discriminability for 2 study images was higher than for 1 study image, $p < 0.001$. The main effect of task type was not significant, $F(1, 76) = 1.51$, $p = 0.223$, $p^2 = 0.02$. The interaction was significant, $F(2, 127) = 3.33$, $p = 0.048$, $p^2 = 0.05$.

Simple effects analysis revealed that in the sequential matching task, discriminability for 3 study images ($M_3 = 1.74$, $SE_3 = 0.10$) was higher than for 2 study images ($M_2 = 1.53$, $SE_2 = 0.09$), $p = 0.010$, and 1 study image ($M_1 = 1.23$, $SE_1 = 0.08$), $p < 0.001$. Discriminability for 2 study images was higher than for 1 study image, $p = 0.001$. In the simultaneous matching task, discriminability for 2 study images ($M_2 = 1.45$, $SE_2 = 0.09$) was higher than for 1 study image ($M_1 = 1.26$, $SE_1 = 0.08$), $p = 0.038$. These results indicate that participants' discriminability in the sequential matching task improved as the number of presented images increased, demonstrating a multiple-image advantage.

More importantly, further analysis of discriminability across the two presentation modes at different image numbers revealed that when 3 study images were presented, discriminability in the sequential task (M_{seq} = 1.74, SE_{seq} = 0.10) was higher than in the simultaneous task (M_{sim} = 1.43, SE_{sim} = 0.10), p = 0.031. When 2 study images were presented, discriminability in the sequential task (M_{seq} = 1.53, SE_{seq} = 0.09) did not differ significantly from the simultaneous task (M_{sim} = 1.45, SE_{sim} = 0.09), p = 0.518. When 1 study image was presented, discriminability in the sequential task (M_{seq} = 1.23, SE_{seq} = 0.08) did not differ significantly from the simultaneous task (M_{sim} = 1.26, SE_{sim} = 0.08), p = 0.756 (see Figure 4 [Figure 4: see original paper]).

Figure 4. Discriminability across different presentation modes for each image number condition

A 2 (face presentation mode) × 3 (number of study images) repeated-measures ANOVA on response criterion revealed a significant main effect of image number, F(2, 134) = 55.79, $p < 0.001$, $p^2 = 0.42$. LSD post-hoc tests showed that the response criterion for 3 study images (M_3 = -0.64, SE_3 = 0.05) was lower than for 2 study images (M_2 = -0.47, SE_2 = 0.04), p = 0.001, and 1 study image (M_1 = -0.22, SE_1 = 0.04), $p < 0.001$. The response criterion for 2 study images was lower than for 1 study image, $p < 0.001$. These results indicate that as the number of presented study images increased, participants adopted a more liberal (more aggressive) response criterion. The main effect of task type was marginally significant, F(1, 76) = 3.04, p = 0.085, $p^2 = 0.04$, with a lower (more liberal) response criterion in the sequential task (M_{seq} = -0.51, SE_{seq} = 0.05) than in the simultaneous task (M_{sim} = -0.38, SE_{sim} = 0.06). The interaction was not significant, F(2, 134) = 2.12, p = 0.130, $p^2 = 0.03$.

A 2 (face presentation mode) × 3 (number of study images) repeated-measures ANOVA on reaction time revealed a significant main effect of image number, F(2, 118) = 6.03, p = 0.006, $p^2 = 0.07$. LSD post-hoc tests showed that reaction time for 3 study images (M_3 = 4419 ms, SE_3 = 89 ms) was shorter than for 2 study images (M_2 = 4688 ms, SE_2 = 116 ms), $p < 0.001$, and 1 study image (M_1 = 4654 ms, SE_1 = 131 ms), p = 0.022. Reaction time for 2 study images did not differ significantly from 1 study image, p = 0.700. The main effect of task type was significant, F(1, 76) = 324, $p < 0.001$, $p^2 = 0.81$, with longer reaction times in the sequential task (M_{seq} = 6429 ms, SE_{seq} = 143 ms) than in the simultaneous task (M_{sim} = 2745 ms, SE_{sim} = 147 ms). The interaction was not significant, F(2, 118) = 0.99, p = 0.356, $p^2 = 0.01$.

