Do image variability and names in missing person appeals improve prospective person memory?

Stefana Juncu[1], Hartmut Blank[1], Ryan J. Fitzgerald[2] and Lorraine Hope[1]

[1]University of Portsmouth

[2]Simon Fraser University

The publisher version is available here: https://doi.org/10.1016/j.jarmac.2020.07.001

Author Note

This research was supported by a South Coast Doctoral Partnership grant (ES/P000673/1) from the Economic and Social Research Council to the first author (project reference: 1956663) as a PhD studentship.

Stefana Juncu, Department of Psychology, University of Portsmouth; Hartmut Blank, Department of Psychology, University of Portsmouth; Ryan J. Fitzgerald, Department of Psychology, Simon Fraser University; Lorraine Hope, Department of Psychology, University of Portsmouth

Corresponding author: Stefana Juncu
Email: Stefana.juncu@port.ac.uk
Phone: +44 23 9284 6318
Address: Department of Psychology, University of Portsmouth, King Henry Building, King Henry I Street, Portsmouth, UK, PO1 2DY

Abstract

Prospective person memory is implicated in searches for missing or wanted individuals. We investigated whether prospective person memory is improved by associating the target of the search with a name and providing photos that reflect variation in the search target's appearance. Participants (N = 242) studied three photographs of each target, taken either at the same event (low variability) or at different events (high variability). For half of the participants, a name was presented alongside the photographs. Both names and high variability photos improved discriminability, suggesting that public appeals for a missing or wanted person should include a name and use images that are representative of the person's variability in appearance across different contexts.

Keywords: prospective person memory, missing people, variability, names

General Audience Abstract

When a person goes missing, the police sometimes release an appeal asking the public to be on the lookout for that missing person. In an online experiment, we asked 242 participants to be on the lookout for a number of target individuals while completing another task. Our participants learned the face of a target individual from either three similar photographs (i.e., low variability) or three dissimilar photographs (i.e., high variability). The similar photographs were taken on the same day and, therefore, provided less information about how that person might look like on a different day. The high variability photographs were taken during three different days and showed how the person might look like in different contexts. For half of our participants we also presented a name next to each target's pictures. The other half of participants only observed their pictures. Learning a face from photographs taken at different events (i.e., high variability) resulted in improved performance on a subsequent search. Associating each face with a name also improved performance. Namely, participants were better at discriminating between targets and other individuals when they learned the targets from multiple high variability images alongside names than from low variability images or without names. Interestingly, providing high variability photographs improved performance more than including names. Including multiple high variability images in appeals is likely to enhance missing person investigations

Do image variability and names in missing person appeals improve prospective person memory?

Prospective memory (PM) is the ability to remember to execute delayed intentions when a particular cue is encountered (McDaniel & Einstein, 1993). In prospective *person* memory (PPM), the cue is *a person*. In PPM tasks, an individual is typically on the lookout for a target person whilst completing other attention-demanding activities (Lampinen, Arnal, & Hicks, 2009). Such scenarios have important real-life applications for the recovery of missing people and the apprehension of wanted criminals. However, research suggests PPM is a surprisingly difficult task (Lampinen, Curry, & Erickson, 2016; Moore, Lampinen, & Provenzano, 2016).

Lampinen and Moore (2016) conceptualised PPM tasks as a series of preconditions: (1) encountering the alert, (2) attending to the alert, (3) encountering the missing person, (4) attending to the missing person, (5) reminding, (6) taking action. Each precondition occurs with some probability, varying on a case by case basis. Researchers usually vary the probabilities of these preconditions and observe the effect on PPM performance. For example, wider distribution of appeals via social media and creating alerts in ways that attract attention (e.g., including images) increases the probabilities of preconditions (1) and (2), respectively, and manipulating the information included in the alert may increase the likelihood of both (2) and (4). Generally, then, successful PPM requires a combination of attention, face recognition, and event-based prospective memory (Lampinen & Moore, 2016).

Prospective memory tasks (e.g., remembering to call the police if you encounter a missing person) are often performed in the context of other ongoing activities (e.g., buying the groceries), and this is mimicked in experimental research via dual-task procedures (Einstein & McDaniel, 2005). The primary task is an ongoing attention-demanding task designed to simulate real-life everyday activities (e.g., grocery shopping). The secondary task involves the prospective memory action (e.g., calling the authorities). Participants are instructed to be on the lookout for a cue (e.g., target person) during the ongoing task and perform an action when the cue is encountered (e.g., press the ‘spacebar’). A good real-life example of this type of dual-task activity is the role of a nightclub doorman, whose job is to verify that everyone who enters a club is above a certain age (primary task) while also keeping on the lookout for anyone who is banned from the club (secondary task). In the present research, we designed an experiment around this scenario.

