

**Title: AN OWN-AGE BIAS IN MIXED- AND PURE-LIST PRESENTATIONS: NO  
EVIDENCE FOR THE SOCIAL-COGNITIVE ACCOUNT**

Short title: *AN OAB IN MIXED- AND PURE- LIST PRESENTATIONS*

Sophie L. Cronin\*<sup>1</sup>, Belinda M. Craig<sup>2</sup>, & Ottmar V. Lipp<sup>1</sup>

<sup>1</sup>School of Psychology, Curtin University, Bentley, WA, 6102, Australia

<sup>2</sup>School of Psychology, University of New England, Armidale, NSW, 2351, Australia

\* Corresponding author information: Sophie Cronin, School of Psychology, Curtin University, Bentley, WA, 6102, Australia, (email: [sophie.l.cronin@postgrad.curtin.edu.au](mailto:sophie.l.cronin@postgrad.curtin.edu.au)).

**Abstract:**

The own-age bias (OAB) is suggested to be caused by perceptual-expertise and/or social-cognitive mechanisms. Bryce and Dodson (2013, Exp 2) provided support for the social-cognitive account, demonstrating an OAB for participants who encountered a mixed-list of own- and other-age faces, but not for participants who encountered a pure-list of only own- or other-age faces. They proposed that own-age/other-age categorisation, and the resulting OAB, only emerge when age is made salient in the mixed-list condition. Our study aimed to replicate this finding using methods typically used to investigate the OAB to examine their robustness and contribution to our understanding of how the OAB forms. Across three experiments that removed theoretically unimportant components of the original paradigm, varied face sex, and included background scenes, the OAB emerged under both mixed-list and pure-list conditions. These results are more consistent with a perceptual-expertise than social-cognitive account of the OAB, but may suggest that manipulating age salience using mixed-list and pure-list presentations is not sufficient to alter categorisation processes.

This is the peer reviewed version of the following article:

Cronin, S.L. and Craig, B.M. and Lipp, O.V. 2019. An own-age bias in mixed- and pure-list presentations: No evidence for the social-cognitive account. *British Journal of Psychology*, which has been published in final form at <https://doi.org/10.1111/bjop.12435>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

**Keywords:**

own-age bias, face recognition, face memory, social-cognition, perceptual-expertise

**Data availability statement:**

The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgements:**

This research was supported by an Australian Government Research Training Program (RTP) Scholarship awarded to SLC and an Australian Research Council Discovery Project (DP150101540) awarded to OVL. We thank Dr Welber Marinovic for his assistance in conducting the Bayesian statistical analyses.

Our ability to recognise people we have encountered before is an important social skill. The ability to correctly recognise unfamiliar faces is, however, prone to error and subject to biases (Hancock, Bruce, & Burton, 2000; Hugenberg, Wilson, See, & Young, 2013). One such bias is the own-age bias (OAB) which is characterised by better recognition memory for own-age relative to other-age faces (Wiese, Komes, & Schweinberger, 2013). The OAB occurs to varying degrees across the lifespan, with the strongest bias seen in young adult observers when the other-age faces are older adults (see Wiese, Komes, et al., 2013 for a review, and Rhodes & Anastasi, 2012 for a meta-analysis). Attempts to explain the cause of the OAB have largely drawn on social-cognitive and perceptual-expertise mechanisms proposed in the related own-race bias field (ORB; Wiese, Komes, et al., 2013).

The social-cognitive account holds that ‘own-group’ biases, like the ORB and OAB, are caused by differential evaluations of, and attention to, own- versus other-group faces. When we view a face we make judgements and evaluations that are suggested to influence the way we remember that face (Rodin, 1987; Sporer, 2001). Judgments relating to in-group/out-group membership, relevance, and motivation can bias us to merely categorise the face and encode category level information, or further individuate it and encode identity level information (Hugenberg, Young, Bernstein, & Sacco, 2010; Levin, 2000). In the case of the OAB, age is a social cue that prompts evaluations about group membership. Within this account, own-age faces are processed with a focus on individuating information, making their identities easier to recognise, and other-age faces are processed with a focus on category information, making their identities more difficult to recognise (Hugenberg et al., 2010; Levin, 2000).

The perceptual-expertise account holds that own-group biases occur due to differential experience with in-group and out-group people (Wiese, Komes, et al., 2013). Generally, we have more exposure to faces from our in-group compared to out-groups, and as

a result we are afforded more opportunities to practice recognising these people, and to gain expertise in in-group face processing. In the case of the OAB, this expertise is said to improve our perceptual processing of own-age faces, enhancing encoding of information useful for recognising identities, and organising our mental representations of identities such that differences are maximised and discrimination is facilitated between different own-age faces (Wiese, Komes, et al., 2013; Valentine, 1991). Other-age faces alternately, do not benefit from expertise and the information encoded from a face is less optimised for identity recognition (Wiese, Komes, et al., 2013). The mental representations of these faces will also be stored in a system optimised to represent differences between own-age faces, resulting in difficulties distinguishing between similar other-age faces (Valentine, 1991). Considering that what is own- and other-age changes as people age, perceptual-expertise accounts often suggest that as new expertise is accumulated, former expertise tends to wane or become dormant and mental representations are adjusted to best represent what is now the new own-age group (see Hills, 2012 for a discussion).

The Categorisation Individuation Model (CIM) proposed by Hugenberg et al. (2010) integrates components of both the perceptual-expertise and social-cognitive accounts. It suggests that categorisation, motivation, and individuation experience can all influence our encoding of faces. The two former components are social-cognitive factors and the latter is a perceptual-expertise factor. This model has been used to describe how recognition can be influenced by both perceptual-expertise and social-cognition, and how these processes interact. An example of this can be seen in Young and Hugenberg's (2012) study on the influence of individuation instructions on the ORB. They hypothesised that instructing participants to pay attention to how other-race faces differed from one another would increase their motivation to remember them and in turn they would become less biased. Results supported this prediction. However, Young and Hugenberg also measured interracial contact

and found that the instruction effect was only present when participants had high levels of interracial experience. They argued that bias reduction caused by motivation to remember other-race faces (a social-cognitive factor) was only effective when there was pre-existing individuation experience (a perceptual-expertise factor) to draw on. This study and the CIM highlight that both perceptual-expertise and social-cognitive factors together can be important to face identity recognition.

Research that has specifically evaluated the mechanism underlying the OAB has yielded mixed results, with evidence emerging for both social-cognitive and perceptual expertise mechanisms. Most of the research has evaluated the perceptual-expertise account. These studies have demonstrated that groups of people who have higher levels of contact with other-age people (e.g. trainee teachers, preschool teachers, geriatric nurses) tend to have a smaller OAB compared to the average person who has low levels of contact with people from other age groups (Harrison & Hole, 2009; Macchi Cassia, Picozzi, Kuefner, & Casati, 2009; Wiese, Wolff, Steffens, & Schweinberger, 2013), and that other-age contact levels correlate with OAB such that as contact increases, the OAB decreases (Ebner & Johnson, 2009; He, Ebner, & Johnson, 2011). Additionally, more efficient holistic processing, a hallmark of expertise, can be seen for own-age relative to other-age faces and has been shown to vary with reported contact (Kuefner, Macchi Cassia, Picozzi & Bricolo, 2008; Macchi Cassia, Picozzi, Kuefner, & Casati, 2009). However, it is unclear if contact or expertise directly cause the OAB. It is likely that contact also co-varies with other more social factors such as positive feelings towards the relevant age out-groups which may explain the results.

