

# **A Gender In-Group Effect on Facial Recall**

Anna Bailey

A report submitted as a partial requirement for the degree of Bachelor  
of Psychological Science with Honours at the University of Tasmania, 2021

### **Statement of Sources**

I declare that this report is my own original work and that contributions of others have been duly acknowledged.

Signed:

Dated: 14/10/2021

## **Acknowledgements**

Peter Tranent for being a respectful and motivational supervisor, fostering a collaborative study where my ideas were heard from the outset. Providing guidance, constructive feedback, an empathetic ear when things did not go as planned, and many long zoom sessions.

The University of Tasmania School of Psychological Sciences for financial support, allowing for participant compensation and access to programming systems.

Ben Howell from PsychStudio for his non-judgmental support with the technical software.

Matt Palmer for access to the stimuli and consultation on project direction.

Christopher Pearn, Michael Bailey, Jennifer Bailey, Sarah Ryan, Samantha Bugg, and Masuka Bailey for your never-ending support not least of which being a consistent supply of comfort foods and laughter.

## Table of Contents

|                                                                           |           |
|---------------------------------------------------------------------------|-----------|
| <b>ABSTRACT.....</b>                                                      | <b>1</b>  |
| <b>INTRODUCTION .....</b>                                                 | <b>2</b>  |
| IMPACTS OF FACE GENDER .....                                              | 3         |
| THE IN-GROUP EFFECT.....                                                  | 5         |
| MODELS FOR CONFIDENCE RATINGS .....                                       | 8         |
| PURPOSE OF THE PRESENT STUDY.....                                         | 10        |
| <b>METHOD .....</b>                                                       | <b>11</b> |
| DESIGN .....                                                              | 11        |
| PARTICIPANTS.....                                                         | 11        |
| MATERIALS.....                                                            | 12        |
| <i>Questionnaire</i> .....                                                | 12        |
| <i>Stimuli</i> .....                                                      | 12        |
| PROCEDURE .....                                                           | 13        |
| <b>RESULTS .....</b>                                                      | <b>15</b> |
| <b>DISCUSSION.....</b>                                                    | <b>20</b> |
| FACE RECALL ACCURACY.....                                                 | 21        |
| CONFIDENCE RATINGS .....                                                  | 29        |
| IMPLICATIONS.....                                                         | 32        |
| LIMITATIONS .....                                                         | 34        |
| CONCLUSION .....                                                          | 36        |
| <b>REFERENCES .....</b>                                                   | <b>38</b> |
| <b>APPENDIX A: EXPERIMENTAL CONDITION ARTICLE .....</b>                   | <b>49</b> |
| <b>APPENDIX B: CONTROL CONDITION ARTICLE.....</b>                         | <b>52</b> |
| <b>APPENDIX C: EXPERIMENTAL CONDITION MANIPULATION CHECK QUESTIONS ..</b> | <b>56</b> |

|                                                                         |           |
|-------------------------------------------------------------------------|-----------|
| <b>APPENDIX D: CONTROL CONDITION MANIPULATION CHECK QUESTIONS .....</b> | <b>57</b> |
| <b>APPENDIX E: EXAMPLES OF FACES USED .....</b>                         | <b>58</b> |
| <b>APPENDIX F: EXAMPLE OF FILLER TASK ACTIVITY.....</b>                 | <b>59</b> |
| <b>APPENDIX G: ETHICS APPROVAL LETTER.....</b>                          | <b>60</b> |
| <b>APPENDIX H: PARTICIPANT INFORMATION SHEET .....</b>                  | <b>62</b> |
| <b>APPENDIX I: PARTICIPANT CONSENT FORM .....</b>                       | <b>65</b> |

## **List of Tables and Figures**

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| TABLE 1: PARTICIPANT DESCRIPTIVE STATISTICS BY GENDER AND CONDITION.....      | 12 |
| TABLE 2: CLASSIFICATIONS OF PARTICIPANT RESPONSES DURING TEST PHASE.....      | 15 |
| FIGURE 1: THE IMPACT OF GENDER ON ACCURACY ACROSS CONDITIONS.....             | 17 |
| TABLE 3: THE IMPACT OF GENDER ON CONTROL CONDITION FACE RECALL ACCURACY ..... | 17 |
| TABLE 4: THE IMPACT OF ARTICLE READ ON FACE RECALL ACCURACY.....              | 19 |

# **A Gender In-Group Effect on Facial Recall**

Anna Bailey

Word Count: 9,991

### **Abstract**

In-group bias in face recall accuracy has been consistently demonstrated with individuals performing better when encountering faces of their own social categorisation. Research has discovered this effect across race and age and has uncovered some evidence of an own gender bias. Discrepancies remain however, especially when alternative in-groups are established. The current study assessed 106 (68 female, 38 male) adult participants' recall accuracy across male and female faces after reading a news article, either highlighting gender inequalities or collective experiences. No significant differences between conditions in either recall or confidence were observed. There were also no significant differences in recall accuracy between genders. Furthermore, no significant differences were exhibited in women's accuracy across face genders. However, a statistically significant medium strength positive difference was found in men's accuracy of male faces when compared to female faces. Additionally, the results revealed men recorded significantly higher confidence ratings than women. These findings advance understanding of the impacts of gender and social identity on recall and confidence ratings. They suggest confidence ratings adhere to a general intelligence model (impacted by both fundamental knowledge and stressors). Therefore, more reliable methods of assessing identification accuracy in eye-witness testimonies than collective confidence ratings should be developed.

Gender effects on facial recognition are widely acknowledged, with females consistently exhibiting higher accuracy rates than their male counterparts (Palmer et al., 2013). A meta-analysis of 43 studies, including 10,100 participants (5373 females) aged between four and 53 years, found female participants recognised significantly more faces than male participants (Herlitz & Loven, 2013). Additionally, research has shown that women display higher levels of facial memory compared to men, even after accounting for gender differences in general cognitive functions (Sommer et al., 2013). Various theories have been utilised in an attempt to explain this phenomenon (Megreya et al., 2011). These arguments include biological brain structure and hormonal differences between genders as well as variances in social interactions.

Memory is thought to be a complex system of components each responsible for specific aspects and located across multiple brain regions (Bryzgalov et al., 2018). Declarative memory is suggested to be one of these components and manages the storage and organisation of episodic and semantic knowledge to allow for conscious recall of the information in future (Pisoni et al., 2015). This memory system is thought to be located in the medial temporal lobe (including the hippocampus), the posterior parietal cortex, and the basal ganglia (Rutishauser, 2019). A subset of declarative memory is recognition memory which allows for future recall of facts and details (e.g., remembering faces) (Maitland et al., 2004). Previous research has demonstrated that on average females possess higher recognition memory abilities (Graves et al., 2017). Moreover, age related decline in recognition memory is faster and more severe in men when compared to women (Gomez et al., 2020). It is theorised that these gender differences may be linked to genetic differences within the dopamine system; with women exhibiting a higher baseline dopamine level than men (Van der Auwera et al., 2021). An increase of dopamine, when viewing and encoding faces, may

also be related to women's higher levels of self-expressed positivity and arousal as a response to human faces, in comparison to men's (Proverbio, 2017).

On the other hand, it is widely accepted that women have a higher skill level in perception and comprehension of non-verbal social cues (Pavlova et al., 2016). These abilities perhaps allow for superior encoding of faces as many non-verbal social cues rely on facial reactions and mannerisms. Additionally, women's face recognition skills may be linked to traditional childhood preferences for social toys (e.g., dolls) in comparison to boys' preferences for inanimate objects (e.g., cars) (Proverbio, 2017). Furthermore, women's interpersonal relationships have been found to be more intimate (Buhrke & Fuqua, 1987) and value personal connection rather than larger group identity (Seeley et al., 2003) or social standing (Kwang et al., 2013) that men's typically do. This is perhaps a contributing factor to the smaller social circles or duos women commonly inhabit in comparison to larger male groups (Mjaavatn et al., 2016). This focus on emotional connection may contribute to the finding that women are better at identifying facial emotions (Lin et al., 2021) and may increase the attentional span dedicated to encoding faces. Moreover, longer durations of eye fixation on the other person's face may elicit social rewards such as an increased likelihood of a responding smile (Gueguen et al., 2008). This positive reinforcement for looking at faces may encourage women to continue looking at faces in future which then assists with later recall. The closeness of female relationships has also been found within cross-gender relationships with men exerting more effort within these connections than their male only groups (Buhrke & Fuqua, 1987). This finding may lend itself to the notion of increased accuracy when identifying female faces.