Performance on this task partly depends on the ability to recognize the target when encountered. In the context of PPM, targets may be previously unfamiliar and, contrary to our exceptional ability to recognize familiar faces, unfamiliar face recognition is difficult (Johnston & Edmonds, 2009). One critical difference between familiar and unfamiliar face recognition is the richness of perceptual information that we have for a familiar face, affording a view-invariant representation (i.e., a face recognition unit; Bruce & Young, 1986). This enables people to recognise familiar faces even from different viewpoints and in different contexts. Correspondingly, when learning a new face, using multiple photos increases the amount of perceptual information and results in better recognition than learning a face from a single photo (Ritchie & Burton, 2017). This effect of learning a new face from multiple images has also been observed in PPM research: participants who saw three – as opposed to just one – photos of each target reported more accurate sightings (Sweeney & Lampinen, 2012).

Importantly, this multi-image advantage may depend on the variability of the images used for face learning. Individuals look different each day (e.g., varying hairstyles, changes in lighting), and encoding of within-person variability is needed to develop a stable representation of their identity (Kramer, Jenkins, Young, & Burton, 2017). Using an old/new recognition paradigm, Ritchie and Burton (2017) showed that using ten high variability images of a target (from different times and locations) results in better face recognition performance than using ten low variability images of the same target (screenshots from one interview). Given that face recognition is required for successful PPM, presentation of high variability photographs at encoding might also improve PPM performance compared to low variability photographs. Notably, the Sweeney and Lampinen (2012) study – which found a multi-image advantage – used high variability photographs that were taken on different days, with different clothing and hairstyles (and were presented simultaneously alongside other information such as name, age, height, and weight). In order to establish that high image variability is a crucial condition for improving PPM performance, it is important to compare it against multiple images with low variability; this was one of the goals of the present study.

In addition to the perceptual information that is encoded when a new face becomes familiar, we also acquire semantic, conceptual and social information about the person (Bruce & Young, 1986). When a familiar face is represented in memory, it is usually associated with social information such as names, occupations, and hobbies. In previous face recognition research, associating unfamiliar faces with socially relevant information during learning improved face recognition (Bower & Karlin, 1974; Courtois & Mueller, 1979; Tanaka & Pierce, 2009). One type of information that is easy to provide in real-life PPM cases (e.g., a missing person investigation) is the target's name.

Associating target people with names should not only help participants discriminate between different individuals but should also help them integrate different images of the same individual into one representation. Tagging each target with a name is a simple way of informing participants about which of the targets is shown in each photograph. Previous researchers (Hugenberg, Young, Bernstein, & Sacco, 2010; McGugin, Tanaka, Lebrecht, Tarr, & Gauthier, 2011; Tanaka & Pierce, 2009) showed that associating faces or objects with labels, such as names, encourages individuation; the act of discriminating among individual faces within a group. Individuation may not be the only benefit of associating faces with names, though. Schwartz and Yovel (2016, 2019a, 2019b) found that, although associating faces with person-related labels (e.g., names, occupation) improved face recognition, this effect was not present when faces were associated with person-unrelated labels (e.g., symbols, object names). Instead, they argue that associating unfamiliar faces with names during encoding improves recognition because it generates a conceptual representation of that *person* in addition to the perceptual representation of that *face***.** Therefore, associating names with faces during familiarization might also improve PPM performance.

When the effect of presenting multiple images of a PPM target was tested previously, all images of a given target were presented simultaneously (Lampinen, Curry, & Erickson, 2016; Sweeney & Lampinen, 2012). In real-life cases, photographs are not always presented simultaneously. For example, if a person goes missing, different appeals might be created and distributed by their friends, family or the police. Therefore, the public might repeatedly see the same, or different pictures of the target (i.e., missing person), whilst encountering other images or other individuals in between. In such cases, the face of the missing person is presented in a distributed fashion along with other faces. Missing person appeals are also increasingly distributed through social media, which allows multiple photographs to be included and presented sequentially (Solymosi, Petcu, & Wilkinson, 2020). Although less discussed in face learning studies, learning stimuli (e.g., nonsense words, word pairs) using a distributed sequence results in better recognition (Russo, Mammarella, & Avons, 2002; Wogalter, Jarrard, & Cayard, 1991). Therefore, we used a distributed presentation of target images during the familiarization phase of our experiment.

When the presentation of high variability images is distributed, participants might not realise that they show the same individual. Jenkins, White, Montfort and Burton (2011) found that when participants were asked to sort 40 images into piles by identity, most participants mistook photos of the same person as photos of different persons, thus overestimating the total number of people portrayed in the photos. With distributed presentation, participants need to recognize whether a person has been shown previously, so they can update their memory representation for that particular person. Therefore, associating each target face with a name might help participants link high variability images into one representation of each target.

The Current Study

The aim of the current research was to explore the effect of using high variability images and name associations on PPM. Using a dual task paradigm, we designed a scenario in which participants assumed the role of nightclub security staff and were on the lookout for people banned from entering the premises while also checking the ages of everyone who tried to enter. Half of the participants learned each target's images alongside their name and half without (IV 1). During the familiarization phase, participants learned new target faces from three photographs taken at the same event (i.e., low variability images) or taken during three different events (i.e., high variability images; IV 2).