In summary, Experiment 1A found that participants' discriminability improved with increasing image numbers only in the sequential matching task (i.e., a multiple-image advantage). More importantly, when three face images were presented, discriminability in the sequential condition exceeded that in the simultaneous condition. These results support the hypothesis that the multiple-image advantage is based on face representation formation rather than increased image information quantity.

A limitation of Experiment 1A was that the average reaction time in the simultaneous matching task was less than 3000 ms, meaning participants' observation time for study images was less than 3000 ms, whereas in the sequential matching task, study images were presented for 5000 ms. Therefore, the experiment could not rule out the possibility that the performance advantage in the sequential matching task resulted from differences in observation time for study images. To exclude this possibility, in Experiment 1B we controlled the presentation duration of study images to a uniform 5000 ms and made presentation mode a within-subjects variable, using three study images.

3.1.1 Participants

Forty-four university students (20 males) with a mean age of 20.8 years (range 17-34, SD = 3.3) participated. All had normal or corrected-to-normal vision and were right-handed. Using G*Power, we calculated that a minimum sample size of N = 44 was required for a paired-samples t-test with statistical power = 0.9 and effect size d = 0.5. All participants volunteered for the experiment, read and signed informed consent forms before the experiment, and received compensation afterward.

3.1.2 Stimuli and Apparatus

This experiment used the same stimuli and apparatus as Experiment 1A. The 36 individual face materials were randomly divided into two groups (n = 18 each), with different groups used for simultaneous and sequential matching tasks. Material assignment between tasks was counterbalanced.

The experiment employed a block design comprising two blocks: simultaneous matching task and sequential matching task, with task order randomized. Each task included 36 trials (half match, half non-match), totaling 72 trials. Each task included 18 different individuals (half male, half female). Before each task, the experimenter informed participants that the image at the top of the screen was the target image, and the three images below it were study images of the same person.

The simultaneous matching task procedure differed from Experiment 1A in that participants could not respond during the first five seconds of target and study image presentation (see Figure 5 [Figure 5: see original paper]). After five seconds, participants made a key response and proceeded to the next trial until all trials in the current block were completed.

The sequential matching task procedure was identical to Experiment 1A (see Figure 3).

Figure 5. Single-trial flowchart for simultaneous matching task (Experiment 1B)

3.2 Results and Discussion

First, data were screened based on familiarity, but no data were excluded in this step. Second, data from one participant with a negative discriminability index were excluded. Third, data from one participant whose reaction time in the simultaneous matching task exceeded three standard deviations were excluded. We then conducted paired-samples t-tests on discriminability and response criterion (c). Participants' discriminability and response criterion results are shown in Table 2 .

Table 2. Descriptive statistical results for Experiment 1B (M±SE)

A paired-samples t-test on discriminability revealed that participants' discriminability in the sequential face-matching task (M_{seq} = 1.92, SE_{seq} = 0.11) was higher than in the simultaneous face-matching task (M_{sim} = 1.68, SE_{sim} = 0.08), $t(41) = 2.13$, $p = 0.039$, Cohen' s $d = 0.33$. These results indicate that under sequential face presentation, the presence of multiple images facilitated the formation of a memorial representation of face identity, which helped participants perform better at face matching.

A paired-samples t-test on response criterion revealed that participants' response criterion in the sequential face-matching task (M_{seq} = -0.32, SE_{seq} = 0.06) was higher (less liberal) than in the simultaneous face-matching task (M_{sim} = -0.48, SE_{sim} = 0.06), $t(41) = 3.21$, $p = 0.003$, Cohen' s $d = 0.49$.

In summary, after eliminating the confound of face presentation duration, we still found that when multiple face images were presented, discriminability in the sequential condition exceeded that in the simultaneous condition. This finding rules out the possibility that the performance advantage in sequential matching tasks resulted from differences in observation time for study images and serves as a complement to Experiment 1A. Combined with the results of both experiments, we can further exclude the possibility that the multiple-image advantage is caused by increased information quantity.

4. Experiment 2: Effects of Image Number on Simultaneous and Sequential Matching Performance for Inverted Faces

Previous research has shown that face representation depends on the integration of configural and featural information (Itier et al., 2007), and face inversion disrupts the normal processing and integration of these two types of information (Tanaka et al., 2014), thereby affecting face representation formation. Based on this, if we invert face images to disrupt face representation formation, the multiple-image advantage in sequential presentation conditions should disappear, and the discriminability advantage of sequential over simultaneous conditions should also vanish.