Research evaluating the social-cognitive account has been much more limited. Most of the research employs manipulations of individuation, often adapted from studies in the ORB field, in an attempt to improve recognition performance of other-age faces. Individuation has been manipulated using instructions, tasks and stimuli, and these studies

have produced mixed support for the role of social-cognition in the OAB. In one study, participants were instructed to individuate, by paying attention to how the other-age faces differed from each other (Craig & Thorne, 2019). Although these instructions have been shown to reduce the ORB (Young & Hugenberg, 2012), they did not reduce the OAB. Another study had participants view faces in pairs and tasked them with deciding if the photos were of the same person (Proietti, Laurence, Matthews, Zhou, & Mondloch, 2018). The authors suggested that this task would necessitate individuation of the faces and found that the OAB was eliminated in a subsequent recognition test.

Studies using emotional expressions have also produced mixed findings. Emotional expressions convey more important and interesting information relative to neutral faces and are hypothesised to improve face recognition memory (Ackerman et al., 2006; D'Argembeau, Van der Linden, Comblain & Etienne, 2003). Where out-group faces may be the subject of cognitive disregard and processed at a categorical level, emotional information can act to increase their relevance to the observer. This in turn increases motivation to remember the face and prompts processing of the face at the individual level. One study examining the effect of emotional expressions on the OAB presented participants with faces posing neutral and emotional expressions (angry, sad or happy; Cronin, Craig & Lipp, 2018). Emotional expressions reduced the OAB seen for faces with neutral expressions by improving the recognition of other-age faces. However, similar experiments have found no effect of emotional expression on the OAB (Ebner & Johnson, 2009), or only marginal effects (Denkinger & Kinn, 2018).

Only one study that we are aware of employs a manipulation targeting categorisation to assess the social-cognitive account of the OAB. Bryce and Dodson (2013, Exp 2) aimed to manipulate age group salience by presenting young adult participants either with a mixed stimulus set ("mixed-list") of young and older adult faces, or a pure set ("pure-list") of only

young or older adult faces. They found that the OAB emerged in the mixed-list but not the pure-list condition. This was due to worse recognition performance for young faces in the pure-list compared to the mixed-list condition. They argued that when age was made salient in the mixed-list condition, own-age/other-age categorisation was induced, causing own-age face processing to be enhanced and facilitating an OAB. In the pure-list condition, where age was not salient, no OAB was observed. Young faces were not categorised as own-age, and as such did not benefit from individuation encoding that would produce an OAB.

A similar approach has been taken in the ORB field, also finding support for the role of social-cognition albeit through a different pattern of results. Young, Hugenberg, Bernstein and Sacco (2009) explored whether increasing race category salience would prompt changes in recognition performance for white own-race faces across two studies. Using a task order manipulation, participants completed two old/new recognition tasks, one with only own-race white faces, and one with only other-race black faces. They found that recognition performance was reduced for own-race faces when this block followed after the other-race face block, while other-race face recognition performance did not differ between conditions. They replicated this change in own-race recognition, with a single white or black face preceding the own-age memory block. They argued that when other-race faces were seen first, they highlighted race as a category, drawing more attention to category-diagnostic facial features in the own-race faces that followed, reducing recognition performance for those faces. This stands in contrast to Bryce and Dodson's (2013) suggestion that category salience should improve own-group recognition through highlighting the relevance of own-group faces.

Bryce and Dodson's (2013) study is one of the few pieces of evidence published that evaluates effects of social cognition in the OAB and is also the only one that attempts to manipulate categorisation processes. In comparison to the Young et al. (2009) study that

examines a similar question for the ORB, the implications for theory are different. While both studies suggest evidence for social-cognitive processes, the type of process proposed differs. Bryce and Dodson suggest that age category salience highlights in-group membership of own-age faces, providing a signal of personal relevance that motivates more individuation style processing of these faces and improves recognition performance. In contrast, Young et al. (2009) suggest that age category salience enhances categorisation of own-race faces, increasing attention to category-diagnostic features of the faces, which in turn reduces recognition performance, though not to the level of the other-race faces. The implication of Bryce and Dodson's study, beyond the role of social cognition in the OAB, is that the different aspects of social cognition may be stronger in the case of age than race, and vice versa. This is consistent with other more recent evidence examining identity recognition and perceived distinctiveness for faces varying in both age and race which suggested some independence between the OAB and ORB, and proposed that the OAB may be based on differences in motivation to deeply encode while the ORB may be based on differences in individuation (Mukudi & Hills, 2019).

There are many components of the Bryce and Dodson (2013) design that differ from typical old/new face recognition studies in the own-group recognition bias field. A basic old/new face recognition paradigm will typically include one encoding phase where participants view faces one at a time. This is followed by a delay period often including an unrelated filler task. Finally, participants complete a recognition test where previously viewed and new faces are presented one at a time for participants to indicate whether they were seen before or not (e.g. Ackerman et al., 2006, Anastasi & Rhodes, 2005; Hills and Lewis, 2011; Young, Bernstein & Hugenberg, 2010). In Bryce and Dodson's study, young adult participants (age range unspecified) completed three encoding-test sessions with 5 min filler tasks separating the sessions rather than the encoding and test phases. In each encoding



phase 24 faces were viewed one at a time for 5s on a variety of background scenes. Participants were asked to remember the face-background pairs. Two additional filler faces that were not later tested were added to the beginning and end of this encoding phase. Immediately afterward, a recognition test was conducted where 24 studied faces were viewed, and an additional 8 new faces were presented in a randomised order, one at a time. Participants indicated if they had seen the face before and rated their confidence in this judgement on a six-point scale. If they reported seeing the face, they were also asked to identify the background it was presented on. Participants either saw a pure-list of young faces or old faces, or a mixed-list of both young and old faces throughout the experiment. Faces were of male and female Caucasians with neutral expressions and background stimuli were a hospital room, basketball court and plain white. Backgrounds were chosen to be age-stereotypical (young-basketball, old-hospital) and hypothesised to differentially influence source memory for young and older adult faces (though they did not). Some of these components were included to test related questions (e.g. background scenes were included to test source memory) and others likely to solve methodological issues (e.g. multiple testing blocks to give sufficient trial numbers to test source memory questions while not overloading memory). Given these idiosyncrasies and complexities in design, a replication using a more typical old/new recognition paradigm would provide further evidence to support the robustness of Bryce and Dodson's results beyond the specifics of their design, and support their generalisability and implications for our understanding of the mechanism underlying the OAB.

The aim of the current series of experiments was to replicate Bryce and Dodson's (2013) finding that the OAB is present under mixed-list but not pure-list conditions, and to examine the robustness of the finding in an alternate paradigm where theoretically unimportant components are removed. Theoretically unimportant components are

components that do not affect the conditions of encoding where the theories of the OAB suggest the OAB is produced. In three experiments, young adult participants were presented with one set of faces to remember followed by a filler task and then a recognition test. Participants viewed a mixed-list of faces (young and older adults), or a pure-list of faces (young or older adults). Relative recognition performance between young and older adult faces was examined. Experiment 1 tested the effect using only male Caucasian faces, Experiment 2 explored the influence of face sex by including male and female faces, and Experiment 3 explored the influence of stereotypical backgrounds. Based on Bryce and Dodson's finding we expect to find an OAB in the mixed- but not the pure-list condition. This would support the notion of social-cognitive processes in the OAB whereby the OAB emerges only when age categorisation is prompted by viewing faces from more than one age category. Alternately, if a perceptual-expertise mechanism underlies the OAB, we expect to see similar OABs in both list conditions.