We predicted that using high variability images during encoding would improve PPM. There are two different ways in which associating names might further improve performance. Based on the notion of individuation (Tanaka & Pierce, 2009), including names alongside high variability photos would increase the likelihood that participants would link together the different photos of a target into a single identity representation. This predicts a name advantage when using high variability images but not when using low variability images (i.e., an interaction). Alternatively, associating photos with names could result in a more generalized benefit to PPM performance by encouraging participants to represent faces in memory as socially meaningful concepts, as opposed to more basic percepts (Schwartz & Yovel, 2016). According to this perspective, including names should result in better discriminability regardless of photo variability (i.e., a main effect of names).

Photo variability and names could also affect response bias. If participants do not correctly link together the images of each target, they might misjudge the number of individuals presented and report more sightings. Therefore, we also hypothesised that response bias would be more liberal in the no-name, high variability condition compared to all other conditions.

Method

Participants

We recruited 270 participants using the *Qualtrics* recruitment services. All participants were paid for their time. All exclusion criteria were set prior to collecting the data. To be eligible to participate, respondents were required to be fluent in English, to complete the task on a computer (not a mobile device), to be at least 18 years old, and to have normal or corrected to normal vision. We included two attention-check questions, and the experiment immediately ended for any participant who failed one of these questions. Participants were also asked if they were familiar with any of the targets and they all responded no. We further excluded 28 participants (14 from each name condition) because they did not report any sightings of either targets or non-targets (suggesting that they did not engage with the prospective memory task). Removing these 28 participants did not qualitatively change the pattern of findings (see Supplemental Materials for detailed analyses of the complete data set). Our final sample included 242 participants (M_{age}= 56.00, SD = 14.82, range: 18-86 years, 40% male, 92% Caucasian).

G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) indicated 264 participants were needed for 80% power to detect a small effect (Cohen's $f = 0.15$) in the between-subjects comparison of names versus no names (alpha = 0.05). The power analysis was performed for a mixed design, to account for the within-subjects manipulation of photo variability. Although the final sample had 20 participants fewer than planned (because we had to exclude more participants than anticipated), we still had 80% power to detect an effect of $f = 0.16$ or larger.

Design

This experiment used a 2 (target name: name vs no name) by 2 (photograph variability: low vs high) mixed design. Both variables were manipulated during the encoding phase. Half of the participants saw photographs of named targets and the other half saw photographs of the same targets without names associated with them. Photograph variability was manipulated within subjects; all participants learned half of the targets from low variability images and half from high variability images. For each participant we counted the correct and incorrect identifications they made and used them to calculate discriminability and response bias measures.

Materials

Across four blocks, participants were instructed to learn four unfamiliar targets. Overall, participants studied images associated with 16 targets (50% male). Targets were online personalities (e.g., YouTubers, TV presenters) from foreign countries (participants were recruited from U.S. and U.K.) and varied in age (estimated range: 20s to 50s; exact ages were rarely available) and ethnicity. A norming sample of $N = 21$ participants were asked to estimate the ages of all targets and to indicate if they looked above, around, or under 25 years; 5 targets were classified as 'definitely under 25', 5 targets as 'definitely above 25', and

6 were classified as 'around 25' (see Figure S1, Supplemental Materials). The assignment of each target to each of the two variability conditions was counterbalanced. We assessed whether the targets' perceived age affected the percentage of participants who correctly identified them. We found no significant difference between targets who were definitely above 25 years old compared with targets who seemed around 25 years old and with targets who were definitely younger than 25 years old; F (2, 13) = 2.68, p = .11.

All names associated with target images were fictitious and were found on online lists of common English names. For the high-variability condition, we used three highly variable photographs from different events, which contained a range of within-person variability (e.g. variability in clothing, make-up, hairstyle, facial hair, but also facial/emotional expression). These images were always captured in different contexts, with different backgrounds, different clothing and from different angles. For the low-variability condition, we used three photographs from the same event (e.g., a single interview, broadcast, or YouTube video). These images varied in facial expressions or head angle, but not on other appearance dimensions such as hair style or makeup.

To find the stimuli, we first searched for foreign personalities who appeared in multiple videos online (e.g., video interview, presentation, YouTube video). Once enough images were found for someone to be used as a target, three of their images from one event were downloaded for the low variability condition. For the high variability condition, we used one randomly picked image from the three images already downloaded, and we downloaded two more images from another two events. For both low and high variability conditions, one final image was downloaded for the testing phase from a fourth event. Therefore, the photograph used during testing was always novel for all participants, regardless of the variability condition.

During the prospective person memory task, targets were intermingled with unfamiliar distractor faces (16 per block). Most images of the distractor faces were obtained from the Faces in the Wild Database (Huang, Ramesh, Berg, & Learned-Miller, 2007) which includes ambient images of foreign personalities. Faces from this database were chosen to be similar enough for the targets to not stand out, and also to vary in age and ethnicity. We also included screen captures of other non-targets from Google videos (resembling the videos that target images were captured from). All images were ambient style and varied in pose, lighting, expression, focal length, etc.

Procedure

Participants were asked to imagine that they were a bouncer at a nightclub. They were instructed that their primary task was to decide who is old enough to enter the club and that the minimum age of entry was 25 years old. Participants were also told that, at the beginning of each testing block, they would study images of individuals who had been banned from the premises and that, if they saw any of these individuals during the testing phase, they should click on their picture to deny their entry (this was the secondary task, i.e., the prospective memory task). The software then randomly assigned participants to one of the two name conditions (i.e., named or not named targets). Each participant completed four blocks of trials: two blocks with low variability photographs and two blocks with high variability photographs. The order of the variability blocks was counterbalanced, with the two low variability and two high variability blocks always presented one after the other.