4.1.1 Participants

Forty-six university students (18 females) with a mean age of 19.7 years (range 18–26, SD = 1.8) participated. All had normal or corrected-to-normal vision and were right-handed. Using G*Power, we calculated that a minimum sample size of N = 46 was required for a repeated-measures ANOVA with statistical power = 0.9 and medium effect size F = 0.25. All participants volunteered for the experiment, read and signed informed consent forms before the experiment, and received compensation afterward.

The experiment employed a two-factor mixed design. The between-subjects variable was face presentation mode (simultaneous, sequential), and the within-subjects variable was number of study images (1, 3).

Before the experiment began, participants were randomly assigned to either the simultaneous face-matching task or the sequential face-matching task. Both tasks included two blocks, each containing 36 trials (half match, half non-match), totaling 72 trials. Participants first completed a block with one-to-one matching between a study image and target image, followed by a one-to-three matching block. Different individuals were used in the one-to-one and one-to-three matching blocks. The trial procedure was identical to Experiment 1B. The simultaneous face-matching task is shown in Figure 6 [Figure 6: see original paper], and the sequential face-matching task is shown in Figure 7 [Figure 7: see original paper].

Figure 6. Single-trial flowchart for simultaneous matching task (inverted)
Figure 7. Single-trial flowchart for sequential matching task (inverted)

4.2 Results and Discussion

First, data were screened based on familiarity, but no data were excluded in this step. Second, reaction times exceeding three standard deviations were excluded, resulting in the removal of 62 data points (1.87% of total data). Third, data from four participants with negative discriminability indices were excluded. We then conducted repeated-measures ANOVAs on discriminability and response criterion (c). Participants' discriminability and response criterion results are shown in Table 3 .

Table 3. Descriptive statistical results for Experiment 2 (M±SE)

A two-factor repeated-measures ANOVA on discriminability revealed (see Figure 8 [Figure 8: see original paper]) a significant main effect of image number,

$F(1, 40) = 5.35$, $p = 0.026$, $p^2 = 0.12$. LSD post-hoc tests showed that discriminability for 3 study images ($M_3 = 0.95$, $SE_3 = 0.06$) was higher than for 1 study image ($M_1 = 0.79$, $SE_1 = 0.06$). The main effect of face presentation mode was significant, $F(1, 40) = 6.91$, $p = 0.012$, $p^2 = 0.15$. LSD post-hoc tests showed that discriminability in the simultaneous face-matching task (M_{sim} = 0.10, SE_{sim} = 0.07) was higher than in the sequential face-matching task (M_{seq} = 0.74, SE_{seq} = 0.07). The interaction was not significant, $F(1, 40) = 0.03$, $p = 0.875$, $p^2 = 0.001$. These results indicate that when faces were inverted, discriminability in the simultaneous face-matching task was higher than in the sequential task regardless of whether study images were single or multiple, and no multiple-image advantage was found in either task.

Figure 8. Discriminability indices across different presentation modes for each image number

A repeated-measures ANOVA on response criterion revealed a significant main effect of image number, $F(1, 40) = 33.67$, $p < 0.001$, $p^2 = 0.43$. LSD post-hoc tests showed that the response criterion for 3 study images ($M_3 = -0.70$, $SE_3 = 0.06$) was lower (more liberal) than for 1 study image ($M_1 = -0.40$, $SE_1 = 0.04$), $p < 0.001$. The main effect of task type was significant, $F(1, 40) = 5.95$, $p = 0.019$, $p^2 = 0.13$, with a lower (more liberal) response criterion in the simultaneous face-matching task (M_{sim} = -0.67, SE_{sim} = 0.07) than in the sequential task (M_{seq} = -0.44, SE_{seq} = 0.07). The interaction was not significant, $F(1, 40) = 0.30$, $p = 0.585$, $p^2 = 0.007$. These results indicate that when faces were inverted, participants adopted a more liberal response criterion as the number of presented images increased, and they were more liberal in the simultaneous matching task.