### **Impacts of Face Gender**

It has also been revealed that women more accurately identify female faces in comparison to male faces (Rehman & Herlitz, 2007). Eye-tracking data has found women

fixate on the eyes of female faces for longer than other gender faces, perhaps allowing for a deeper processing and therefore easier recall (Man & Hills, 2016). Most of the evidence, however, points towards this effect being an outcome of the gender bias. This concept is that individuals are more motivated to remember faces of their own gender and is thought to happen due to higher occurrences of interactions being with same gendered people (Wright & Sladden, 2003). A social preference for same gender groups is seen even in school aged children (Mjaavatn et al., 2016). Additionally, Hills and colleagues (2018) recently found that girls attending single-sex schools rated female faces as more attractive than girls attending coeducation schools. The study also uncovered both male and female participants rated female faces as more attractive than male faces (Hills et al., 2018). Given that facial recognition studies have found attractive faces to be more easily remembered than unattractive faces (Malloy et al., 2021) this preference for female faces may give cause to the high accuracy rates exhibited.

Superior accuracy for female faces (in comparison to male faces) has also been demonstrated in men. Research has found both males and females exhibit higher recall performances for female faces across ages (Wang et al., 2020), suggesting it may stem from early childhood exposure. This notion is supported by Marquis and Sugden's (2019) meta-analysis of 76 samples which uncovered infants display early preferential attention to faces that parallel the same gender and race of their primary caregiver. The study found that the level of attention declined as the infant aged, however the preference for the familiar gender remained (Marquis & Sugden, 2019). Arguments in support of these findings have included notions that women are subjected to heightened exposure to physical attractiveness pressures. Women's physical attractiveness has been linked with their perceived level of humanity (Alaei et al., 2021), intelligence (Hernandez-Julian & Peters, 2017), and trustworthiness (Ch'ng, 2021). Tian and colleagues (2019) also found that both male and female participants

perceived time as passing faster when looking at an attractive female face. It is thought that this emphasis on female appearance may lead to an increased focus for women on facial features and on female faces overall (Palmer et al., 2013).

The rate of which men more accurately recall female faces (if at all), is contentious and has been linked to a childhood bias proving stronger than an adult own-group preference (Palmer et al., 2013). A number of studies have failed to find a male recall preference for female faces and instead encountered either no effect (Man & Hills, 2016) or a male face preference (Wolff et al., 2014). Previously, no effect has been linked to external influences, such as face attractiveness (Malloy et al., 2021) or an alternative in-group preference (Hehman et al., 2010). Recently another explanation has come to light. El Haj and colleagues (2019) found that males who identified as not in a relationship had a significantly higher recognition of female faces than male faces, whereas males in relationships showed no significant gender preference. The same effect was not seen in females. Female participants consistently outperformed males on facial recall and demonstrated a preference for female faces, regardless of relationship status (El Haj et al., 2019).

As a result, the current study will examine gender as an in-group and attempt to increase this bias by making experimental group participants aware of gender inequalities prior to encoding. It is thought that this will increase recognition of own gender.

### **The In-Group Effect**

In-group bias has been consistently evident across a variety of facial recognition tasks (Hungenberg et al., 2010). Some of the most robust results have been regarding the own-race bias and the own-age bias. These biases are reflective of the findings that participants consistently recall faces of their own race and age at a significantly higher accuracy rate and significantly lower false alarm rate than faces of other ethnicities (Meissner & Brigham, 2001) or ages (Rhodes & Anastasi, 2012). It has been suggested that the own-ethnicity bias

may stem from the viewer focusing on facial features that would be distinctive for their ethnicity, however, are not for the ethnicity of the face. For example, Hills and colleagues (2020) found that participants who fixated on the eyes of the face were significantly less accurate in identifying ethnically Black faces in comparison to ethnically White faces. Another possibility for this effect may be the familiarity of same ethnicity faces due to increased time spent interacting within these groups, whether within social circles or the family home.

Similarly, it is thought that own-age bias is due to individuals' higher percentage of interactions with peers of a similar age (e.g., in the corporate or education sectors). However, recently Denkinger and Kinn (2018) found younger adults (18-27 years) displayed an own-age face recognition bias whereas older adults (62-80 years) did not. The researchers suggest this is due to older adults engaging more within the community and therefore interacting with people from various age groups (Denkinger & Kinn, 2018). Moreover, it has been found that specialists who work with older adults regularly, such as geriatric nurses, do not exhibit the same-age bias in facial recognition tasks (Wiese et al., 2013). These findings support the notion of a familiarity effect underlying the in-group recognition bias, however, seemingly leaves room for discrepancies within the phenomenon.

Recently the persistence of the in-group effect has been a source of contention. This is evident in the historically strong replication across racial differences (Bernstein et al., 2007). Ng and colleagues (2020) recently found the effect of race on facial recognition was not as strong in first generation immigrants as in second generation immigrants. As a result, the researchers concluded the bias is more linked to social groups, particularly within Western societies, than race itself (Ng et al., 2020). Furthermore, Hills and colleagues' (2020) study of 231 participants (131 female) found that those with higher Intelligence Quotient (IQ) scores correctly recognised more faces of both Black and White ethnicities and therefore exhibited

lesser rates of the in-group effect. The researchers suggest that this may be due to high levels of intelligence allowing the viewer to encode a novel face via a variety of methods rather than being limited to methods only suitable for in-group recognition (Hills et al., 2020).

The impact of race as an in-group has also been overpowered by experimentally induced alternative in-groups, such as university groupings (Hehman et al., 2010). These findings suggest that the in-group bias may be able to be manipulated to either increase or decrease facial recognition results. However, Harrison and colleagues' (2020) recent study found that an in-group bias could only be experimentally induced when the categorisation was a source of community polarisation. In their study the experimenters found that university alumni were not a sufficient grouping to elicit an in-group facial familiarity effect within a pool of United Kingdom participants. Harrison and colleagues' (2020) were able to find an effect when manipulating groups based on Referendum vote (to "leave" or "remain" in the European Union). This may explain why racial grouping biases can be superseded by current societal or community issues. Furthermore, considering the high quantity of media attention on inequalities (e.g., gender and race), this may provide insight into why facial recall remains impacted by in-group associations.

Therefore, the current study will manipulate the existing in-group of gender, in an effort to better understand facial recognition bias. A recent study found that societal gender inequality is correlated with increased performance recalling celebrity female faces (Mishra et al., 2019). Mishra and colleagues (2019) found an increased divide between male and female participant recognition accuracy for those from low gender equality countries (e.g., India and Brazil) when compared to high gender equality countries (e.g., Sweden and Netherlands). This study, however, did not consider whether the inequality was well known to participants or salient at the time of recall. Indeed, there is currently no research on the impact of making participants conscious of gender equalities prior to learning and recalling

faces. As a result, the present study will attempt to manipulate the sense of polarisation between males and females within the experimental group through discussing popular issues such as the gender pay gap. It will do this utilising readily available media materials, strengthening external validity of any results observed.

### **Models for Confidence Ratings**

Several models attempt to explain the cognitive processes underlying face recognition bias, including crystallised and fluid intelligences (Dunn, 2004). Measurement of individual confidence in accuracy can be reflective of this. The level of the confidence-accuracy relationship is crucial in situations such as eye-witness testimonies as this can carry large weightings in the justice system. Given that false identifications can lead to dire consequences understanding confidence rating reliability is necessary. This is evidenced through eye-witness testimonies accounting for the majority of convictions that have later been overturned through DNA evidence (Innocence Project, 2021). It is widely accepted that confidence ratings directly reflect the strength of the underlying memory signal (Delay & Wixted, 2021). Therefore, in general it is expected that confidence will increase with accuracy (Davis et al., 2019). However, there are results that suggest confidence may also be impacted by external factors such as stress, suggesting that the confidence-accuracy effect may be more complicated (Davis et al., 2019).

Davis and colleagues (2019) attempted to replicate the psychological stress witnesses are often subjected to during the encoding stages of their memory acquisition. The experimenters required 111 participants to submerge their non-dominant hand in water with the temperature of the water being approximately 22 degrees Celsius for those in the control condition and zero degrees Celsius for 55 experimental participants. Participants in the stress condition displayed less accurate memory, however, they reduced their confidence ratings accordingly. While this study is still reflective of the confidence-accuracy theory this finding

suggests that confidence ratings may reflect an element of fluid intelligence systems as they adapt in relation to external or internal stressors. In comparison, a crystallised intelligence system would be expected to remain rigid regardless of external factors as it is based on fundamental knowledge. However, given the participants still conformed to the confidence-accuracy model, confidence may actually be influenced by a combination of fluid and crystallised knowledges. A general processing system (an amalgamation of both crystallised and fluid models) would remain relatively consistent as it is based on prior and more fundamental knowledge, such as social constructs, but may be somewhat impacted by external stressors (Cattell, 1963). This model may therefore be a more accurate assessment of the metacognitive process of confidence ratings.

However, there are potential limitations to Davis and colleagues' (2019) study. Particularly regarding the disputable direct correlation between physical stress (such as holding a hand in ice water) and psychological distress. In conjunction, whether this is sufficient to accurately interpret the impact of psychological stressors on the confidence-accuracy relationship. Furthermore, if the unique and conscious stressor of holding a hand in ice water is generalisable to more insidious stressors encountered in daily life (e.g., pressures from the media). As a result, the literature is still lacking a reliable assessment of the impact of psychological stressors on self-perceived accuracy.