Each block included a familiarization phase and a test phase. In the familiarization phase, participants encoded four targets (i.e., those individuals banned from the club), with three photographs per identity. Participants were instructed to try to remember all target faces but were not told how many targets they would see. All images were shown one at a time using a distributed presentation. Specifically, the three images of any given target were separated by images of other targets (see Figure 1). The familiarization phase was self-paced, such that participants had unrestricted time to look at each photo.

After participants saw all 12 images (3 images of 4 targets each), the familiarization phase was completed and they were reminded of the main task (i.e., to decide if an individual is 25 years or older) and the secondary task (i.e., to click on their photo if they think the individual is banned). Then they saw 20 individuals (4 targets and 16 non-targets), one person at a time, and had to respond with ‘Yes’ or ‘No’ to the question ‘Are they old enough to go in?’. For each person presented, they could also click on the photograph to identify them as one of the target individuals, but this was not cued in any way during individual trials. During the testing phase, all individuals were presented without names, and a novel image of each target was used. After each block was completed, participants were informed that a new block would commence by showing the images of four new targets. They were permitted to take a self-timed break before a new block started. After participants completed all four blocks, they were thanked for participating and the experiment ended.

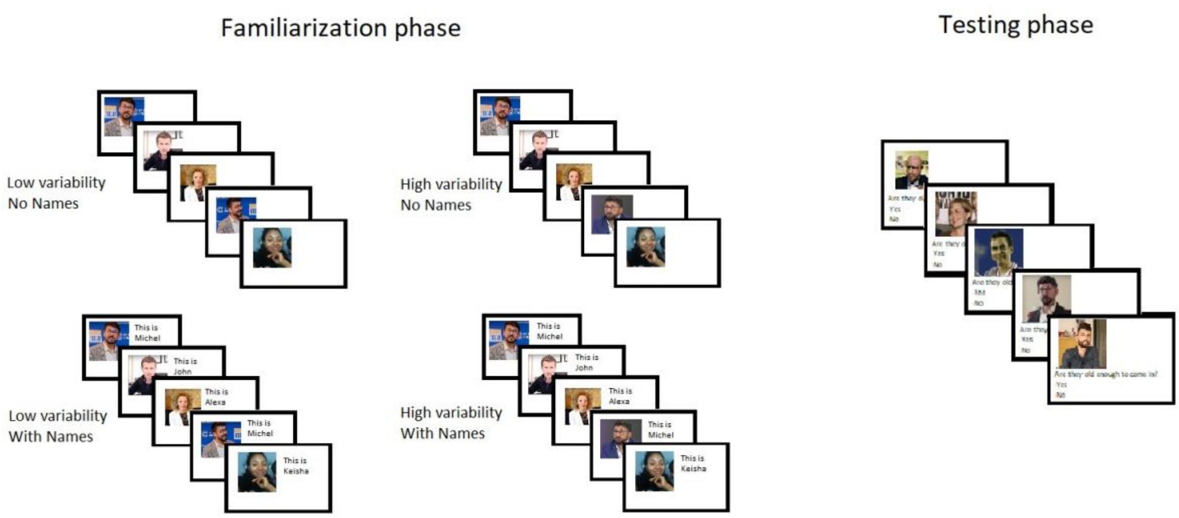

Figure 1. Examples of low and high variability images, with and without names presented during the familiarization phase and examples of target and nontarget images presented during the testing phase. In the experiment, all images were in colour.

Results

For each participant, we calculated the proportion of targets correctly identified (hits) and the proportion of non-targets mistakenly identified (false alarms). Following standard Signal Detection Theory methodology (Macmillan & Creelman, 2004), we used each participant's hit and false alarm rates to compute discriminability (*d'*) and response bias (*c*) measures. For participants with hits or false alarm rates of one or zero we used the Macmillan and Kaplan (1985) correction to adjust their scores. Specifically, we subtracted or added 0.5/8 (N_{targets}) from the number of hits and subtracted or added 0.5/32 ($N_{\text{non-targets}}$) to the number of false alarms. Table 1 presents descriptive statistics for each experimental condition, including hits, false alarms, discriminability (*d'*) and response bias (*c*).

Table 1

Means (SDs) for number of hits, false alarms, discriminability and response bias

	Without Names		With Names	
	Low Variability	High Variability	Low Variability	High Variability
Hits	0.61 (0.25)	0.70 (0.22)	0.66 (0.24)	0.73 (0.24)
False alarms	0.08 (0.10)	0.08 (0.10)	0.07 (0.08)	0.05 (0.08)
Discriminability (*d'*)	1.96 (0.78)	2.32 (0.81)	2.18 (0.79)	2.54 (0.82)
Response bias (*c*)	0.65 (0.55)	0.53 (0.53)	0.59 (0.52)	0.55 (0.52)

Discriminability

To test our main hypotheses, we ran a 2 (names) × 2 (variability) mixed ANOVA on *d'* scores. The model revealed a significant effect of names; F (1, 240) = 6.64, p = .01, n^2_p =

.03. Presenting names alongside photographs (*M* = 2.36, *SD* = 0.66) increased discriminability relative to presenting photographs with no names (*M* = 2.14, *SD* = 0.67; see Figure 1). We also observed a much stronger main effect of variability, *F* (1, 240) = 41.01, *p* < .001, n^2_p = .15. Participants were better at discriminating between targets and non-targets when high variability photographs (*M* = 2.43, *SD* = 0.82) were used compared with low variability photographs (*M* = 2.07, *SD* = 0.79). The interaction was not significant, *F* (1, 240) < 0.01, *p* = .97, n^2_p = 0.01.