In summary, when faces were presented inverted, no multiple-image advantage was observed in either simultaneous or sequential face-matching tasks. Moreover, regardless of whether single or multiple face images were presented, discriminability in the simultaneous face-matching task was higher than in the sequential task. These results indicate that face inversion disrupted face representation formation, preventing observation of the multiple-image advantage phenomenon. The findings support the hypothesis that the multiple-image advantage is based on face representation formation rather than simply increased image information quantity.

This study examined people' s ability to identify faces from 1 to 3 images in simultaneous and sequential matching tasks, yielding four key findings. (1) A multiple-image advantage—where discriminability improved with increasing image numbers—was found only in the sequential face-matching task. (2) When multiple face images were presented, discriminability in the sequential condition exceeded that in the simultaneous condition. (3) When faces were inverted, discriminability in the sequential task was lower than in the simultaneous task regardless of whether study images were single or multiple. (4) When faces were inverted, no multiple-image advantage appeared in either task.

The first finding—that discriminability improved with increasing image numbers only in the sequential face-matching task (i.e., a multiple-image advantage)—is consistent with Sandford and Ritchie' s (2021) results. The simultaneous and sequential face-matching tasks in Experiment 1 correspond to those used in Experiments 1A and 2 of Sandford and Ritchie (2021). The experimental procedures and stimulus layouts were essentially identical. Although we used face images of Hong Kong celebrities sourced from their social media posts, did not strictly control image variability (lighting, angle, expression, etc.), and our participants were of a different ethnicity, we consistently found a multiple-image advantage in sequential face-matching tasks. Ritchie et al. (2021) argued that simultaneous and sequential face-matching tasks differ in nature, and participants process multiple images differently across tasks. Sequential face-matching tasks separate the presentation of multiple study images and the target image, requiring participants to remember the study faces to complete the task, thus placing high demands on memory. Simultaneous face-matching tasks present multiple study images and the target image concurrently, allowing participants to scan back and forth between target and study images, relying more on perceptual processes. Menon et al. (2015b) similarly argued that when tasks involve memory components, participants abstract stable representations from multiple face images, thereby improving face recognition performance. This view is supported by other face recognition studies involving memory (Bindemann & Sandford, 2011; Mileva & Burton, 2019; Mileva et al., 2021). For example, Mileva et al. (2021) used a face search task where one or four face images of the same individual were presented simultaneously along with surveillance footage that might contain the target face. Participants had to judge whether the target face appeared in the surveillance footage based on the face images. The results showed that presenting four face images improved discriminability for unfamiliar face recognition. In this study, even though multiple face images and target video were presented simultaneously, the search task itself required memory involvement, and thus a multiple-image advantage was observed. In summary, these results suggest that the multiple-image advantage is based on the formation of robust representations that require memory for construction and storage.

The second finding—that discriminability in the sequential condition exceeded that in the simultaneous condition when multiple face images were presented—deserves attention. In general object recognition, simultaneously presented stimuli provide more information for identification and judgment. In contrast, sequential presentation requires participants to extract and memorize object information within a certain time frame and make identifications based on memorized information, implying some degree of information loss during processing. However, in the current study, this information loss did not impair recognition performance but instead improved participants' ability to discriminate between different faces. Burton et al. (2005) proposed that when learning new faces, people form an abstract average representation from multiple face images of the same individual. This representation retains identity-related information while eliminating surface effects caused by lighting, photographic equipment, and variations due to emotion or health status. This aligns with Devue and de Sena' s (2023) cost-efficient theory, which similarly posits that people tend to assign higher representational weight to stable facial features. That is, people prioritize encoding facial features that are identity-related and stable across multiple images while discarding surface information that is identity-unrelated and variable across images (image variability). This neglected image variability is often considered an important cause of poor unfamiliar face recognition performance (Bruce et al., 1999; Megreya & Burton, 2006, 2008; Kramer & Ritchie, 2016). This explains the improved discriminability in sequential tasks. However, this processing was only observed in memory-involved sequential face-matching tasks, suggesting that representation formation from multiple images depends on memory involvement. Li and Li (2018) provided further theoretical explanation for this view. When participants form representations based on cues, they need to extract corresponding information from memory, form representations, and store them in memory. Meanwhile, representational precision decreases due to decay or interference during memory maintenance. In summary, people construct and store abstract representations of multiple face images in memory. On one hand, these representations contain stable face identity information across images (Peng et al., 2019); on the other hand, the formation process discards image-irrelevant distracting information, a processing mechanism that facilitates face recognition.