## **Experiment 1**

### **Method**

**Participants.** Given the modified ANOVA analysis (see Data Analysis below) used in the Bryce and Dodson (2013, Exp 2) study, a power analysis using G\*Power (Faul, Erdfelder, Lang, & Buchner, 2007) was conducted using an effect size estimate from a previous study investigating the OAB (paired samples t-test:  $d_{av} = 0.64$ , Cronin et al., 2018). This suggested that 32 participants would be required to have an 80% chance of finding the OAB in a mixed-list design. Considering this, along with Bryce and Dodson's (2013, Exp 2) sample of 24 participants per list-type, counterbalancing needs, and potential participant attrition, we aimed to recruit 64 participants per list-type.

One-hundred and sixty eight participants' data were analysed (94 male, 73 female, 1 undisclosed;  $M_{age}=26.33$ ,  $SD_{age}=3.07$ , range=18-31). Demographic requirements for this experiment were that participants be aged between 18 and 31 years old and Caucasian. This was so that participants' age and race matched the characteristics of the "young" stimuli used in the experiments. Participants were sourced from Amazon Mechanical Turk and compensated USD2.80 for their time. There were 56 participants in each list-type group (mixed, pure-young, and pure-old). Additional participants were removed prior to analyses because they did not fit age and race demographic requirements (10 participants), completed the experimental sequence incorrectly (4 participants), or needed to be removed so that group numbers were equal and the required analysis method could be conducted (15 participants - 11 from the pure-young condition, 4 from the pure-old condition; see Data Analysis for further detail). Ethics approval was obtained from an Australian university.

**Stimuli.** Ninety-six images of Caucasian male young and older adult faces with neutral expressions were used in this experiment. Half were sourced from the CAL/PAL database (Minear & Park, 2004; the database used in the Bryce & Dodson, 2013 study) and half from the FACES database (Ebner, Riediger, & Lindenberger, 2010). Young adult posers were between the ages of 18 and 31 years, and older adult posers were 69 years and older. A total of 48 young adult and 48 older adult posers were used. Images were in colour and  $335 \times 419$  pixels in size. To maintain consistency across stimulus sets, backgrounds were removed with faces cropped around the face and hair outline.

**Procedure.** A previously used intentional learning old/new recognition paradigm was employed (Cronin et al., 2018). Participants were randomly assigned to one of three stimulus list-types; mixed, pure-young or pure-old. In the pure conditions, participants only saw faces of one age group throughout the experiment (young or old, respectively). In the mixed

condition age was varied within participants with half of the faces viewed at each stage of the experiment young, and half old.

Participants first viewed 24 faces one at a time in an exposure phase, and were instructed to “pay attention to the faces as you will be asked to remember them later”. Those in the mixed-list condition saw 12 young and 12 older adult faces. Those in the pure-young condition saw 24 young faces, and those in the pure-old condition saw 24 older adult faces. In line with previous experiments investigating the OAB with online participants (Cronin et al., 2018), a secondary response time (RT) task was included to ensure participants were paying attention to the faces. To facilitate this, half of the faces were presented for 1000ms and half for 1500ms, with an inter-stimulus interval of 1500ms. The presentation duration of each poser was counterbalanced between-participants and in the mixed-list condition an equal number of young and older adult faces were presented for each duration. Participants were instructed to press the spacebar when the face disappeared from the screen as quickly as possible. Text prompts to “remember the face” and “press spacebar after it disappears” were presented at the bottom of the screen on each trial.

Next, participants completed a filler task. This was an irrelevant task designed to remove the faces from working memory and allow the recognition test to be a test of long term memory. The task involved eight word puzzles in which participants had 30 seconds to generate the longest word from a grid of nine letters in which the centre letter had to be included. Including reading of instructions, word generating, and responding, the task took approximately five minutes to complete.

Following the filler task, participants completed a recognition test. They were presented with 48 images, 24 that they saw at the beginning of the experiment and 24 new faces. Posers that were studied in exposure, and posers that were new at test were

counterbalanced between-participants. Participants saw faces of the same ages in the exposure and test phases (i.e. participants in the mixed-list condition saw young and old, participants in the pure-young condition saw only young, and participants in the pure-old condition saw only old). Participants were instructed to press ‘e’ if they had “seen” the face before, and ‘i’ if they had “not seen” the face before (response mapping was counterbalanced between-participants). If they indicated they had seen the face before, a follow-up remember/know/guess question was included to probe memory strength. Participants were instructed to respond “remember” if they recalled seeing the face, “know” if they could not recall it specifically but it was familiar, and “guess” if they were guessing (Gardiner & Richardson-Klavehn, 2000). Although this measure in the Cronin et al. (2018) paradigm is not a necessary feature as it appears post-encoding, we retained it to keep this alternate paradigm consistent with its past use.

At the end of the experiment a brief survey collected demographics (age, race, sex), and measures of age group contact and age group social identification. The contact measure asks two questions per age group: *“How often do you have personal (i.e., face-to-face) contact with young adults/older adults (approx. between 18 to 30 years of age/approx. 65 years of age and older)?”* and *“How often do you have other types of contact (e.g., phone, e-mail, letter) with young adults/older adults (approx. between 18 to 30 years of age/approx. 65 years of age and older)?”*. Responses were taken on a 1-8 scale from *“daily”* to *“less than once a year”* and averaged to form a single contact score in line with previous research (Ebner & Johnson, 2009). The identification measure asks one question per age group: *“I identify with older adults (65 years+)”* and *“I identify with young adults (18-30 years)”* on a 1-7 Likert scale from *“strongly disagree”* to *“strongly agree”* (Postmes, Haslam, & Jans, 2013; Reysen, Katzarska-Miller, Nesbit, & Pierce, 2013).

**Data analysis.** The design of this study includes an independent variable (face age) that is manipulated both within and between participants. In order to analyse such data, Erlebacher's (1977) modified ANOVA method was employed. This method was also used by Bryce and Dodson (2013) and is a method that was created specifically to test if the effect of an independent variable changes depending on whether the design is within- or between- participants. Error degrees of freedom in the method are estimated for each effect using Satterthwaite's (1946) method. For more details refer to Erlebacher (1977) and Satterthwaite (1946). The analysis was conducted using R (version 3.4.3; R Core Team, 2017) and code from Merritt, Cook, and Wang (2014). In order for the analysis to be completed, all groups (mixed, pure-young, and pure-old) need to have an equal number of participants. As mentioned above, some participants' data was removed prior to analysis (those who had participated last were removed first).

To supplement the main analysis, data were analysed using Bayesian generalised linear mixed models and the *brm* function from the BRMS R package (Bürkner, 2018) with model likelihoods compared. In this analysis, scores were modelled as a normally distributed variable using a Gaussian family function, the priors were set as a Cauchy distribution using a non-informative prior (0.707), and the number of iterations was set to 4000.