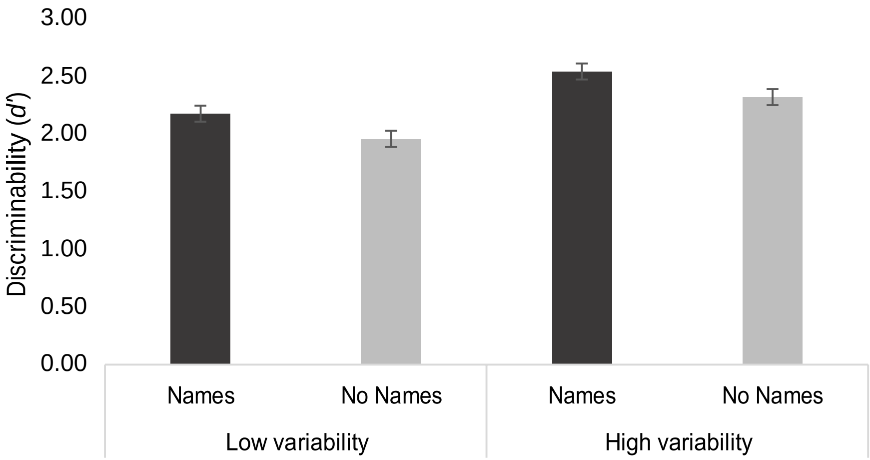

Figure 1. Mean discriminability (*d'*). Error bars represent standard errors.

Response bias

The same ANOVA on the response bias measure revealed no significant effect of names, *F* (1, 240) = 0.12, p = .73, n^2_p = .01, a significant main effect of variability, *F* (1, 240) = 4.16, *p* = .04, n^2_p = .02, and no significant interaction, *F* (1, 240) = 0.90, *p* = .34, n^2_p = .01. Participants reported more sightings for high variability targets than low variability targets.

In the no names condition, we hypothesised that high variability photographs would increase sightings compared with low variability photographs (i.e., c will decrease). A repeated measures *t*-test including only participants from the no-names condition revealed a significant difference between the low and high variability conditions, *t* (120) = 1.99, *p* =

.049, $d = 0.18$. As seen in Table 1, this difference was driven by an increase in correct sightings, t (120) = 3.43, $p = .001$, $d = 0.31$. The number of incorrect sightings was not significantly different; t (120) = 0.27, $p = .79$, $d = 0.02$.

We also hypothesised that within the high variability condition, participants would have a more liberal response bias if no names are presented than if names are presented. Contrary to our hypothesis, an independent samples *t*-test revealed no significant difference between the name and no name conditions, t (240) = 0.32, $p = .75$, $d = 0.04$.

Primary Task Performance

In further analyses, we sought to put the main findings into the context of primary task performance. First, we assessed whether we successfully emulated a dual task procedure as in a real PPM situation. Theoretically, participants in our study could have ignored the primary task and focused exclusively on identifying the targets. To answer this, we used decision accuracy as an indicator of participants' engagement with the primary task (i.e., deciding if the non-target nightclub visitors were above or below 25 years); if participants were focusing on the primary task, they should perform at a level clearly above chance. Assessing decision accuracy is not trivial, though, as (a) there is no absolute ground truth (we do not know the actual ages of the non-targets) and (b) better than chance performance can only be expected for non-targets clearly above or below the threshold age.

We resolved these issues through our norming study (see above and Supplemental Materials), by drawing on non-targets who were seen as either clearly above or below 25 years with a high degree of agreement. More than 90% of norming participants agreed that 5 non-targets were definitely below, and 25 non-targets were definitely above 25 years old. Decision accuracy ($M = 0.95$, $SD = 0.10$) for this subset of non-targets was significantly higher than chance; t (241) = 70.86, $p < .001$ (based on a one-group t-test against 50%

performance). Excluding participants who scored less than 85% correct on this measure of ongoing task performance did not qualitatively change the pattern of findings (see Supplemental Materials for detailed analyses).

Second, it would be reassuring to know that these indicators of primary task engagement did not differ between any of our conditions (and therefore cannot compromise any of our main findings regarding those conditions). We ran a 2 (names) x 2 (variability) mixed ANOVA on decision accuracy. There were no significant differences in primary task performance between participants in the name (M = 0.96, SD = 0.10) or no-name conditions (M = 0.95, SD = 0.11); F (1, 240) = 0.45, p = .51; or between the low variability (M = 0.96, SD = 0.11) and the high variability conditions (M = 0.95, SD = 0.12); F (1, 240) = 3.80, p = .052.