A more thorough understanding of the impact of psychological stressors on metacognition is necessary. Especially considering the prior literature stipulates confidence ratings are most accurate when directly following memory recall (Smalarz et al., 2021) and the trauma often related to events eyewitnesses are required to recount. Therefore, the current study will look to clarify the existing literature regarding this effect. This will be achieved through a manipulation based on gender, which is highlighting society's inequality towards women. It is expected that making this inequality salient will induce psychological distress on

participants allowing insight into the confidence-accuracy relationship after exposure to psychological stressors.

### **Purpose of the Present Study**

Overall, the current study aims to evaluate the impact of enhancing gender in-group membership through manipulation, in the form of a news article discussing prominent gender inequalities, on facial recall accuracy. It is expected that the article will also induce psychological distress in participants (particularly female). Consequently, by requesting confidence ratings a more in-depth understanding of stressor impact on the confidence-accuracy relationship will be obtained. The results of this study will assist in further developing the current literature. Additionally, through making the manipulation simulate information encountered everyday through the media, the study is expected to be more readily generalisable when compared to the previous research. As a result, providing a better understanding of social cognitive behaviour and eyewitness recall reliability.

Therefore, the current study proposes six hypotheses.

H1a; Female control participants will accurately recall significantly more faces overall than male control participants.

H1b; Control participants will accurately recall significantly more female faces than male faces.

H2a; Female participants in the experimental condition will recall significantly more female faces than females in the control condition.

H2b; Males in the experimental condition will recall significantly more male faces than males in the control condition.

H3a; Females in the experimental condition will record significantly lower confidence ratings than females in the control condition.

H3b; Males in the experimental condition will record significantly higher confidence ratings than those in the control condition.

## **Method**

### **Design**

In this study the nominal independent variables are the article read (neutral/gender inequality), participant gender (male/female), and gender of face (male/female). The dependent variables are the accuracy rate of face identifications and confidence in the identification, both measured on a continuous scale.

### **Participants**

The study was promoted utilising a mixture of psychology undergraduate announcements and posts on social media. Participants needed to be between the ages of 18 and 65, have no known memory impairments, and adequate natural/corrected vision. This criterion ensured that the data is a representation of facial recognition rather than a reflection of cognitive or visual limitations. Additionally, participants were required to enter their relationship status, ethnicity, and gender orientation information as necessary for the final analysis. Therefore, when participants elect ‘Prefer not to say’ or ‘Other’ for gender orientation, the data was excluded from the final study. These participants, however, still received course credit or were entered into the prize draw to ensure they were not disadvantaged.

Of the 125 participants who completed the study, two identified as “other” for gender, four were outside of the age restrictions, and 13 did not pass the manipulation check. This left a total participant pool of 106 (68 female) with a mean age of 30.12 years ( $SD = 13.15$ ). The majority of participants identified as Caucasian ( $n = 89$ ), with Asian ( $n = 11$ ), Indigenous ( $n = 5$ ), and African ( $n = 1$ ) ethnicities also represented. As shown in Table 1 on participant descriptive statistics, 60% of participants reported being in a relationship.

**Table 1***Participant Descriptive Statistics by Gender and Condition*

| Variable                  | Female Participants |              | Male Participants |              | Total |
|---------------------------|---------------------|--------------|-------------------|--------------|-------|
|                           | Control             | Experimental | Control           | Experimental |       |
| <i>n</i>                  | 30                  | 38           | 13                | 25           | 106   |
| Mean age                  | 27.50               | 27.66        | 27.77             | 38.24        | 30.12 |
| In a relationship (%)     | 50                  | 61           | 62                | 72           | 60    |
| Not in a relationship (%) | 50                  | 32           | 31                | 24           | 35    |
| Other (%)                 | 0                   | 3            | 23                | 4            | 5     |

*Note.* Other = “It’s complicated” or “Prefer not to say” responses to relationship status.

**Materials*****Questionnaire***

Initially information regarding the individual’s gender, ethnicity, relationship status, and age was obtained, for data and eligibility requirements. Next, participants responded to five true or false questions, regarding the stimulus. If the participant responded correctly to at least three of the questions (60% accuracy) it was assumed that the stimulus had been read and understood. In the test of recall stage of the experiment participants were required to indicate if the presented face was “new” or “old”. Finally, participants rated their confidence in the accuracy of their novelty response on a 6-point scale from chance (50%) to certainty (100%).

***Stimuli***

Two articles were composed from four recent news reports (Davey, 2021; Hare, 2020; Hutchens et al., 2021; Janda & Pupazzoni, 2021). One worked to strengthen the divide between genders (Appendix A), outlining things such as the pay gap and inequalities in the work force. This article utilised exclusive language (e.g., females and males). The neutral

article (Appendix B) emphasised the shared human experience of wage decreases and used inclusive language (e.g., humanity and people).

A pool of 68 (34 female) different faces was compiled from the databases of previous studies. From this pool two groups of 40 (20 female) images were randomly selected utilising an online number generator. The faces are all of a neutral expression and images were non-pixelated. Each face had two images, one facing directly at the camera and one slightly off centre. This allowed for accurate assessment of face recognition rather than image recall (Bruce, 1982).

The software program PsychStudio was utilised to construct and run the experiment.

## **Procedure**

Participants completed the study remotely, using a desktop or laptop computer (sign up instructions stipulated this requirement). After reading and agreeing to the experiment outline and consent document, participants were asked to complete a general information survey including age, ethnicity, relationship status, and gender. This information was completely anonymous and the participants were able to elect not to respond. In the event participants elected “prefer not to say” or “other” (in the case of gender) they were able to complete the study and gain compensation, however, the data was excluded from analysis. This is due to the finding that in-group bias in facial recognition may be a Western societal effect and the lack of representation of gender non-binary participants in existing literature (as well as the limited availability to compile an adequate sample size).

Participants were then randomly allocated to either the control or experimental condition and to one of two groupings of 40 (20 female) faces. Next, all participants were presented with one of the two articles to read. The articles were followed by five comprehension questions (Appendices C and D).

Participants were then shown a series of 20 (10 female) emotion neutral faces, looking either straight into the camera or slightly off-centre (study phase) (Appendix E). Each face was presented for 2 seconds with a 250-millisecond gap between each to align with previous research (Bernstein et al., 2007; Palmer et al., 2013).

Participants then completed a short (approximately five minutes) unrelated and gender-neutral filler task in order to reduce subject rehearsal and effects of prospective memory (McBride et al., 2011). During this task participants were shown a series of black and white squares for three seconds. The sequence would then immediately reappear with one square having changed colour. In order to progress to the next image participants were required to select the button that was numbered the same as the different square (Appendix F). The filler task consisted of nine sequences ranging from four to 15 squares and increasing in difficulty. Participants were required to select the correct response before progressing to the next task. This filler task was constructed for this study to coincide with previous research (Palmer et al., 2013).

At the completion of the filler task participants were consecutively shown 40 faces (20 female) which include the 20 previously shown faces randomly dispersed throughout (test phase). The images presented in the test phase were the opposite to those shown during the study phase. For example, if female face 15 was shown slightly off-centre in the study phase female face 15 would be presented facing straight at the camera in the test phase. This was to ensure the participants were recalling the face instead of the picture itself (Bruce, 1982). Furthermore, to strengthen the validity of the analysis the faces and positioning were randomly selected prior to the study through the use of an online number generator with equal forward and off-centre male and female faces in both the learning and test phases. The order of the pre-selections was then randomised by the software for each trial.

For each face participants were asked if the face had already been presented or if it was new. Finally, participants were asked to provide their level of confidence in their response with 50% (chance) being the lowest and 100% (certainty) the highest (Palmer et al., 2013). Once both responses had been made the next face was shown until all 40 faces had been assessed.

On completion participants were presented with a debrief and gratitude screen with contact details made available should they have had any questions or concerns; this needed to be acknowledged (through a tick box) before compensation was granted. In all the study took approximately half an hour to an hour to complete. Participants were compensated for their time through either course credit (1 credit) or entry into the draw to win one of two e-gift cards (valued at \$50 AUD each).

## Results

As outlined in Table 2, responses in the test phase of the study were categorised into four classifications. These were based on whether the face was shown during the learning phase (target) or not (lure). In order to remain consistent with previous research the current analysis considered both total accuracy rates (sum of hits and correct rejections) (Malloy et al., 2021) and  $d'$  values, calculated using hit and false alarm rates (Palmer et al., 2013).