The third finding—that discriminability in the sequential task was lower than in the simultaneous task regardless of whether image sets were single or multiple —can be interpreted in two ways. First, in general object recognition, simultaneous presentation tasks allow participants to scan different stimuli repeatedly, providing more information for identification and judgment. Sequential presentation tasks require participants to observe and memorize objects within a limited time, causing some information loss. Therefore, simultaneous presentation tasks confer a processing advantage for general object recognition. The results of Experiment 2 align with this explanation, suggesting that discrimination of inverted faces follows the pattern of general object recognition. Second, because face inversion disrupts holistic processing and representation of faces, both simultaneous and sequential tasks were compared under more "fair" conditions, revealing the inherently higher difficulty of sequential tasks. This finding also suggests that previous research and theoretical perspectives may have overlooked the potential influence of task difficulty differences when interpreting performance differences between the two tasks, thereby highlighting the theoretical necessity of testing and excluding the "interaction between image information quantity and task difficulty" hypothesis in the current study.

The fourth finding—that no multiple-image advantage was observed in either task when faces were inverted—not only perfectly matches the a priori predictions of the "memory facilitates representation"hypothesis but also demonstrates that the "image information quantity" variable does not play a significant role in either task when images are inverted. Therefore, these results suggest that both face representation formation and face image perceptual processing depend on upright face presentation. This aligns with findings and perspectives from previous research. Due to long-term experience perceiving and remembering upright faces, people develop expertise in upright face recognition that is strongly orientation-dependent. Combined with these views, we can conclude that participants' construction of face representations from multiple images in sequential tasks is a cognitive process based on face expertise.

Additionally, this study found that as the number of faces in the presented image set increased, participants adopted increasingly liberal response criteria. This result is consistent with previous research (Matthews & Mondloch, 2018; Menon & Kemp, 2015b; Ritchie et al., 2021; Sandford & Ritchie, 2021). On one hand, Tanaka et al. (1998) proposed that a single face representation can be activated by multiple face inputs. This many-to-one mapping from stimulus input to face memory is called an attractor field. When a face image falls into a specific attractor field, it is categorized as the corresponding individual. When face representations integrate more images, the attractor field for that individual becomes larger, leading to higher face recognition accuracy and more "same" responses (Menon et al., 2015a). This partially explains our finding that participants had more liberal response criteria and higher accuracy in sequential matching tasks. On the other hand, this response bias may be related to participants' cognitive load. In a study by Mileva et al. (2019), participants were asked to observe face videos at the top of the screen and find matching faces in surveillance videos at the bottom. The results showed poor recognition performance. The researchers argued that presenting too much information simultaneously created excessive cognitive load, preventing participants from utilizing additional information. This may lead participants to be more inclined to make "same" judgments (Ritchie et al., 2021). This explanation can account for the results of Experiment 2. When faces were inverted, task difficulty increased, and simultaneously presenting multiple face images overloaded participants' cognitive capacity, causing them to be more inclined to make "same" judgments.

6. Limitations and Future Directions

First, this study only considered the effect of image number on face representation formation within a limited range. In reality, people can benefit from a larger number of face images. Current research has found that when the number of study images is small (1 to 4), face recognition performance improves significantly as the number of study images increases (Bindemann & Sandford, 2011; White et al., 2014; Menon et al., 2015a; Menon et al., 2015b; Mileva & Burton, 2019; Ritchie et al., 2021; Sandford & Ritchie, 2021). However, when the number of study images reaches a certain level (5 or more), face recognition performance shows little change with increasing image numbers. We therefore speculate that there exists a "minimum number of images required to form a robust face representation." Future research could examine the effects of larger numbers of images on face representation formation or investigate this question using computational modeling approaches.

Second, this study only used a 5-second presentation duration for multiple images, so we cannot know whether face representation would form under shorter presentation durations. Previous research indicates that this representation formation process is extremely rapid (White et al., 2014; Dunn et al., 2018). For example, White et al. (2014) used a face-matching paradigm and found that varying the time from image presentation to when participants could respond (3, 6, or 9 seconds) did not affect recognition performance. Therefore, representation formation for four face images may occur within 3 seconds. Future research should shorten image presentation durations to explore the time course of face memory representation formation.