Further, primary task performance was positively correlated with both low variability (r (242) = .18, p = .01) and high variability discriminability in the PPM task (r (242) = .15, p = .02). This probably reflects that participants with better attentional and cognitive capacities, or who paid more attention overall, did better in both tasks; in any case, PPM performance did not come at the expense of primary task performance.

Briefly, these findings show that our participants paid attention to the primary task. Decision accuracy did not vary based on experimental conditions, and excluding participants who performed poorly did not affect the pattern of findings. Hence, these additional results support the robustness of our main findings.

Discussion

In this experiment, we tested whether factors known to improve face recognition would have similar benefits on prospective person memory. In line with previous findings, both increased variability of encoded instances of each target (Ritchie & Burton, 2017) and associating identities with names (Schwartz & Yovel, 2016) improved PPM performance independently.

The variability effect is consistent with the encoding variability principle (Dempster, 1987; Young & Bellezza, 1982). Given that the probability of retrieval is related to the degree of match between the original encoding context and subsequent retrieval contexts (Tulving, 1974), using multiple high variability photographs results in a higher chance that, in at least one of those events, the target looked similar to how they look at test. The ‘nearest neighbour classifier’ argument in the automatic face recognition literature (Xu, Zhu, Chen, & Pan, 2013) also draws on the relative similarity of encoding and test images. Specifically, one of the three images in the high variability condition will always be a closer match to the test image (i.e., the ‘nearest neighbour’ in this set), but there is no guarantee that the images in the low variability condition will contain this nearest neighbour. To gain a better understanding of these contingencies, future studies should more systematically assess the similarity between images used during encoding and testing, in both low and high variability conditions.

The positive effect of increased variability could also be a result of the development of view-independent representations of target faces (Biederman & Gerhardstein, 1993; Bruce & Young, 1986). Based on this principle, knowing how an individual looks in different contexts results in a more accurate representation of their face, accounting for intra-individual variability (Ritchie & Burton, 2017). Learning how a target might appear in different contexts (i.e., from high variability images) is particularly relevant for PPM tasks because the targets are encountered incidentally during another, attention-demanding task. When investigating PM tasks, Nowinsky and Dismukes (2005) showed that contextual cues available at both encoding and retrieval increase the success of PM, and that highly associated or typical targets support retrieval better. However, identifying the target during a PPM task is considerably more ambiguous than in standard word-based PM experiments. In most PM research, participants see the same cue (e.g., the word 'duck') in the same form during both encoding and testing. In PPM research, participants are presented with different images of a target face during testing than at encoding. Therefore, the encoded information needs to be enough not only to match, but also to prompt retrieval when the person is encountered.

Overall, participants reported more sightings when high variability photographs were used. However, the hit and false alarm rates suggest that increasing photo variability only increased accurate sightings, not inaccurate ones. This effect of variability on response bias therefore, does not support the hypothesis that participants overestimate the number of targets but rather supports the idea that using high variability images results in a higher probability that the test picture will match participants' encoding of the target (i.e., increased hits) without affecting the false alarm rates. The beneficial effect of high variability images could in principle also be due to participants spending more time studying these images than the low variability images. Although we did not examine at the effect of encoding time on PPM, differences in study-duration could mediate the effect of variability. Future studies should assess how encoding time differs based on the images used and investigate if variability is still beneficial when encoding time is reliably controlled for.

Associating each target's images with a name also resulted in better discriminability. Prospective person memory partly depends on the ability to switch attention from the ongoing task to the PPM task when a target is encountered. According to McDaniel and Einstein (2000), this is based on an automatic associative-memory system, which depends on how well the memory cue (i.e., target individual) is processed during encoding. Providing names results in a deeper encoding of each individual and therefore, in a stronger association with the prospective memory intention (i.e., to click on their image if encountered during the main task). Our results support Schwartz and Yovel's (2019a) argument that names provide an additional layer of information about an individual, changing that face's representation from a percept (i.e., an image) to a socially relevant concept (i.e., an identity).

We did not find a significant interaction between the effect of variability and the effect of names. In line with the individuation effect, we hypothesised that names would help participants connect the different images of each target, particularly if the images are highly variable. In contrast, we predicted no effect of names on low variability images because these should be similar enough for participants to realise that they are of a single person even when no names are presented. The absence of the predicted interaction might have been because either the low variability images were harder to connect than anticipated or the high variability images were easier to connect than anticipated. Note that, if images were presented simultaneously, associating names would not be required for connecting them. Therefore, future research should further differentiate between the potential mechanisms of the beneficial effect of name associations by presenting multiple images simultaneously or sequentially, and with or without names.

When interpreting our results, it is important to consider the amount of variability included in our stimuli. We compared highly variable images with less variable images. Although we used the 'low' and 'high' variability labels to describe our conditions, these are relative and do not denote absolute levels of variability. Quite possibly, image variability can be even higher than in our present 'high variability' condition (e.g., if the person had aged between photos). As a consequence, we cannot be sure that high variability images will always improve PPM; there may be a point beyond which the beneficial effect diminishes and performance decreases. Assessing the amount of variability beneficial for PPM would be a worthy topic for future research.