**Table 2**

*Classifications of Participant Responses During Test Phase*

| Classification    | Definition                                                    |
|-------------------|---------------------------------------------------------------|
| Miss              | Target face presented but incorrectly identified as lure face |
| Hit               | Target face presented and correctly identified                |
| Correct Rejection | Lure face presented and correctly identified                  |
| False Alarm       | Lure face presented but incorrectly identified as target face |

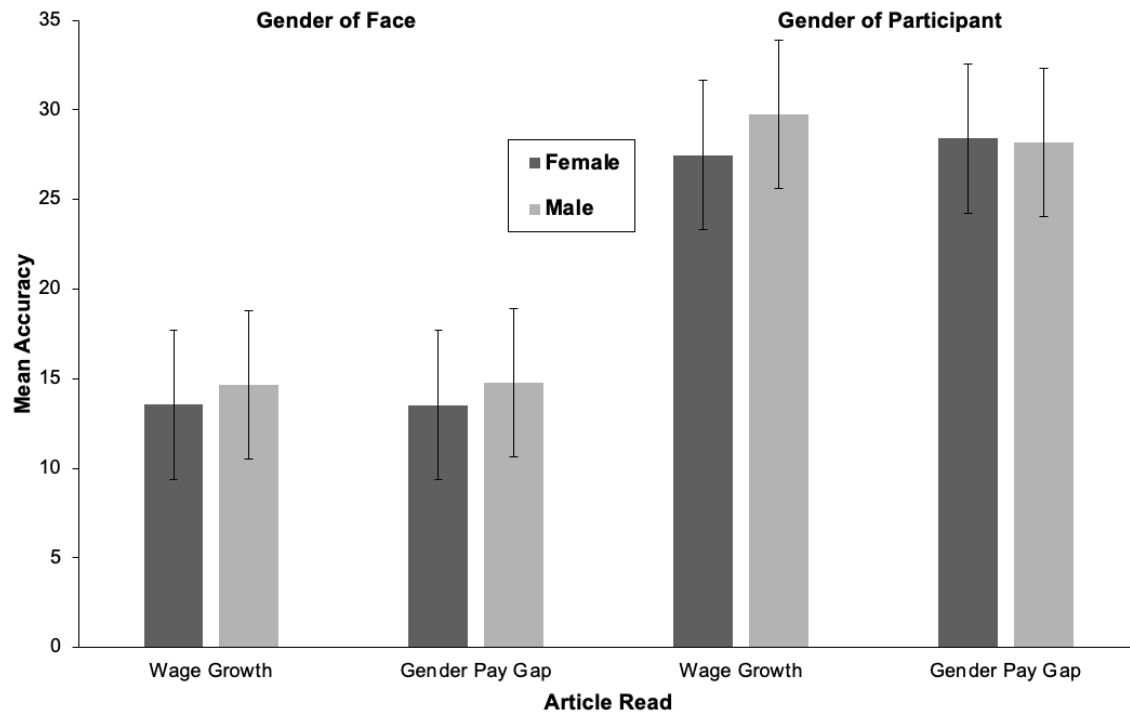
Conforming with the existing literature, the results were initially analysed utilising a 2 (participant gender) x 2 (face gender) x 2 (article read) mixed analysis of variance (ANOVA) with a series of Bonferroni adjusted t-tests conducted post-hoc to better interpret the results (Man & Hills, 2016; Mishra et al., 2019; Palmer et al., 2013). Additionally, to assess changes in confidence ratings a 2 (participant gender) x 2 (article read) ANOVA was used, again with post-hoc Bonferroni adjusted t-tests. While the results uncovered some minor evidence directional to what was anticipated in H2a, H2b, and H3a there were no statistically significant differences in support of the hypotheses. Additionally, the inverse to H1b was uncovered with control participants recalling more male faces than female faces however this was also not statistically significant.

Please see Figure 1 for a comparison of the impact of gender on recall accuracy within the control and experimental conditions. The results indicate male control participants accurately recalled more faces overall and recorded a higher  $d'$  value ( $d' = 1.52$ ,  $SD = 0.54$ , 95% CI [1.22,1.81]) than female control participants ( $d' = 1.23$ ,  $SD = 0.67$ , 95% CI [0.89,1.37]). Completing an independent samples t-test, as depicted in Table 3, revealed this difference was non-significant. This finding contradicts the hypothesis that female control participants would record significantly higher accuracy rates.

As represented in Figure 1, participants recorded more hits and correct rejections for male faces than female faces after reading the wage growth (control) article. Conducting an independent t-test analysis reveals this to be a not statistically significant when accounting for a Bonferroni adjustment (see Table 3). This was replicated when accounting for false alarm rates as participants displayed a  $d'$  value of 1.32 ( $SD = 0.81$ , 95% CI [1.07,1.56]) for female faces compared to a  $d'$  of 1.44 ( $SD = 0.99$ , 95% CI [1.14,1.73]) for male faces. These results are inconsistent with the hypothesis that control participants would recall significantly more female faces than male faces instead indicating the reverse.

**Figure 1**

*The Impact of Gender on Accuracy Across Conditions*



*Note.* Mean correct response (hit and correct rejection) scores across conditions for male and female participants and male and female faces. Error bars show standard errors.

**Table 3**

*The Impact of Gender on Control Condition Face Recall Accuracy*

| Variable           | <i>t</i> (101) | <i>p</i> | Cohen's <i>d</i> | 95% CI       |
|--------------------|----------------|----------|------------------|--------------|
| Correct responses  |                |          |                  |              |
| Participant gender | 1.64           | .629     | 0.62             | [1.00,4.80]  |
| Face gender        | 2.24           | .164     | 0.36             | [0.15,2.04]  |
| <i>d'</i>          |                |          |                  |              |
| Participant gender | 1.11           | 1.00     | 0.61             | [-0.04,0.82] |
| Face gender        | 0.63           | 1.00     | 0.11             | [-0.21,0.46] |

*Note.* *p* = statistical significance with Bonferroni adjustment applied; CI = confidence interval; *d'* = *d* prime.

Table 4 displays the statistics for recognition accuracy across participant gender and condition. After reading the gender pay gap article women correctly recalled marginally more female faces ( $M = 13.71$ ,  $SD = 2.56$ , 95 % CI [12.90,14.52]) than those in the control condition ( $M = 13.30$ ,  $SD = 2.40$ , 95 % CI [12.44,14.16]), however, the  $d'$  value remained consistent across conditions. Conducting an independent samples t-test uncovered female participants' accuracy difference between conditions to be not statistically significant,  $t(101) = -1.00$ ,  $p = 1.00$ ,  $d = -0.17$ , 95% CI [-1.62,0.80]. These findings contradicted the hypothesis that female participants in the experimental condition would recall significantly more female faces than those in the control condition.

Furthermore, female participants recalled more male faces in the experimental condition than the control condition (Table 4), however, according to an independent samples t-test neither the  $d'$  statistic ( $t(101) = -0.45$ ,  $p = 1.00$ ,  $d = -0.11$ , 95% CI [-0.55,0.34]) and total correct responses ( $t(101) = -0.85$ ,  $p = 1.00$ ,  $d = -0.22$ , 95% CI [-1.69,0.65]) were significantly different.

As Table 4 indicates, female participants also recalled more male faces than female faces overall. Following a Bonferroni adjusted t-test neither the differences in the  $d'$  value ( $t(101) = -0.39$ ,  $p = 1.00$ ,  $d = -0.04$ , 95% CI [-0.32,0.22]) nor the total correct responses ( $t(101) = 2.61$ ,  $p = .079$ ,  $d = 0.32$ , 95% CI [0.22,1.63]) were significant. These findings oppose the gender in-group effect hypothesised.

Males in the control condition recorded more hit and correct rejections of male faces than those in the experimental condition (Table 4). An independent samples t-test uncovered this finding to be not statistically significant ( $t(101) = 0.62$ ,  $p = 1.00$ ,  $d = 0.28$ , 95% CI [-1.11,2.57]). Additionally, males who read the article on the gender pay gap (experimental) performed at a higher  $d'$  value than those in the control condition. Again, completing an independent samples t-test revealed these differences to be non-significant ( $t(101) = 0.12$ ,  $p =$

.33,  $d = 0.14$ , 95% CI [-0.56,0.84]). This contradicts the hypothesis that men in the experimental condition will accurately recall significantly more male faces than those in the control condition.