Finally, normal face familiarization involves a longer time dimension (Wang et al., 2023), whereas this study focused on the process of face representation formation from scratch. Currently, it remains unclear how these initial face representations transform into robust familiar face representations. Future research should further investigate the roles of image number, image variability, and other factors in this process over longer time spans. Additionally, research by Baker and Mondloch (2023) suggests that participants' ability to benefit from multiple images is related to their own face recognition ability. Future studies should incorporate participants' face recognition ability as a variable.

This study found that a multiple-image advantage existed only under sequential presentation conditions; when multiple images were used for discrimination, performance in the sequential condition was superior to that in the simultaneous condition. This indicates that the multiple-image advantage in face identity discrimination is based on face representation formation rather than simply increased image information quantity. Furthermore, we found that when faces were inverted, participants showed no multiple-image advantage, and performance in the simultaneous condition was superior to that in the sequential condition. This indicates that when the face representation formation process is disrupted, the multiple-image advantage disappears. These findings reveal that face representation formation depends on memory processes, providing a new perspective for understanding the cognitive mechanisms of face identity recognition and guiding future research.

References

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology, 68*(10), 2041−2050.

Baker, K. A., & Mondloch, C. J. (2019). Two sides of face learning: Improving between−identity discrimination while tolerating more within−person variability in appearance. *Perception, 48*(11), 1124−1145.

Baker, K. A., & Mondloch, C. J. (2023). Unfamiliar face matching ability predicts the slope of face learning. *Scientific Reports, 13*(1), 5248.

Baker, K. A., Laurence, S., & Mondloch, C. J. (2017). How does a newly encountered face become familiar? The effect of within−person variability on adults' and children' s perception of identity. *Cognition, 161*, 19−30.

Baker, K. A., Stabile, V. J., & Mondloch, C. J. (2023). Stable individual differences in unfamiliar face identification: Evidence from simultaneous and sequential matching tasks. *Cognition, 232*, 105333.

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo−IDs of the same person. *Perception, 40*(5), 625−627.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*(4), 339−360.

Burton, A. M., Jenkins, R., Hancock, P. J., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology, 51*(3), 256−284.

Davis, E. E., Matthews, C. M., & Mondloch, C. J. (2021). Ensemble coding of facial identity is not refined by experience: Evidence from other-race and inverted faces. *British Journal of Psychology, 112*(1), 265−281.

Devue, C., & de Sena, S. (2023). The impact of stability in appearance on the development of facial representations. *Cognition, 239*, 105569.

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *Quarterly Journal of Experimental Psychology, 69*(1), 1−10.

Dunn, J. D., Kemp, R. I., & White, D. (2018). Search templates that incorporate within−face variation improve visual search for faces. *Cognitive Research: Principles and Implications, 3*(1), 37.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*(2), 159−170.

Honig, T., Shoham, A., & Yovel, G. (2022). Perceptual similarity modulates effects of learning from variability on face recognition. *Vision Research, 201*, 108128.

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early face processing specificity: it' s in the eyes! *Journal of Cognitive Neuroscience, 19*(11), 1815−1826.

Jones, S. P., Dwyer, D. M., & Lewis, M. B. (2017). The utility of multiple synthesized views in the recognition of unfamiliar faces. *Quarterly Journal of Experimental Psychology, 70*(5), 906-918.

Kramer, R. S., & Reynolds, M. G. (2018). Unfamiliar face matching with frontal and profile views. *Perception, 47*(4), 414–431.

Kramer, R. S., & Ritchie, K. L. (2016). Disguising superman: How glasses affect unfamiliar face matching. *Applied Cognitive Psychology, 30*(6), 841–845.

Kramer, R. S., Hardy, S. C., & Ritchie, K. L. (2020). Searching for faces in crowd chokepoint videos. *Applied Cognitive Psychology, 34*(2), 343-356.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology: Section A, 53*(2), 513–536.

Li, X., & Li, H. (2018). The relationship between imagery, perception, and memory from the perspectives of representation and cognitive processes. *Psychological Science, 41*(03), 520–525.

Longmore, C. A., Santos, I. M., Silva, C. F., Hall, A., Faloyin, D., & Little, E. (2017). Image dependency in the recognition of newly learnt faces. *Quarterly Journal of Experimental Psychology, 70*(5), 863–873.