A general limitation of our study (which also applies to other PPM studies; Lampinen, Peters, & Gier, 2012; Sweeney & Lampinen; 2012) was that we did not have any filler task in between encoding of target faces and the testing phase. Therefore, it is possible that at least some of our participants constantly maintained the PM intention in conscious awareness, which involves vigilance processes (Brandimonte et al., 2001). It is widely accepted that vigilance plays a critical role in PM tasks (Brandimonte et al., 2001; Maylor, 1996; Rose et al., 2010). However, whilst vigilance refers to the maintenance of a single goal, PM tasks require alternating between the performance of two tasks (Rose et al., 2010). Although our instructions followed the common PM procedure of embedding the PM task in the ongoing task, future studies should systematically vary the amount of time between encoding and testing, the activities participants engage with during this time, as well as the ratio of targets to non-targets. This could highlight how attentional resources and vigilance influence PPM performance.

The findings of this experimental study may have implications outside the lab. In missing person cases, for example, the public might encounter multiple appeals for the same individual sporadically over a certain period of time. Intuitively, some might think that being consistent and showing the same image of the missing person will make it easier for the public to remember them. However, our findings suggest that using multiple pictures from different contexts results in an increased number of correct sightings. Including names of the target individuals also improved performance. Extrapolating from this, including other types of socially relevant information (e.g., occupation, hobbies) may also be helpful; it might affect how the public processes the appeal and, consequently, lead to improved PPM accuracy. Future research should further explore the effects of providing more contextual and social information on PPM performance. In practice, most missing person appeals include the name of the person. In the current experiment, increasing the within-person variability shown during encoding (e.g., in the missing person appeal) increased performance more than including names. However, varying the type of photographs used is rarely systematically implemented in applied contexts and should, based on the current research, receive more attention from those who create and distribute missing person appeals.

Author Contributions

The first three authors developed the study concept and designed the experiment. The first author created experimental stimuli, organized data collection, performed all data analyses and drafted the manuscript. Second, third, and fourth authors provided feedback on the research and reviews on the research paper.

References

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(6), 1162–1182. https://doi.org/10.1037/0096-1523.19.6.1162

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, *103*(4), 751–757. https://doi.org/10.1037/h0037190

Brandimonte, M. A., Ferrante, D., Feresin, C., & Delbello, R. (2001). Dissociating prospective memory from vigilance processes. *Psicológica*, 22(1), 97-113.

Bruce, V., & Young, A. (1986). Understanding face recognition. British Journal of *Psychology*, *77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Courtois, M. R., & Mueller, J. H. (1979). Processing multiple physical features in facial recognition. *Bulletin of the Psychonomic Society*, *14*(1), 74–76. https://doi.org/10.3758/BF03329404

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, *79*(2), 162–170. https://doi.org/10.1037/0022-0663.79.2.162

Einstein, G. O., & McDaniel, M. A. (2005). Prospective Memory. *Current Directions in Psychological Science*, *14*(6), 286–290. https://doi.org/10.1111/j.0963-7214.2005.00382.x

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007) Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. *University of Massachusetts, Amherst, Technical Report 07-49,*

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. https://doi.org/10.1037/a0020463

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*(5), 577–596. https://doi.org/10.1080/09658210902976969

Kramer, R. S. S., Jenkins, R., Young, A. W., & Burton, A. M. (2017). Natural variability is essential to learning new faces. *Visual Cognition*, *25*(4–6), 470–476. https://doi.org/10.1080/13506285.2016.1242522

Lampinen, J. M., Arnal, J. D., & Hicks, J. L. (2009). Prospective person memory. In M.Kelley (Ed.), *Applied memory* (pp. 167–184). Hauppauge, NY: Nova.

Lampinen, J. M., Curry, C. R., & Erickson, W. B. (2016). Prospective Person Memory: The Role of Self-Efficacy, Personal Interaction, and Multiple Images in Recognition of

Wanted Persons. *Journal of Police and Criminal Psychology*, *31*(1), 59–70. https://doi.org/10.1007/s11896-015-9164-7

Lampinen, J. M., & Moore, K. N. (2016). Prospective Person Memory in the Search for Missing Persons. In *Handbook of Missing Persons* (pp. 145–162). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40199-7_11

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide.* Psychology press.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185

Maylor, E. A. (1996). Does prospective memory decline with age? In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), Prospective memory: Theory and applications (p. 173–197). Lawrence Erlbaum Associates Publishers.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: a multiprocess framework. *Applied Cognitive Psychology*, *14*(7), S127–S144. https://doi.org/10.1002/acp.775

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-Specific Perceptual Discrimination Improvement Following Short Individuation Training With Faces. *Cognitive Science*, *35*(2), 330–347. https://doi.org/10.1111/j.1551-6709.2010.01148.x

Moore, K. N., Lampinen, J. M., & Provenzano, A. C. (2016). The Role of Temporal and Spatial Information Cues in Locating Missing Persons. *Applied Cognitive Psychology*, *30*(4), 514–525. https://doi.org/10.1002/acp.3242