**Table 4**

*The Impact of Article Read on Face Recall Accuracy*

| Variable            | Control Condition |               | Experimental Condition |               |
|---------------------|-------------------|---------------|------------------------|---------------|
|                     | <i>M(SD)</i>      | 95% CI        | <i>M(SD)</i>           | 95% CI        |
| Female Participants |                   |               |                        |               |
| Correct responses   |                   |               |                        |               |
| Total               | 27.47(4.14)       | [25.99,28.95] | 28.40(3.69)            | [27.22,29.57] |
| FF                  | 13.30(2.40)       | [12.44,14.16] | 13.71(2.56)            | [12.90,14.52] |
| MF                  | 14.17(2.72)       | [13.19,15.14] | 14.68(2.11)            | [14.01,15.35] |
| $d'$                |                   |               |                        |               |
| Total               | 1.13(0.67)        | [0.89,1.37]   | 1.31(0.64)             | [1.11,1.52]   |
| FF                  | 1.21(0.86)        | [0.90,1.52]   | 1.21(0.96)             | [1.13,1.74]   |
| MF                  | 1.33(1.01)        | [0.96,1.69]   | 1.43(0.83)             | [1.17,1.69]   |
| Confidence (%)      | 76(8)             | [73,79]       | 75(7)                  | [73,78]       |
| Male Participants   |                   |               |                        |               |
| Correct responses   |                   |               |                        |               |
| Total               | 29.77(2.46)       | [28.43,31.10] | 28.20(3.98)            | [26.64,29.76] |
| FF                  | 14.07(1.61)       | [13.20,14.07] | 13.24(2.13)            | [12.41,14.07] |
| MF                  | 15.70(2.39)       | [14.39,16.99] | 14.96(2.78)            | [13.87,16.05] |
| $d'$                |                   |               |                        |               |
| Total               | 1.52(0.54)        | [1.22,1.81]   | 1.30(0.63)             | [1.05,1.55]   |
| FF                  | 1.56(0.66)        | [1.20,1.92]   | 1.30(0.85)             | [0.96,1.63]   |
| MF                  | 1.70(0.91)        | [1.21,2.19]   | 1.56(1.06)             | [1.15,1.98]   |
| Confidence (%)      | 80(8)             | [75,85]       | 80(10)                 | [76,84]       |

*Note.* CI = confidence interval; FF = female face; MF = male face;  $d'$  =  $d$  prime.

However, Table 4 shows that male participants accurately recalled more male faces across both conditions. The difference between  $d'$  values was found to be insignificant following a matched pairs t-test analysis ( $t(101) = 1.00, p = 1.00, d = 0.20, 95\% \text{ CI } [-0.15, 0.60]$ ). A matched pairs t-test of the total hits and correct rejections, however, uncovered a statistically significant, medium, positive difference between male and female faces recall ( $t(101) = 2.99, p = .021, d = 0.56, 95\% \text{ CI } [0.70, 2.67]$ ). This finding supports a gender in-group effect for male participants.

Refer to Table 4 for statistics related to participants' confidence ratings according to gender and condition. Females in the experimental condition recorded a lower average confidence rating than females who read the article on wage growth (control). While this difference was directional to that hypothesised in H3a an independent t-test analysis found this to be insignificant ( $t(101) = 0.11, p = 1.00, d = 0.03, 95\% \text{ CI } [-3.74, 4.21]$ ).

Additionally, male participants' confidence ratings remained consistent across the two conditions (Table 4). Therefore, hypothesis H3b was also not supported by the data.

Interestingly, an independent samples t-test analysis found there was a statistically significant, medium, positive difference in confidence ratings between genders ( $t(101) = 4.73, p = .013, d = 0.54, 95\% \text{ CI } [1.48, 8.38]$ ). This indicated that male participants recorded a higher confidence percentage ( $M = 80, SD = 9, 95\% \text{ CI } [77, 83]$ ) than female participants ( $M = 75, SD = 8, 95\% \text{ CI } [73, 77]$ ) over the two conditions.

## Discussion

This study aimed to build on previous research investigating the impact of both face and participant gender on facial recall accuracy. Additionally, the study worked to explore the impact of mild psychological distress on conforming to gender in-group biases and the influence this has on recall and confidence ratings. Research in this area is crucial in formulating a more in-depth understanding of societal impacts on cognition and eyewitness

testimony reliability. The results found there to be no significant differences in support of the hypotheses.

While the data was directional to that proposed in H2a, H2b, and H3a these were all non-significant in nature. Similarly, although the results were inverse to that expressed in the first hypothesis this was not statistically significant. There was also no impact on men's confidence between conditions as predicted in H3b. Finally, there was no statistically significant difference revealed between the recall rates of face gender in the control condition. However, the results indicated that participants were minimally more accurate in recalling male faces compared to female faces, contrary to what was hypothesised in H1b.

### **Face Recall Accuracy**

The findings relating to the first hypothesis directly contradicted expectations, with male control participants recording higher accuracy rates than female control participants. While non-significant the effect size of the difference between male and female participants in the control condition was moderate, suggesting the test may have been underpowered. The number of male participants recruited into the control condition ( $n = 13$ ) did not meet the recommended cell total of 20 for data analysis (Simmons et al., 2011). Additionally, a post-hoc calculation of power on this data finds there is a 38% chance the hypothesis has been incorrectly rejected (type II error). This supports the notion that the test was underpowered and therefore not completely reliable. Regardless, these findings challenge previous research that consistently demonstrates women outperform men in facial recall tasks (Herlitz & Loven, 2013).

On the other hand, the non-significant effect of participant gender on facial recall accuracy reflects outcomes uncovered by Man and Hills (2016). Contrary to previous studies, the researchers' experiment on eye-tracking behaviours in recall tasks, found that male and female participants did not significantly differ in processing of faces. Instead, the researchers

found that participants attended to facial features at relatively the same level of importance regardless of their gender (Man & Hills, 2016). Furthermore, prior research has indicated that own-gender bias is largely associated with the encoding process rather than underlying neural mechanisms (Wolff et al., 2014). Considering this alongside the findings supporting gender similarities while encoding faces, it is understandable that accuracy results would yield insignificant differences.

Additionally, these results may be reflective of the participant demographics. Previous research has suggested that female proficiency in facial recall tasks may be a result of their superior perception and comprehension of non-verbal social cues (Pavlova et al., 2016). Given the small sample size is compiled largely of psychology undergraduate students the data may be representing a more detail orientated and perceptive male cohort in comparison to the larger population. Furthermore, as previously discussed, high IQ has also been associated with recall accuracy across in- and out-groups (Hills et al., 2020). The underlying traits which perpetuated the participants to undertake their university degree as well as psychology units on counselling skills and recognition of non-verbal cues may have contributed to this outcome. Therefore, it is possible that the current results are reflective of male learned abilities and natural tendencies bridging the gender gap and are perhaps not generalisable to the wider population.

Furthermore, previous face recall studies have been completed under laboratory conditions with participants attending in person to complete the task (Pavlova et al., 2020; Rehnman & Herlitz, 2007; Wang et al., 2020). This model allows researchers to remove the confounding divided attention variable. Therefore, by the current study being conducted online with participants completing the tasks in their own time and place of choosing, this cannot be controlled. However, this does make the results more akin to events later requiring an eye-witness testimony. In instances of crime individuals are likely to encounter an

abundance of competing information. This could include loud noises, multiple perpetrators, other bystanders, or weapons. Research has found recall accuracy to be reduced when a weapon is present at the scene (Carlson et al., 2016). Prior research has also demonstrated that divided attention alone leads to significantly reduced recall accuracy (Naveh-Benjamin & Brubaker, 2019). Consequently, these results suggest credibility previously given to women, in comparison to men, in face recall may not be warranted in scenarios outside of experimental settings.

The finding that control participants correctly recalled non-significantly more male faces than female faces contradicts previous outcomes. Firstly, it is important to consider whether the face stimuli used played a role in the results uncovered. Research demonstrates that participants who fixate on differentiating facial features, such as the eyes in ethnically White faces (Hills et al., 2020), display enhanced later recall. There has been some evidence suggesting women may focus more on male noses than female noses and that both genders focus more on the eyes of faces of their own gender compared to eyes of other gendered faces (Man & Hills, 2016). It may be that the current results are reflecting the distinctive features of the stimulus set utilised. Perhaps the eyes of the faces were not as easily distinguishable as larger features such as noses.

Additionally, the attractiveness of the faces used within the study could have influenced the results. Previous literature shows that the participant's rating of attractiveness of a presented face is positively associated with later recall accuracy and has been seen to eradicate female gender effects (Malloy et al., 2021). Moreover, the distinguishability of the faces may not have been as strong for female faces as male faces perpetuating the finding of a higher male face accuracy. However, the faces were collected from a pool of previously composed images depicting a range of individuals, all of roughly the same age and of neutral

expressions. As the images were not computer generated, they could be more generalisable to everyday situations whereby people are required to later recall faces (e.g., eyewitnesses).

This data may also be indicative of societal changes. As discussed earlier, infants demonstrate an early preference for faces of the same gender group to their primary caregiver (Marquis & Sugden, 2019). Traditionally early infancy was expected to be dominated by interactions with the mother as primary caregiver. This is becoming less prominent with fathers becoming increasingly involved (Polivanova, 2018). Possibly the results of this study reflect changes in infant interactions, leading to reduced familiarity effects and nullified in-group gender effects during facial recognition tasks. While there is still a gender difference evident within parental childcare arrangements in Australia (Australian Bureau of Statistics, 2020) males are expressing a desire to spend increased time with their children (Harrington et al., 2017). This suggests that infant care arrangements may continue to equalise into the future. Therefore, the impact this may have on future facial recall tasks, based on the present findings, could grow.