Corrected recognition scores ( $P_r$ ), hit rates, false alarm rates, and response bias ( $B_r$ ) were calculated as dependent variables (Snodgrass & Corwin, 1988).  $P_r$  is calculated by subtracting the false alarm rate from the hit rate.  $B_r$  is calculated by dividing the false alarm rate by 1 minus  $P_r$ . Scores above 0.5 indicate a tendency to respond "seen" (liberal response bias) and scores below 0.5 indicate a tendency to respond "not seen" (conservative response bias). Hit rates represent the proportion of test trials in which the participant correctly responded "seen" to a face that was seen before. False alarm rates represent the proportion of test trials in which the participant incorrectly responded "seen" to a face that had not been

seen before. Loglinear adjustments were applied to  $P_r$  and  $B_r$  calculations, in line with recommendations by Snodgrass and Corwin (1988). This is such that before hit and false alarm rates are calculated 0.5 is added to the raw counts of hits and of false alarms, and 1 is added to the number of trials with seen faces and to the number of trials with new faces.

Where hit and false alarm rates are reported below, these are the unadjusted rates.

Sensitivity ( $d'$ ) was also calculated and produced the same pattern of results as  $P_r$  throughout the experiments in this paper and is reported in the Supplement.

Remember/know/guess analyses did not contribute meaningfully to the interpretation of results of these experiments and can be found in the Supplement.

## Results

One-way ANOVAs were conducted on the group contact data (young contact and old contact) with a between-groups factor of list-type (mixed, pure-young, and pure-old). Results indicate that the list-type groups did not differ on amount of contact with young adults ( $F(2, 164)=1.47, p=.234, \eta^2=.02$ ) but there was a marginal difference in the amount of contact with older adults ( $F(2, 163)=2.97, p=.054, \eta^2=.04$ ). This marginal difference emerged as those in the pure-young group reported slightly less contact with older adults than those in the pure-old group ( $t(108)=2.29, p=.024, d_s=0.44$ ). Given that those in the pure-young group did not view any older adult faces we do not believe this to be problematic. Paired samples t-tests also confirm that our sample had higher levels of contact with young than older adults ( $t(165)=16.11, p<.001, d_{av}=1.74$ ) and greater social identification with young than older adults ( $t(165)=24.92, p<.001, d_{av}=3.08$ ).

An analysis of the  $P_r$  data using Erlebacher's (1977) method with a face age (young, old) factor and a design (mixed [within-participants], pure [between-participants]) factor was conducted (see Figure 1). There was a main effect of face age such that the OAB was present

overall ( $F(1, 160)=12.13, p<.001, \eta^2=.04$ ). There was, however, no effect of design ( $F(1, 119)=0.95, p=.333, \eta^2=.01$ ), nor an interaction between face age and design ( $F(1, 160)=0.56, p=.457, \eta^2<.01$ ). Bayesian analyses were conducted to compare the likelihood of our pattern of results (main effect of age only) to the pattern Bryce and Dodson (2013) found (main effect of age and an interaction between age and design). We found that a model containing only the age effect was 410.75 times more likely than a model containing the age and interaction effects given the observed data.

Analysis of the hit rates produced no significant effects of face age ( $F(1, 146)<0.01, p=.955, \eta^2<.01$ ), design ( $F(1, 128)=1.18, p=.280, \eta^2=.01$ ), nor an interaction between the two ( $F(1, 146)=0.03, p=.865, \eta^2<.01$ ). An analysis of the false alarm rates produced a main effect of face age ( $F(1, 136)=11.98, p<.001, \eta^2=.05$ ) such that fewer false alarms were made to young adult faces than older adult faces (see Table 1). There was no main effect of design ( $F(1, 131)=0.03, p=.858, \eta^2<.01$ ) nor an interaction between face age and design ( $F(1, 136)=0.79, p=.377, \eta^2<.01$ ) for false alarms. Analysis of  $B_r$  produced a main effect of age ( $F(1, 137)=4.29, p=.040, \eta^2=.02$ ) such that responding was more conservative for young than older adult faces. There was no effect of design ( $F(1, 136)=0.38, p=.538, \eta^2<.01$ ), nor an interaction between age and design ( $F(1, 137)=0.26, p=.609, \eta^2<.01$ ) in the  $B_r$  data.

-----Figure 1-----

-----Table 1-----

## Discussion

These results do not replicate the finding by Bryce and Dodson (2013) that the OAB varies as a function of stimulus set composition. We observed the standard OAB overall, and this did not vary as a function of whether participants saw one age group or two. Bryce and Dodson argue that pure-list conditions should produce comparable recognition performance



for differently aged faces because age is not salient and therefore age categorisation processes are not engaged. If this underlying social-cognitive mechanism of categorisation is responsible for the OAB it may be that some of the differences between our paradigm and Bryce and Dodson's influenced this mechanism and may explain why we find our results. One difference that may have influenced performance is that we did not vary face sex. The faces used in Experiment 1 were restricted to male faces while in the Bryce and Dodson study both male and female faces were used. We used only male faces to remove variation in the potential in-group/out-group cue of face sex. It may be, however, that varying sex allows the salience of age to be reduced in the pure-list conditions because there is another cue to categorise by. To explore this possibility, Experiment 2 included both male and female faces in all conditions. If face sex variation aids in the reduction of age salience and categorisation, consistent with the findings of Bryce and Dodson, we expect to find an OAB in the mixed-, but not in the pure-list condition. Alternately, if a perceptual-expertise mechanism underlies the OAB, we expect to see similar OABs in both list conditions, replicating Experiment 1.

## Experiment 2

### Method

**Participants.** One-hundred and seventy seven participants' data were analysed (92 male, 85 female;  $M_{\text{age}}=26.11$ ,  $SD_{\text{age}}=3.04$ , range=18-31). Participants were sourced from Amazon Mechanical Turk and compensated USD2.80 for their time. There were 59 participants in each list-type group (mixed, pure-young, and pure-old). Additional participants were removed prior to analyses because they did not fit demographic requirements (7 participants), completed the experimental sequence incorrectly (9

participants), or needed to be removed so that group numbers were equal and the required analysis method could be used (9 participants - 5 pure-young, 4 pure-old).

**Stimuli.** The same stimuli as Experiment 1 were included, with an additional ninety-six Caucasian female faces with neutral expressions (48 young and 48 older adult posers) taken from the same databases. Visible jewellery was digitally removed. While the age of faces that participants saw varied based on the list-type they were allocated to, all participants saw both male and female faces throughout the experiment.

**Procedure and data analysis.** As per Experiment 1 with the following differences. Of the faces presented during encoding and during test, half were male and half were female. For participants in the mixed-list condition, there were an equal number of faces of each sex for each age condition. Presentation time counterbalancing in exposure was also such that an equal number of male and female faces were presented for each duration. Given the now larger stimulus set, the subset of faces chosen for each participant was counterbalanced between-participants.

## Results

One-way ANOVAs on the group contact data (young contact and old contact) with a between-groups factor of list-type (pure-young, pure-old, mixed) show that our groups did not differ in the amount of contact reported with young ( $F(2, 173)=1.00, p=.370, \eta^2=.01$ ) or older adults ( $F(2, 172)=1.84, p=.162, \eta^2=.02$ ). Paired-samples t-tests confirmed that our sample had higher contact with young than older adults ( $t(173)=15.64, p<.001, d_{av}=1.65$ ), and identified more with young than older adults ( $t(175)=23.64, p<.001, d_{av}=3.00$ ).