Nowinski, J. L., & Dismukes, K. (2005). Effects of ongoing task context and target typicality on prospective memory performance: The importance of associative cueing. *Memory*, 13(6), 649-657

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. Quarterly Journal of Experimental Psychology, 70(5), 897–905. https://doi.org/10.1080/17470218.2015.1136656

Rose, N. S., Rendell, P. G., McDaniel, M. A., Aberle, I., & Kliegel, M. (2010). Age and individual differences in prospective memory during a" Virtual Week": The roles of working memory, vigilance, task regularity, and cue focality. *Psychology and aging*, 25(3), 595

Russo, R., Mammarella, N., & Avons, S. E. (2002). Toward a unified account of spacing effects in explicit cued-memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(5), 819–829. https://doi.org/10.1037/0278-7393.28.5.819

Schwartz, L., & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*(11), 1493–1511. https://doi.org/10.1037/xge0000220

Schwartz, L., & Yovel, G. (2019). Learning faces as concepts rather than percepts improves face recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1733–1747. https://doi.org/10.1037/xlm0000673

Schwartz, L., & Yovel, G. (2019). Independent contribution of perceptual experience and social cognition to face recognition. *Cognition*, *183*, 131–138. https://doi.org/10.1016/j.cognition.2018.11.003

Solymosi, R., Petcu, O., & Wilkinson, J. (2020, January 27). Exploring public engagement with missing person appeals on Twitter. https://doi.org/10.31235/osf.io/wugxs

Sweeney, L. N., & Lampinen, J. M. (2012). The effect of presenting multiple images on prospective and retrospective person memory for missing children. *Journal of Applied Research in Memory and Cognition*, *1*(4), 235–241. https://doi.org/10.1016/j.jarmac.2012.08.001

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 122–131. https://doi.org/10.3758/CABN.9.1.122

Tulving, E. (1974). Recall and recognition of semantically encoded words. *Journal of Experimental Psychology*, *102*(5), 778–787. https://doi.org/10.1037/h0036383

Wogalter, M. S., Jarrard, S. W., & Cayard, J. A. (1991). Massed versus distributed exposure and imaging of faces: Changing the test view. *Current Psychology*, *10*(4), 281–288. https://doi.org/10.1007/BF02686900

Xu, Y., Zhu, Q., Chen, Y., & Pan, J. S. (2013). An improvement to the nearest neighbor classifier and face recognition experiments. *Int. J. Innov. Comput. Inf. Control*, 9(2), 543-554.

Young, D. R., & Bellezza, F. S. (1982). Encoding variability, memory organization, and the repetition effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(6), 545–559. https://doi.org/10.1037/0278-7393.8.6.545

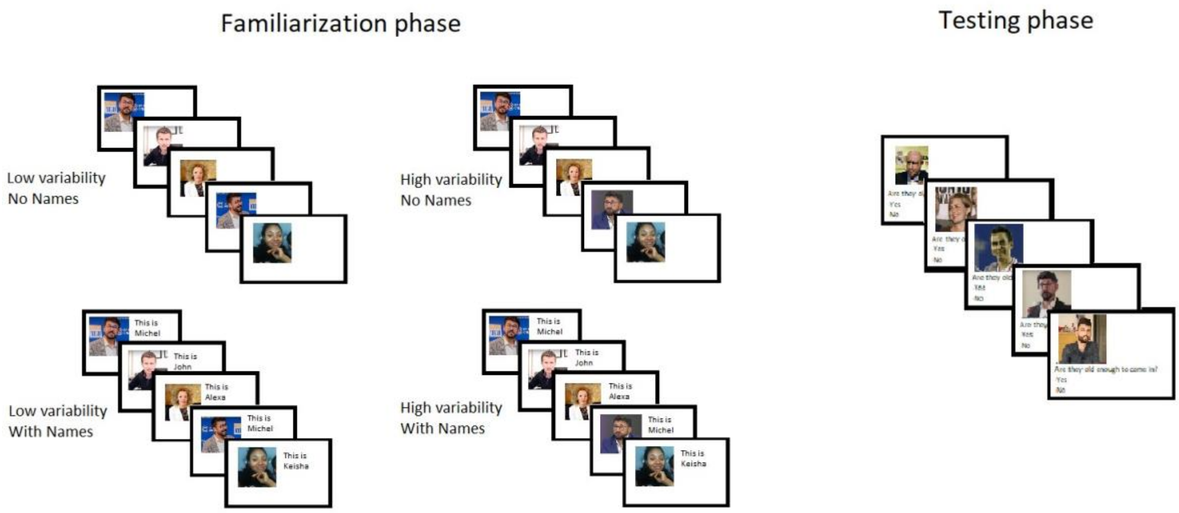

Figure 1. Examples of low and high variability images, with and without names presented during the familiarization phase and examples of target and nontarget images presented during the testing phase.

Author Statement

S. Juncu, H. Blank, and R. Fitzgerald developed the study concept and designed the experiment. S. Juncu created experimental stimuli, organized data collection, performed all data analyses and drafted the manuscript. H. Blank, R. Fitzgerald and L. Hope provided feedback on the research and reviews on the research paper.