Matthews, C. M., & Mondloch, C. J. (2018). Finding an unfamiliar face in a line-up: Viewing multiple images of the target is beneficial on target-present trials but costly on target-absent trials. *British Journal of Psychology, 109*(4), 758–776.

Matthews, C. M., & Mondloch, C. J. (2022). Learning faces from variability: Four-and five-year-olds differ from older children and adults. *Journal of Experimental Child Psychology, 213*, 105259.

Matthews, C. M., Ritchie, K. L., Laurence, S., & Mondloch, C. J. (2024). Multiple images captured from a single encounter do not promote face learning. *Perception, 53*(5-6), 299-316.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34*, 865–876.

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14*(4), 364–372.

Megreya, A. M., White, D., & Burton, A. M. (2011). The other-race effect does not rely on memory: Evidence from a matching task. *Quarterly journal of experimental psychology, 64*(8), 1473-1483.

Menon, N., Kemp, R. I., & White, D. (2018). More than a sum of parts: robust face recognition by integrating variation. *Royal Society Open Science, 5*(5), 172381.

Menon, N., White, D., & Kemp, R. I. (2015a). Variation in photos of the same face drives improvements in identity verification. *Perception, 44*(11), 1332–1341.

Menon, N., White, D., & Kemp, R. I. (2015b). Identity−level representations affect unfamiliar face matching performance in sequential but not simultaneous tasks. *Quarterly Journal of Experimental Psychology, 68*(9), 1841−1851.

Mileva, M., & Burton, A. M. (2019). Face search in CCTV surveillance. *Cognitive Research: Principles and Implications, 4*, 37.

Mileva, V. R., Hancock, P. J. B., & Langton, S. R. H. (2021). Visual search performance in 'CCTV' and mobile phone−like video footage. *Cognitive Research: Principles and Implications, 6*, 63.

Murphy, J., Ipser, A., Gaigg, S. B., & Cook, R. (2015). Exemplar variance supports robust learning of facial identity. *Journal of Experimental Psychology: Human Perception and Performance, 41*(3), 577-581.

Peng, S., Kuang, B., & Hu, P. (2019). Memory of ensemble representation was independent of attention. *Frontiers in Psychology, 10*, 228.

Pitcher, D., Caulfield, R., & Burton, A. M. (2023). Provoked overt recognition in acquired prosopagnosia using multiple different images of famous faces. *Cognitive Neuropsychology, 40*(3-4), 158-166.

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology, 70*(5), 897−905.

Ritchie, K. L., Kramer, R. S., Mileva, M., Sandford, A., & Burton, A. M. (2021). Multiple−image arrays in face matching tasks with and without memory. *Cognition, 211*, 104632.

Ritchie, K. L., Mireku, M. O., & Kramer, R. S. (2020). Face averages and multiple images in a live matching task. *British Journal of Psychology, 111*(1), 92−102.

Sandford, A., & Ritchie, K. L. (2021). Unfamiliar face matching, within−person variability, and multiple−image arrays. *Visual Cognition, 29*(3), 143−157.

Tanaka, J., Giles, M., Kremen, S., & Simon, V. (1998). Mapping attractor fields in face space: the atypicality bias in face recognition. *Cognition, 68*(3), 199-220.

Tanaka, J. W., Kaiser, M. D., Hagen, S., & Pierce, L. J. (2014). Losing face: impaired discrimination of featural and configural information in the mouth region of an inverted face. *Attention, Perception, & Psychophysics, 76*, 1000−1014.

Wang, Z., Ni, H., Feng, D., Yan, L., & Sun, Y. (2023). Regional asynchrony and eye region-specificity in part-based processing and holistic processing during face familiarization. *Acta Psychologica Sinica, 55*(6), 861-876.

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo−ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied, 20*(2), 166−173.

Xu, B., & Tanaka, J. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. *Journal of Vision, 13*(2), 22−22.

Yin, R. K. (1969). Looking at upside−down faces. *Journal of Experimental Psychology, 81*(1), 141−145.

Young, A. W., & Burton, A. M. (2018). Are we face experts? *Trends in Cognitive Sciences, 22*(2), 100−110.

Note: Figure translations are in progress. See original paper for figures.