Analysis of the  $P_r$  data using Erlebacher's (1977) method with factors of face age (young, old) and design (mixed, pure) was conducted (see Figure 2). There was a main effect of face age such that the OAB was present overall ( $F(1, 141)=9.83, p=.002, \eta^2=.04$ ). There

was, however, no effect of design ( $F(1, 131)=0.08, p=.774, \eta^2<.01$ ) nor an interaction between face age and design ( $F(1, 141)=0.63, p=.427, \eta^2<.01$ ). Bayesian analyses found that a model containing only the age main effect was 423.18 times more likely than a model containing the age and interaction effects to explain the data. Stimulus sex effects were not significant in the data and are not reported further.

An analysis of the hit rates produced no significant effect of face age ( $F(1, 125)=0.21, p=.650, \eta^2<.01$ ), design ( $F(1, 104)=1.28, p=.261, \eta^2=.01$ ), nor an interaction between the two ( $F(1, 125)=0.63, p=.428, \eta^2<.01$ ). An analysis of the false alarm rates produced a main effect of face age ( $F(1, 149)=16.94, p<.001, \eta^2=.06$ ) such that fewer false alarms were made to young adult faces than older adult faces (see Table 2). There was no main effect of design ( $F(1, 119)=0.23, p=.632, \eta^2<.01$ ), nor an interaction between face age and design ( $F(1, 149)<0.01, p=.961, \eta^2<.01$ ) for false alarms. Analysis of  $B_r$  produced a significant effect of age ( $F(1, 143)=7.58, p=.007, \eta^2=.03$ ) such that responding to young faces was more conservative than to older adult faces. There was no effect of design ( $F(1, 118)=1.87, p=.174, \eta^2<.01$ ), nor an interaction between the age and design ( $F(1, 143)=0.02, p=.895, \eta^2<.01$ ) for the  $B_r$  data.

-----Figure 2-----

-----Table 2-----

## Discussion

The results of this experiment replicate those of Experiment 1 where an OAB was observed, which did not differ between the mixed-list and pure-list conditions. Again, these results are in contrast to Bryce and Dodson's (2013) findings. The inclusion of female faces did not influence the pattern of results. Another feature of Bryce and Dodson's procedure that might influence categorisation processes was the inclusion of a source memory task using age

stereotypical backgrounds. Rather than placing all faces on a neutral background, Bryce and Dodson presented faces on one of three backgrounds; a white rectangle, a basketball court or a hospital room. The first served as a neutral control and the latter two were suggested to be stereotypic of age groups (young and old, respectively). However, the backgrounds may also be stereotypic of sex (male and female, respectively). In the pure-list conditions where age does not vary, the salience of face sex may be enhanced by the sex stereotypes associated with the backgrounds. Additionally, asking participants to remember the faces and the backgrounds they are presented on may influence the way in which faces are encoded. To examine if the inclusion of these background stimuli influences the pattern of results, the source memory task was included in Experiment 3. If the inclusion of stereotypical backgrounds aids in the reduction of age salience and categorisation we expect to see an OAB in the mixed- but not in the pure-list condition. Alternately, if a perceptual-expertise mechanism underlies the OAB, we expect to see similar OABs in both list conditions, replicating Experiment 1 and 2.

### Experiment 3

#### Method

**Participants.** One-hundred and fifty six participants' data were analysed (70 male, 86 female;  $M_{age}=25.75$ ,  $SD_{age}=3.00$ , range=18-30). Participants were sourced from Amazon Mechanical Turk and compensated USD3.70 for their time. There were 52 participants in each list-type group (mixed, pure-young, and pure-old). Additional participants were removed prior to analyses because they did not fit demographic requirements (24 participants), completed the experimental sequence incorrectly (16 participants), or needed to be removed so that group numbers were equal and the required analysis method could be used (13 participants - 4 mixed, 9 pure-old).

**Stimuli.** The same stimuli as in Experiment 2 were used, with the addition of background stimuli. Background stimuli in exposure were a picture of a hospital room, a basketball court, and a white rectangle outlined in black. In the test phase, the background was a grey rectangle. These stimuli were  $636 \times 478$  pixels in size.

**Procedure and data analysis.** As per Experiment 1 with the following differences. Faces were presented overlaid on a background and participants were instructed to remember the face-background pairings. The possible backgrounds were the white rectangle, hospital room, and basketball court. An equal number of young and old, and male and female faces were presented on each background. Presentation time was also extended to 5000ms and 5500ms so that participants had sufficient time to remember the face-background pairs. On test trials, all faces were presented on the grey background. Where participants indicated that they had seen a face, they were prompted with the remember/know/guess question and then an additional source memory question. The source memory question presented the three backgrounds seen in the exposure phase on-screen and asked participants to select the background that the face had been presented on. As participants were instructed to remember the backgrounds, this source memory question was included, however, these data are not reported as they were not central to our predictions and too few trials were available for analysis.

## Results

One-way ANOVAs on the group contact data (young contact and old contact) with a between-groups factor of list-type (pure-young, pure-old, mixed) show that our groups did not differ in the amount of contact reported with young ( $F(2, 152)=0.49, p=.614, \eta^2=.01$ ) or older adults ( $F(2, 152)=0.22, p=.804, \eta^2<.01$ ). Paired-samples t-tests confirmed that our

sample had higher contact with young than older adults ( $t(153)=15.05, p<.001, d_{av}=1.74$ ), and identified more with young than older adults ( $t(153)=21.56, p<.001, d_{av}=2.91$ ).

Analysis of the  $P_r$  data using Erlebacher's (1977) method with factors of face age (young, old) and design (mixed, pure) was conducted (see Figure 3). There was a main effect of face age such that the OAB was present overall ( $F(1, 151)=6.41, p=.012, \eta^2=.03$ ). There was, however, no effect of design ( $F(1, 145)=1.85, p=.175, \eta^2=.01$ ), nor an interaction between face age and design ( $F(1, 151)=0.89, p=.347, \eta^2<.01$ ). Bayesian analyses found that a model containing only the age main effect was 102.74 times more likely than a model containing the age and interaction effects to explain the data.

An analysis of the hit rates produced no significant effect of age ( $F(1, 119)=0.48, p=.491, \eta^2<.01$ ), design ( $F(1, 139)=0.09, p=.761, \eta^2<.01$ ), nor an interaction between the two ( $F(1, 119)=0.29, p=.594, \eta^2<.01$ ). An analysis of the false alarm rates produced a main effect of face age ( $F(1, 131)=14.21, p<.001, \eta^2=.05$ ) such that fewer false alarms were made to young adult faces (see Table 3). There was also a main effect of design ( $F(1, 139)=8.51, p=.004, \eta^2=.05$ ), such that there were higher false alarm rates in the pure-list conditions. There was no interaction between face age and design ( $F(1, 131)=0.68, p=.411, \eta^2<.01$ ). Analysis of  $B_r$  produced a significant effect of age ( $F(1, 153)=6.01, p=.015, \eta^2=.02$ ) such that responding to young faces was more conservative than to older adult faces. There was also a significant effect of design ( $F(1, 141)=4.36, p=.039, \eta^2=.02$ ) such that responding was more conservative in the mixed than in the pure-list condition. There was no interaction between age and design ( $F(1, 153)=0.63, p=.428, \eta^2<.01$ ) for the  $B_r$  data.

**Supplementary analysis.** A supplementary analysis examined the relationship between self-reported contact with older adults and OAB magnitude (young  $P_r$  - old  $P_r$ ). Bias scores could only be calculated in a meaningful way for participants in the mixed-list

condition of each experiment so a combined analysis including data from all three experiments was conducted (findings were the same when each experiment was analysed individually). No relationship was found between self-reported contact with older adults and OAB magnitude ( $r(166) = .004, p = .956$ ).