Previously, research has proposed that participants demonstrated increased recall accuracy for female faces when compared to male faces, due to increased societal pressures regarding female physical attractiveness (Palmer et al., 2013). It is believed that this emphasis on appearance may lead to increased attention dedicated to female facial features, allowing for easier later recall. However, society may be becoming more attentive to male physical appearance. A recent study found that the majority of Instagram posts utilising hashtags promoting “fitspiration” were depicting lean muscular males (Gultzow et al., 2020). Furthermore, idealised and often unachievable body standards for men are appearing increasingly frequently in films (Roberts et al., 2021), advertisement (Hanan et al., 2021), and even children’s programs (Gonzalez et al., 2020). Research has shown that exposure to sexualised male images via social and popular media has a negative impact on men’s body

image (Hanan et al., 2021; Sumter et al., 2021) indicating awareness of the pressures imposed. Therefore, individuals may now be attending to male faces at a higher rate, perpetuating increased recall accuracy supported by the current study.

The results indicated that women did recall more female faces after reading the gender pay gap article (experimental) compared to those who read the pay growth article (control), however this was nonsignificant. This demonstrates that the article may have worked, to an extent, in making gender differences salient and supporting female in-group affiliation. Interestingly, the  $d'$  value remained consistent across conditions suggesting that the manipulation did not impact female participants' false alarm rate. Suggestive that the  $d'$  statistic may be more reliable than the sum of hits and correct rejections in recall tasks. The non-significance of these results could be associated with various factors including the manipulation, participant pool, or societal setting.

The manipulation, via reading a news article on the gender pay gap in Australia, may not have been sufficient to facilitate an in-group effect. As discussed earlier, the existing literature indicates participants need to perceive the in-group categorisation as a sufficient source of community polarisation for an effect to be established (Harrison et al., 2020). It may be that the topic of the gender pay gap was not sufficient to elicit participant in-group affiliation, perhaps due to it being so widely acknowledged. Therefore, constructing an article based on more novel impacts of gender inequality may have induced a greater division between conditions. However, the article presented to participants was based on readily available media in an effort to simulate that encountered in everyday life. Furthermore, those who are motivated to engage with media reporting lesser-known gender inequality statistics may already hold higher biases, potentially leading to increased in-group membership. As a result, the findings that consumption of widely acknowledged media, regarding negative

gender differences, does not significantly impact the rate at which women recall female faces may be more readily generalisable to the wider population.

Additionally, Mishra and colleagues (2019) recently conducted a study examining the impact of gender on facial recall accuracy recruiting 417 participants across 10 countries. The researchers discovered that the gender in-group effect was not evident in participants from countries with high rates of gender equality. Presently, Australia is ranked as number 50 of 156 in the Global Gender Gap Index (World Economic Forum, 2021) indicating that this may not be as substantial of an issue. However, gender inequalities remain a source of contention within society and popular media. The current participant group, however, was not limited to Australian individuals instead including a range of ethnicities. Therefore, participants from more divisive countries may have been more influenced by the manipulation article and conformed at a higher rate to their gender in-group during the recognition task. Alternatively, participants from countries with higher levels of gender inequality may have read the article on pay gap concerns in Australia and not associated this as sufficient community polarisation, given their own experiences. This suggests that individual differences amongst the participants may have impacted the degree of manipulation salience. It is possible that this contributed to the insignificant result.

Moreover, the finding that women did not recall more female faces than male faces may be representative of divided attention. Palmer and colleagues (2013) investigated the impacts of divided attention on facial recall through requiring participants listen to a series of tones and categorise the sounds based on pitch while viewing a set of 20 faces. The researchers found that in the divided attention condition female participants' recall of female faces was more severely impacted than their recall of male faces. As previously discussed, participants in the current study completed the task remotely meaning that external attentional demands could not be accounted for. Additionally, participants were required to read an

article and respond to questions related to the information prior to the face learn phase. The intention of the experimental article was to induce a small level of psychological distress and encourage assimilation to own gender. Given that the articles were designed to impact participants while undertaking the encoding of faces this may have resulted in a level of divided attention, even if not conscious to the participant. This may account for the reduced rate of recall of female faces.

However, the results showing women recalled more male faces than female faces, although not significant, are inconsistent with previous literature. Research has overwhelmingly found women are better at remembering female faces (Rehman & Herlitz, 2007). As discussed previously potential factors impacting this could be increased intimacy in female friendships (Buhrke & Fuqua, 1987), physical appearance pressures (Palmer et al., 2013), and higher perceptive abilities (Pavlova et al., 2016). Although, as indicated earlier, society may be progressing away from such gender divides with increasing emphasis being placed on male appearance. Parallel to the current results, Mishra and colleagues (2019) did not uncover a female gender bias in their facial recall study. The researchers suggested that these findings are representative of sociocultural impacts such as reduced inequality between genders. Furthermore, even though a female gender in-group was not exhibited the researchers did find evidence of male gender bias (Mishra et al., 2019).

In the present study, while men did recall significantly more male faces than female faces, the rate of this only minimally differed between conditions. Furthermore, those in the control condition had a higher accuracy rate for male faces than those in the experimental condition, opposing the original hypothesis. The impact of divided attention may also explain this finding with the additional stressor of the gender pay gap article leading to reduced attention during encoding. On the other hand, this may be due to the gender in-group effect already providing a strong bias within male participants, making further manipulations

meaningless. However, men did display a reduced recall ability for female faces after reading the gender pay gap article. While this finding was insignificant it suggests that the manipulation may have worked in reducing attentional motivations for their gender out-group.

Evidence for a gender in-group effect for men is supported by some prior research (Mishra et al., 2019; Wright & Sladden, 2003). While attempting to understand the underlying mechanisms of this bias, Wolff and colleagues (2014) uncovered male participants rated own gender faces as more distinctive than female faces. This mimics findings in other in-group effects such as race biases (Wan & Crookes, 2017). Wright and Sladden (2003) also uncovered that the presence of hair in images increased the gender bias effect, perhaps due to providing a proficient memory cue. The current study utilised facial images that were inclusive of hair, thus potentially encouraging later increases in recall, particularly in men for male faces.

Research also suggests the in-group effect is largely related to the participant's self-identity and perceived social membership (Wolff et al., 2014). Identity is multifaceted with some components more impactful to the individual than others (Drummond, 2021). This is demonstrated when Australian third generation Greek immigrants strongly self-identify as Greek (Papadelos, 2021). Suggesting the identity aspect of being Greek remains significant even with generational separation. A strong affiliation with a specific element of identity may lead to the formation of an in-group. This may explain the previously discussed findings that second-generation immigrants to the United States record higher race in-group bias during facial recall tasks compared to first-generation immigrants (Ng et al., 2020). This, in conjunction with men's tendencies to belong to large social circles (Mjaavatn et al., 2016), may explain the results displayed. Men may put extra weighting towards the male aspect of their identity in comparison to women's level of importance to identifying as female. This

may be represented in men's preference for large social circles as opposed to women's memberships in intimate friendship groups (Seeley et al., 2003). In turn, this may lead males to conform with the gender in-group bias and more accurately recall male faces during facial recall tasks, such as this one.

As discussed earlier, recent research has uncovered men who self-reported as not being in a relationship recall significantly more female faces (El Haj et al., 2019). El Haj and colleagues (2019) also found that men who were in a relationship recalled more male faces than female faces, although the difference was not significant. The researchers proposed this occurred due to single males being motivated by primal mating instincts, leading to increased attentional effort applied to fixating on the faces of the opposite gender. In the current study more male participants indicated that they were in a relationship ( $n = 26$ ) than reported being single ( $n = 10$ ). Therefore, the improved performance seen in recalling male faces (compared to female faces) may be attributed to the majority displaying a reduced motivation in attending to female faces.

### **Confidence Ratings**

The data showed a minor decrease in women's confidence after reading the gender pay gap article compared to women who read the wage growth article. While this change was not statistically significant it may signify an impact of gender inequality salience on confidence. Prior studies have uncovered a strong correlation between societal gender-inequalities and gender disparities in mental health rates (Yu, 2018). Bracke and colleagues' (2020) study of over 116,000 participants across 29 countries found that the relationship between gender inequality and depression rates was cumulative with age. Providing evidence that accrued knowledge of inequalities can impact mental wellbeing. Unfortunately, due to the nature of gender pay gaps this illuminates discrimination against women. Therefore, these findings are supportive of the experimental condition article having elicited a level of

psychological distress within women. Furthermore, given the evidence that increased psychological distress is associated with decreased self-esteem (Thompson et al., 2019), the distress caused by the gender pay gap article could have led to the reduction in confidence, evidenced in the current results.

As previously discussed, it is widely accepted that identifications are more reliable when an eyewitness provides a high confidence rating directly following recall (Smalarz et al., 2021). It is crucial to consider this in light of the current study's finding that confidence scores appear somewhat malleable when the individual is experiencing psychological distress. The current result, while not statistically significant, indicates that lesser confidence ratings in identification scenarios may not be an accurate indicator of a false alarm. Especially critical as female participants in the experimental performed better in facial recall (across both male and female faces) than in the control condition but still recorded lower confidence. This is indicative of the confidence-accuracy relationship being of limited value when psychological distress is present.