-----Figure 3-----

-----Table 3-----

### **General Discussion**

The aim of this study was to replicate Bryce and Dodson's (2013) finding that the OAB emerges in young adult participants presented with a mixed- but not with a pure-list of young and older adult faces, and to further our understanding of the mechanism/s underlying the OAB. We predicted that the results would replicate and be robust when removing theoretically unimportant components of the Bryce and Dodson old/new recognition paradigm. Across three experiments that removed components of the original paradigm (Experiment 1) and included components replicating the encoding conditions (Experiment 2 and 3), the finding that the OAB only occurs in a mixed-list was not replicated. In each of the three experiments, the OAB was evident regardless of list-type.

Our results deviate from Bryce and Dodson's (2013) finding that suggests social-cognitive mechanisms influence the OAB. Bryce and Dodson argued that when participants view only one age group of faces, they do not engage in the age categorisation that is necessary to produce the OAB. They suggest that age group variance in a mixed-list context acts as a situational cue to make age salient and prompt age categorisation. Our results instead show that encountering a list of age mixed faces (or not) does not change performance. Additionally, these results indicate that the disparate pattern of results seen in comparison to results noted in the ORB field (where category salience reduces own-group

recognition performance rather than enhances it; Young et al., 2009) may not be reliable. This tempers the implication from Bryce and Dodson's study that the OAB and ORB are distinct social-cognitive phenomena, and instead suggests that social cognition may play a more important role for remembering differently raced faces, than differently aged faces.

Our results overall, are more consistent with a perceptual-expertise account. A perceptual-expertise account predicts that manipulations of category salience will have no influence on the OAB. Instead, the OAB should be driven by participants having more contact with own- than other-age people. Our results support these predictions, with the list-type manipulation having no effect on the OAB, and our participants reporting higher contact with young adults than older adults. However, unlike in previous studies of the OAB (Ebner & Johnson, 2009; He, Ebner, & Johnson, 2011), we did not find that participants' level of other-age contact was associated with their OAB. As only participants in the mixed-list had scores for both young and older adult recognition we were unable to determine if this pattern was the same for other participants. Additionally, while we used the same measure as the above mentioned studies, these are the only two cases we are aware of where this evidence has been produced. Most of the evidence supporting the role of experience in the OAB instead comes from group comparisons where participants with distinct experience profiles are compared (e.g. trainee teachers with high contact with children, versus young adult participants with no contact with children; Harrison & Hole, 2009).

The failure to replicate Bryce and Dodson's (2013) results in our paradigm may indicate that the perceptual-expertise account best explains the OAB, but it might also indicate that the use of mixed- versus pure-list conditions did not successfully manipulate social categorisation or face age salience in the context of the current experiments. The social categorisation literature suggests that category activation is fast and spontaneous (Brewer, 1988). Determining which categories are activated and under which conditions has been the



topic of much research (Macrae & Bodenhausen, 2000). Factors such as category salience, stereotype salience, beliefs and prejudices, motivations, and processing capacity have all been suggested to influence whether, and if so, which category is activated (Macrae & Bodenhausen, 2000). While our manipulation aimed to target category salience, other factors may have prevented the effectiveness of the manipulation. Perhaps for our participants, age categorisation occurred automatically regardless of the condition the faces were encountered in.

There remain methodological differences between the studies that may also account for the different patterns of results observed (see Table 4 for a summary). Some notable differences are that Bryce and Dodson's (2013) study included a single passive encoding task (versus our additional RT task), fewer new faces than studied faces at test (versus our equal number of new and studied faces), no delay between encoding and test (versus our 5 min delay), confidence ratings on all trials (versus our remember/know/guess question on "seen" response trials), and three encoding-test sessions (versus our one session). We did not examine the influence of these features as the RT task was designed to assist encoding in an online setting by requiring attention to the faces, and the remaining differences occur post-encoding and should not have influenced the OAB which is suggested to be created during the encoding phase (Wiese, Komes, et al., 2013; Young, Bernstein & Hugenberg, 2010). We would argue that even though there remain possible boundary conditions to the influence of using a mixed- versus pure-list on the OAB, these conditions are not theoretically relevant to the predictions made by social-cognitive models like the CIM (Hugenberg et al., 2010). If presenting a mixed- versus pure-list of faces is sufficient to manipulate social categorisation, and the OAB is partly a product of social categorisation, the influence of the manipulation should be seen regardless of the remaining untested methodological differences. If these methodological differences are responsible for producing different patterns of results, this is

likely indicative of their influence on other processes such as response bias or working memory rather than on the encoding processes that are thought to produce the OAB.

-----Table 4-----

Even when considering these differences, it is difficult to conceive how they may have produced Bryce and Dodson's (2013) results. The inclusion of an RT task in encoding could be argued to influence depth of encoding with shallower encoding seen in our experiments than in Bryce and Dodson's suggested by the overall poorer recognition performance in our study (corrected recognition ranged from approximately .20-.45 across our experiments versus .60-.85 in Bryce and Dodson's). However, the observation of the OAB in all conditions, suggests that encoding of the face stimuli was sufficiently deep to enable the categorisation and individuation processes that are said to be required for the emergence of the OAB. The inclusion of fewer new faces at recognition may influence response bias, making "seen" responses more frequent. However, given there were equal numbers of new own- and other-age faces, this response bias shift should be the same across face age types, maintaining the relative difference in performance. We can also see in our data a response bias OAB consistent to that reported in the Rhodes and Anastasi (2012) meta-analysis which indicated that young adult participants respond more conservatively to young than older adult faces. There was a significant effect of design in Experiment 3, however, this was not present in Experiment 1 or 2 and likely reflects differences in the participant groups. Similarly, the benefit conferred by faces being retained in working memory due to a lack of delay between encoding and test should be equal among the list-type conditions. The inclusion of confidence ratings and multiple testing rounds might produce learning effects whereby participants perform differently over time based on confidence or lack thereof in their performance. However, we would still assume that the first encoding-test session should yield a pattern of results that would match our own. Assuming that it did, this would require

that improvements in subsequent encoding-test sessions differed across list-types and face age (i.e. larger improvement for old faces in the pure list condition only) in order to yield the pattern of results reported by Bryce and Dodson. Such a selective improvement seems rather unlikely and is inconsistent with Bryce and Dodson's notion that performance should not differ across age in the pure-list condition.

Overall, the findings of this study do not replicate Bryce and Dodson's (2013) and do not provide support for the social-cognitive account of the OAB. While there remain methodological differences in the approach that we and Bryce and Dodson took, we are confident that the findings we have presented across three separate experiments reflect the effect of mixed- and pure-lists on the production of the OAB. These results are more consistent with the perceptual-expertise account but further research is required to fully assess the influence of categorisation on the OAB. The literature would benefit from research that more strongly manipulates and measures categorisation. Ambiguity manipulations such as those used in the ORB and own-sex bias literature where cues are used to indicate group membership of ambiguous faces (MacLin & Malpass, 2001; Huart, Corneille & Becquart, 2005) may be one such approach, although creating ambiguity would be much harder for age than it is for race or sex.