The outcome of women in the control condition being non-significantly less accurate but also more confident compared to those in the experimental condition contradicts the confidence-accuracy effect. This finding challenges those presented by Davis and colleagues (2019) whereby participants in the psychological distress condition recorded lower confidence but also displayed lower accuracy, therefore adhering to confidence-accuracy expectations. Instead, the current study shows evidence of a more fluid intelligence explanation to confidence ratings (Cattell, 1963) as lower confidence did not reflect lower accuracy in females. The confidence rating was instead seemingly more associated with stressors. This is inconsistent with the inflexible expectations of a crystallised intelligence model (Cattell, 1963). Additionally, the finding that men recorded their confidence consistently high, regardless of their decreased accuracy in the experimental condition, may

suggest a fluid ability. Moreover, Burns and colleagues (2016) provided an explanation for this in their finding that older adults were significantly more likely to over rate confidence in comparison to accuracy. Considering that the researchers classified young adults as under 30 (Burns et al., 2016) and the men in the current study's experimental condition had a mean age of 38.24 this seems to align with the prior research results.

Discrepancies within confidence rating patterns between genders were evident. Women displayed a decrease in confidence from the control to experimental condition, however this was not statistically significant. While men recorded no change in confidence between conditions. Given the participants performed to a high level of accuracy the corresponding high confidence fits with the continuous model expected in the accuracy-confidence assumption. However, both genders also displayed elements of fluid abilities with females decreasing in confidence based on psychological distress and men maintaining overly high confidence regardless of reduced accuracy. So, while this finding may be moving towards a fluid model of confidence it seems like a more comprehensive effect may be taking place. A general intelligence model (Cattell, 1963) may better describe the changes in confidence evident in the results. Such a model would predict confidence to remain predominantly consistent, based on prior knowledge and therefore conform to the principles of the confidence-accuracy model, while also making allowances for external factors such as psychological stressors. As previously discussed, this model would also work in explaining the results in Davis and colleagues' (2019) where participants' confidence ratings reduced according to accuracy and stressors.

This study discovered that men were significantly more confident than women regardless of article read. As previously discussed, no significant differences between genders in recall accuracy were revealed. Men's high confidence compared to accuracy is echoed in Herbst's (2020) findings that men significantly overrated their own abilities. The

present results, however, indicate that both genders performed at a high level of accuracy with women's confidence relatively reflective of this (Table 4). Therefore, the difference in confidence may be suggestive of internal self-rating processes (Stankov, 2019). For example, considering how likely the response is incorrect versus how likely it is correct. Future research should aim to identify the metacognitive processes underlying confidence ratings and the impact gender has on this. In-depth knowledge in this area would allow development of more effective questions for law enforcement officers working with witnesses, encouraging stable confidence assessments across genders.

On the other hand, men's higher confidence ratings may be more to do with typical personality differences between genders. Research suggests that men generally display higher levels of self-confidence and lower a degree of fear of failing in comparison to women (Risse et al., 2018). This parallels the current study's results indicating men were statistically significantly more confident than women. This finding is supportive of an ongoing trait difference between genders. This is important to note, as it is crucial that leaders and the general population consider gender differences in confidence, particularly in making decisions relating to political elections and employment. Considering metacognitive differences such as self-perceived abilities in community and corporate decisions could reduce inequalities currently evident in salary and upper management statistics (Workplace Gender Equality Agency, 2020).

### **Implications**

While the discovery that within the control condition men outperformed women in recall accuracy was not significant, it is an important development to previous research. As men are appearing to become increasingly sexualised in popular media (Gultzow et al., 2020) previous physical attractiveness inequalities may be diminishing. This may also be attributed to men increasingly attending to facial features. Future research would benefit from exploring

a causal effect in relation to these proposals. Additionally, as the findings may be related to decreased mental wellbeing, researchers should aim to develop effective interventions to assist in combating this.

The finding that participants recalled more male faces than female faces was only significant when not accounting for false alarm rates. In eyewitness scenarios the implications of this result are concerning as this is synonymous to false accusation. As discussed previously misidentification via eyewitness testimony is responsible for the most convictions later overturned due to discoveries of new and conflicting evidence (Innocence Project, 2021). Therefore, the current results should encourage caution and future research should consider implementing  $d'$  statistics within all facial recognition studies.

The notion of the in-group bias being exhibited by men but not women implies the effect is based on self-perception of group membership and expands on previous research (Harrison et al., 2020). As discussed previously this may have been evident due to men's affiliation with being male and masculine rather than a more intimate social circle. This provides insights into why in-group bias is not always reliable especially in those trying to integrate into new community groups (Ng et al., 2020). The field would benefit from future studies investigating this effect through both quantitative and qualitative designs. Potentially utilising self-reporting measures to isolate integral components of participants identity. This would allow for a more cohesive understanding of the cognitive processes behind in-group conformity.

As discussed previously, the potential presence of psychological distress may have led to the marginal reduction in female confidence between genders. This is highly applicable to the confidence ratings in eyewitness testimonies. Considering, the potentially trauma inducing nature of observing a crime as well as the stress of later recounting the events, psychological stress seems inevitable. As the current study indicates distress may reduce

confidence ratings, they may not be an effective assessment identification validity.

Furthermore, this research suggests that men may be prone to overrating accuracy. Therefore, in the case of eyewitness testimonies high confidence ratings may not be indicative of a correct identification. Future researchers should consider these findings and progress towards a more reliable assessment of recall. Moreover, law officials should display caution when accepting (or rejecting) eyewitness recall based on confidence and provide education to jurors to reduce risk of false conviction.

The current research is also supportive of a general ability model (Cattell, 1963) of confidence. This calls into question the validity of the confidence-accuracy theory as the results were inconsistent with a continuous effect. Instead, the results indicate that confidence is impacted by outside factors (e.g., stressors) as well as fundamental prior knowledge. Future studies should work on expanding this notion through assessing confidence over varying tasks encompassing memory and other cognitive assessments.

Considering gender inequalities are still evident within society it is crucial to consider ways in which disparities can be minimised. Therefore, the finding that women were significantly less confident than men could be implemented practically, particularly in the corporate sphere. An opportunity to do this is via developing educational tools to encourage deeper knowledge about women's tendency to rate their performance significantly lower than men. Administering this to community leaders and employers could encourage more insightful assessments of potential new recruits, self-evaluations, and requests for salary increases.

### **Limitations**

There were several limitations to this study that warrant addressing. Firstly, as previously mentioned the sample size for the male participants was insufficient, leading to a

underpower in results. This means the results are potentially representative of a few unique individuals instead of an accurate depiction of the wider community.

Secondly, there was a large disparity between the mean age of the male experimental condition and the other conditions meaning that again the results may be representative of an alternate cohort. Future studies would benefit from setting more restrictive age limitations. Additionally, future researchers should consider conducting a cross-sectional study, to better determine whether age impacts on in-group bias. Conducting such research would provide a more detailed understanding of the influence societal changes have on in-group self-identification and recall accuracy.

Thirdly, the lack of control for external influences such as fatigue and conflicting attentional demands. However, as discussed previously this may actually provide a stronger generalisability to practical implementation such as eyewitness testimonies.

Fourthly, the strength and novelty of the manipulation article could have been heightened to potentially elicit more convincing results. Furthermore, having the control article discuss the impacts of COVID-19 on nation wage growth may not have differed enough from the experimental article. Especially considering the greater impact this has had on women. This decision was made in an attempt to minimise factor differences between conditions. Care was taken to ensure the article was overall positive and inclusive language was used throughout. However, there is a possibility that this confounded the results.

Fifthly, whilst the inclusion of true or false comprehension questions confirmed manipulation was comprehended there was no measure of emotional influence. Therefore, it is not possible to completely ascertain that psychological distress was successfully achieved.

Sixthly, the current study only assessed participants who identified as male or female and there was no consideration given to sexual orientation. Prior research is suggestive of increased recall accuracy within own gendered faces for participants who identify as

homosexual (Steffens et al., 2013). Future researchers should consider collecting data regarding sexual orientation within the personal information questionnaire. Future research could also consider assessing face recall and confidence ratings based on a gender continuum to allow for greater generalisability.

Seventhly, the short delay between encoding and recall, while interluded by a filler activity, is not equivalent to practical situations where a witness may need to attempt an identification after extended periods of time.

Finally, the use of stagnant images instead of videos depicting target movement and mannerisms limits the findings to purely visual recall. This is not aligned to complex scenarios which may allowing for additional memory cues or more competing information that eyewitnesses may encounter.