Our results qualify evidence in favour of a distinct social-cognitive account of the OAB. Beyond their implications to our understanding of the OAB and its differences and similarities to the broader own-group bias literature, our results are important for a field that almost exclusively uses designs with mixed-lists of own- and other-age faces during encoding. Our results indicate that the biases we are examining exist outside of these stimulus set composition constraints and are not a product of our experimental procedures. In real life we can encounter heterogeneous or homogenous groups, or interact with a single person. Our

results indicate that regardless of the nature of these encounters, an age-based recognition bias will be present.

### References

- Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., . . . Schaller, M. (2006). They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity. *Psychological Science, 17*(10), 836-840. doi:10.1111/j.1467-9280.2006.01790.x
- Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review, 12*(6), 1043-1047.  
doi:10.3758/BF03206441
- Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.
- Bryce, M. S., & Dodson, C. S. (2013). Cross-age effect in recognition performance and memory monitoring for faces. *Psychology and Aging, 28*(1), 87.
- Bürkner, P. C. (2018). Advanced bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395-411. doi:10.32614/RJ-2018-017
- Craig, B. M., & Thorne, E. M. (2019). Social categorisation and individuation in the own-age bias. *British Journal of Psychology*. Advance online publication.  
doi:10.1111/bjop.12376
- Cronin, S. L., Craig, B. M., & Lipp, O. V. (2018). Emotional expressions reduce the own-age bias. *Emotion*. Advance online publication. doi:10.1037/emo0000517
- D'Argembeau, A., Van der Linden, M., Comblain, C., & Etienne, A. (2003). The effects of happy and angry expressions on identity and expression memory for unfamiliar faces. *Cognition and Emotion, 17*(4), 609-622.
- Denkinger, B., & Kinn, M. (2018). Own-Age Bias and Positivity Effects in Facial Recognition. *Experimental Aging Research, 44*(5), 411-426.

- Ebner, N. C., & Johnson, M. K. (2009). Young and older emotional faces: Are there age group differences in expression identification and memory? *Emotion, 9*(3), 329-339. doi:10.1037/a0015179
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*(1), 351-362. doi:10.3758/BRM.42.1.351
- Erlebacher, A. (1977). Design and analysis of experiments contrasting the within-and between-subjects manipulation of the independent variable. *Psychological Bulletin, 84*(2), 212.
- 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.
- Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing.
- Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330-337.
- Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review, 16*(2), 264-269. doi:10.3758/PBR.16.2.264
- He, Y., Ebner, N., & Johnson, M. (2011). What predicts the own-age bias in face recognition memory? *Social Cognition, 29*(1), 97-109. doi:10.1521/soco.2011.29.1.97
- Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology, 48*(2), 499-508. doi:10.1037/a0026524
- Hills, P. J., & Lewis, M. B. (2011). The own-age face recognition bias in children and adults. *The Quarterly Journal of Experimental Psychology, 64*(1), 17-23.

- Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorisation, context-based categorisation, and distortions in the recollection of gender ambiguous faces. *Journal of Experimental Social Psychology, 41*(6), 598-608.
- Hugenberg, K., Wilson, J. P., See, P. E., & Young, S. G. (2013). Towards a synthetic model of own group biases in face memory. *Visual Cognition, 1*-26.  
doi:10.1080/13506285.2013.821429
- Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorisation-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*(4), 1168-1187. doi:10.1037/a0020463
- Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 811.
- Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*(4), 559-574.  
doi:10.1037/0096-3445.129.4.559
- Macchi Cassia, V., Picozzi, M., Kuefner, D., & Casati, M. (2009). Short article: Why mix-ups don't happen in the nursery: Evidence for an experience-based interpretation of the other-age effect. *Quarterly Journal of Experimental Psychology, 62*(6), 1099-1107. doi:10.1080/17470210802617654
- MacLin, O. H., & Malpass, R. S. (2001). Racial categorisation of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law, 7*(1), 98.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual review of psychology, 51*(1), 93-120.

- Merritt, P., Cook, G., & Wang, M. (2014). Erlebacher's method for contrasting the within and between-subjects manipulation of the independent variable using R and SPSS. Working paper.
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4), 630-633.
- Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, 52(4), 597-617.
- Proietti, V., Laurence, S., Matthews, C. M., Zhou, X., & Mondloch, C. J. (2018). Attending to identity cues reduces the own-age but not the own-race recognition advantage. *Vision Research*.
- R Core Team, (2017). R: A language and environment for statistical computing. [Computer software]. Vienna, Austria: R Foundation for Statistical Computing
- Reysen, S., Katzarska-Miller, I., Nesbit, S. M., & Pierce, L. (2013). Further validation of a single-item measure of social identification. *European Journal of Social Psychology*, 43(6), 463-470. doi:10.1002/ejsp.1973
- Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, 138(1), 146-174. doi:10.1037/a0025750
- Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition*, 5(2), 144-165.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2(6). 110-114.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34.



- Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, 7(1), 36.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 43(2), 161-204.
- Wiese, H., Komes, J., & Schweinberger, S. R. (2013). Ageing faces in ageing minds: A review on the own-age bias in face recognition. *Visual Cognition*, 21(9-10), 1337-1363. doi:10.1080/13506285.2013.823139
- Wiese, H., Wolff, N., Steffens, M. C., & Schweinberger, S. R. (2013). How experience shapes memory for faces: an event-related potential study on the own-age bias. *Biological psychology*, 94(2), 369-379.
- Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2009). Interracial contexts debilitate same-race recognition. *Journal of Experimental Social Psychology*, 45(5), 1123-1126. doi:10.1016/j.jesp.2009.05.009
- Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When do own-group biases in face recognition occur? Encoding versus post-encoding. *Social Cognition*, 28(2), 240-250.
- Young, S. G., & Hugenberg, K. (2012). Individuation motivation and face experience can operate jointly to produce the own-race bias. *Social Psychological and Personality Science*, 3(1), 80-87. doi:10.1177/1948550611409759

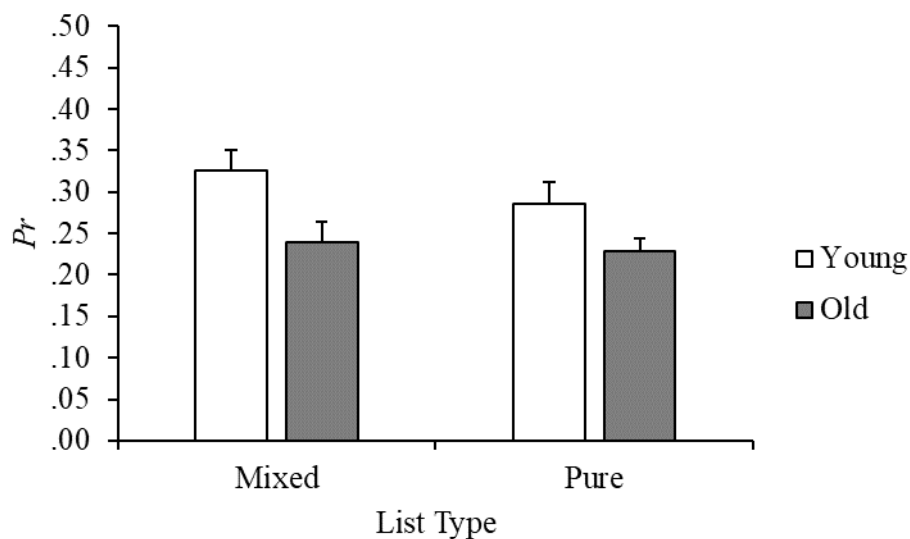


Figure 1. Corrected recognition scores ( $P_r$ ) to young and older adult faces as a function of list type in Experiment 1. Error bars are the standard error of each mean.

Table 1

*Hit rates, false alarm rates, and response bias for young and older adult faces in each list type group in Experiment 1*

	Hits		False Alarms		Response Bias ( $B_r$ )	
	Young	Old	Young	Old	Young	Old
Mixed	.55(.21)	.55(.22)	.19(.16)	.29(.20)	.34(.21)	.41(.22)
Pure-Young	.52(.20)	-	.22(.16)	-	.33(.19)	-
Pure-Old	-	.51(.18)	-	.28(.16)	-	.38(.21)

Note. Standard deviations are reported in brackets.

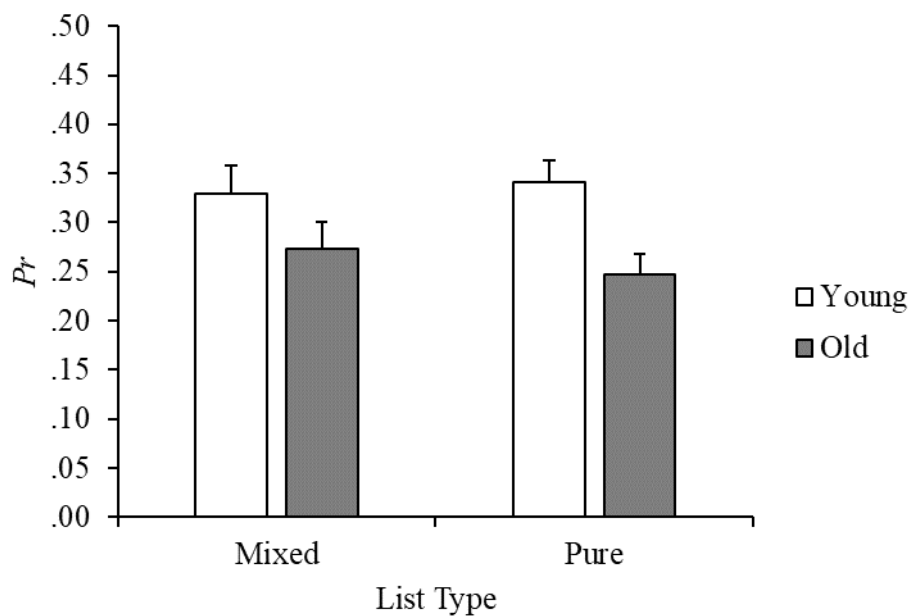


Figure 2. Corrected recognition scores ( $P_r$ ) to young and older adult faces as a function of list type in Experiment 2. Error bars are the standard error of each mean.

Table 2

*Hit rates, false alarm rates, and response bias for young and older adult faces in each list type group in Experiment 2*

	Hits		False Alarms		Response Bias ( $B_r$ )	
	Young	Old	Young	Old	Young	Old
Mixed	.56(.24)	.59(.24)	.21(.18)	.30(.23)	.35(.22)	.42(.24)
Pure-Young	.55(.16)	-	.20(.12)	-	.31(.16)	-
Pure-Old	-	.54(.17)	-	.29(.18)	-	.38(.19)

*Note.* Standard deviations are reported in brackets.

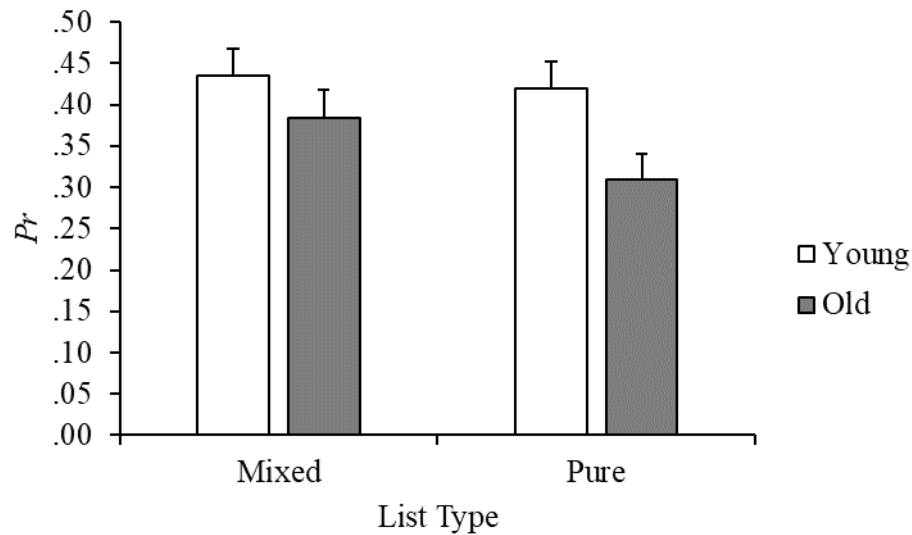


Figure 3. Corrected recognition scores ( $P_r$ ) to young and older adult faces as a function of list type in Experiment 3. Error bars are the standard error of each mean.

Table 3

*Hit rates, false alarm rates, and response bias for young and older adult faces in each list type group in Experiment 3*

	Hits		False Alarms		Response Bias ( $B_r$ )	
	Young	Old	Young	Old	Young	Old
Mixed	.57(.23)	.57(.21)	.11(.11)	.16(.16)	.25(.16)	.29(.19)
Pure-Young	.60(.20)	-	.16(.14)	-	.29(.18)	-
Pure-Old	-	.56(.19)	-	.24(.17)	-	.37(.22)

*Note.* Standard deviations are reported in brackets.

Table 4

*Summary of methodological features and differences in Bryce and Dodson's (Exp 2, 2013) and the current study's experiments*

Feature	Difference	Bryce and Dodson (Exp 2, 2013)	Current Study
Sample	Sample larger in the current study	N = 72 Young adults	Exp 1: N = 168 Exp 2: N = 177 Exp 3: N = 156 18-31 years old
Participation Setting	In-person vs. online	In-person	Online
Stimuli	1 vs. 2 face databases	Male and female faces (CAL/PAL database)	Male and female (Exp 1 male only) faces (CAL/PAL and FACES databases)
		Background scenes (basketball court, hospital room)	Background scenes (Exp 3: basketball court, hospital room)
Experimental sequence	3 vs. 1 old/new recognition tasks	Encoding 1, Test 1, Filler 1, Encoding 2, Test 2, Filler 2, Encoding 3, Test 3	Encoding, Filler, Test
Encoding Phase	1 vs. 2 tasks	Intentional encoding of 26 face-background pairs	Intentional encoding of 24 faces (Exp 3: 24 face- background pairs), RT task
		5s presentations	1-1.5s (Exp 1, 2) and 5-5.5s (Exp 3) presentations
Filler Task	Maze task vs. word puzzles	Maze task	Word puzzles
	Between vs. within the old/new recognition tasks	Between each old/new recognition task sequence	Between encoding and test
Encoding-Test Transition	Immediate vs. delayed	Immediate transition	Filler task between encoding and test
Test Phase	3:1 vs. 1:1 ratio of studied and new faces	24 studied faces, 8 new faces	24 studied faces, 24 new faces
	Confidence vs. Remember/Know/Guess question	Old/New judgement, confidence rating (after all trials), background selection	Seen/Not Seen judgement, remember/know/guess question (after 'seen' trials), background selection (Exp3)