## **Conclusion**

The learnings from this study are essential in further developing understandings within both theoretical and practical avenues. Although the hypotheses were not supported important evidence for a number of gender differences and societal changes was uncovered. The findings that there are limitations in the confidence-accuracy model and that this may reflect a more general intelligence ability, is essential in recalculating the emphasis the criminal justice system places on eyewitnesses' confidence in identification. Future, research should work to better understand this insight and work towards establishing more accurate ratings of eyewitness identifications. Furthermore, a significant difference in confidence between genders was displayed. These findings suggest future researchers and community leaders should acknowledge differences in metacognition between genders and work to reduce gender inequalities by considering this when selecting career or promotional positions. The study was also suggestive of the changing of societal norms, particularly surrounding male appearance pressures, and the influence this has on cognition. It is important that future

researchers attempt to develop a better understanding of this change and work to implement interventions for men who may be encountering negative psychological consequences.

## References

- Alaei, R., Deska, J. C., Hugenberg, K., & Rule, N. O. (2021). People attribute humanness to men and women differently based on their facial appearance. *Journal of Personality and Social Psychology*. Advanced online publication.  
<https://doi.org/10.1037/pspi0000364>
- Australian Bureau of Statistics. (2020). *Gender indicators, Australia*.  
<https://www.abs.gov.au/statistics/people/people-and-communities/gender-indicators-australia/latest-release#work-and-family-balance>
- Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorisation is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, 18(8), 706-712. <https://doi.org/10.1111/j.1467-9280.2007.01964.x>
- Bracke, P., Delaruelle, K., Dereuddre, R., & Van de Velde, S. (2020). Depression in women and men, cumulative disadvantage and gender inequality in 29 European countries. *Social Science & Medicine*, 267, Article 113354.  
<https://doi.org/10.1016/j.socscimed.2020.113354>
- Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, 73(1), 105-116.  
<https://doi.org/10.1111/j.2044-8295.1982.tb01795.x>
- Bryzgalov, D. V., Kuznetsova, I. L., & Rogaev, E. I. (2018). Enhancement for declarative memory: From genetic regulation to non-invasive stimulation. *Biochemistry (Moscow)*, 83(9), 1124-1138. <https://doi.org/10.1134/S0006297918090146>
- Buhrke, R. A., & Fuqua, D. R. (1987). Sex differences in same- and cross-sex supportive relationships. *Sex Roles*, 17(5-6), 339-352. <https://doi.org/10.1007/BF00288457>

- Burns, K. M., Burns, N. R., & Ward, L. (2016). Confidence – More a personality or ability trait? It depends on how it is measured: A comparison of young and older adults. *Frontiers in Psychology*, 7, Article 518. <https://doi.org/10.3389/fpsyg.2016.00518>
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22. <https://doi.org/10.1037/h0046743>
- Ch'ng, K. (2021). Attractiveness, trust and trustworthiness: An experimental study. *Malaysian Journal of Economic Studies*, 58(1), 45-57. <https://doi.org/10.22452/MJES.vol58no1.3>
- Carlson, C. A., Pleasant, W. E., Weatherford, D. R., Carlson, M. A., & Bednarz, J. E. The weapon focus effect: Testing an extension of the unusualness hypothesis. *Applied Psychology in Criminal Justice*, 12(2), 87-100. <https://psycnet.apa.org/record/2017-07766-002>
- Davey, M. (2021, March 26). 'Gender apathy': Australian pay gap unlikely to close for 26 years due to views that 'devalue' women – report. *The Guardian*. <https://www.theguardian.com/australia-news/2021/mar/26/gender-apaty-australian-pay-gap-unlikely-to-close-for-26-years-due-to-views-that-devalue-women-report>
- Davis, S. D., Peterson, D. J., Wissman, K. T., & Slater, W. A. (2019). Physiological stress and face recognition: Differential effects of stress on accuracy and the confidence-accuracy relationship. *Journal of Applied Research in Memory and Cognition*, 8(3), 367-375. <https://doi.org/10.1016/j.jarmac.2019.05.006>
- Delay, C. G., & Wixted, J. T. (2021). Discrete-state versus continuous models of the confidence-accuracy relationship in recognition memory. *Psychonomic Bulletin & Review*, 28, 556-564. <https://doi.org/10.3758/s13423-020-01831-7>

- Denkinger, B., & Kinn, M. (2018). Own-age bias and positivity effects in facial recognition. *Experimental Aging Research*, 44(5), 411-426.  
<https://doi.org/10.1080/0361073X.2018.1521493>
- Drummond, J. J. (2021). Self-identity and personal identity. *Phenomenology and the Cognitive Sciences*, 20, 235-247. <https://doi.org/10.1007/s11097-020-09696-w>
- Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, 111(2), 524-542. <https://doi.org/10.1037/0033-295X.111.2.524>
- El Haj, M., Moustafa, A. A., & Nandrino, J. (2019). Singles and faces: High recognition for female faces in single males. *Advances in Cognitive Psychology*, 15(4), 301-307.  
<https://doi.org/10.5709/acp-0277-x>
- Gomez, P., Von Gunten, A., & Danuser, B. (2020). Recognizing images: The role of motivational significance, complexity, social content, age, and gender. *Scandinavian Journal of Psychology*, 61(2), 183-194. <https://doi.org/10.1111/sjop.12593>
- Gonzalez, M. P. L., Infantes-Paniagua, A., Thornborrow, T., & Jordan, O. C. (2020). Associations between media representations of physical, personality, and social attributes by gender: A content analysis of children's animated film characters. *International Journal of Communication*, 14, 6026-6048.  
<https://ijoc.org/index.php/ijoc/article/view/16149>
- Graves, L. V., Moreno, C. C., Seewald, M., Holden, H. M., Van Etten, E. J., Uttarwar, V., McDonald, C. R., Delano-Wood, L., Bondi, M. W., Woods, S. P., Delis, D. C., & Gilbert, P. E. (2017). Effects of age and gender on recall and recognition discriminability. *Archives of Clinical Neuropsychology*, 32(8), 972-979.  
<https://doi.org/10.1093/arclin/acx024>

- Gueguen, N., Fischer-Lokou, J., & Lefebvre, L. (2008). Women's eye contact and men's later interest: Two field experiments. *Perceptual and Motor Skills*, 106(1), 63-66.  
<https://doi.org/10.2466/PMS.106.1.63-66>
- Gultzow, T., Guidry, J., Schneider, F., & Hoving, C. (2020). Male body image portrayals on Instagram. *Cyberpsychology, Behavior, and Social Networking*, 23(5).  
<https://doi.org/10.1089/cyber.2019.0368>
- Hanan, M. A., Arshad, A., & Saleem, N. (2021). Media and other socializing agents influence on male body-shaping behavior: Body esteem as a mediator. *Asia-Pacific Social Science Review*, 21(1), 115-126. <http://apssr.com/wp-content/uploads/2021/03/RA-9.pdf>
- Hare, J. (2020, July 21). Yes, women outnumber men at university. But they still earn less after they leave. *The Conversation*. <https://theconversation.com/yes-women-outnumber-men-at-university-but-they-still-earn-less-after-they-leave-142714>
- Harrington, B., Fraone, J. S., & Lee, J. (2017). *The new dad: The career-caregiving conflict*. Boston College Carroll School of Management: Centre for Work & Family.  
<http://www.bc.edu/content/dam/files/centers/cwf/research/publications/researchreports/BCCWF%20The%20New%20Dad%202017.pdf>
- Harrison, V., Hole, G., & Habibi, R. (2020). Are you in or are you out? The importance of group saliency in own-group biases in face recognition. *Perception*, 49(6), 672-687.  
<https://doi.org/10.1177/0301006620918100>
- Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology*, 46(2), 445-448.  
<https://doi.org/10.1016/j.jesp.2009.11.008>

- Herbst, T. H. H. (2020). Gender differences in self-perception accuracy: The confidence gap and women leaders' underrepresentation in academia. *SA Journal of Industrial Psychology*, 46, Article a1704. <https://doi.org/10.4102/sajip.v46i0.1704>
- Herlitz, A., & Loven, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition*, 21(9-10), 1306-1336. <https://dx.doi.org/10.1080/13506285.2013.823140>
- Hernandez-Julian, R., & Peters, C. Student appearance and academic performance. *Journal of Human Capital*, 11(2), 247-262. <https://doi.org/10.1086/691698>
- Hills, P. J., Lowe, L., Hedges, B., & Teixeira, A. R. (2020). The role of extraversion, IQ and contact in the own-ethnicity face recognition bias. *Attention, Perception, & Psychophysics*, 82, 1872-1882. <https://doi.org/10.3758/s13414-019-01947-6>
- Hills, P. J., Pake, J. M., Dempsey, J. R., & Lewis, M. B. (2018). Exploring the contribution of motivation and experience in the post pubescent own-gender bias in face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 44(9), 1426-1446. <https://doi.org/10.1037/xhp0000533>
- 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>
- Hutchens, G., Pupazzoni, R., & Nogel, R. (2021, March 19). Unemployment rate falls to 5.8 per cent, down from 6.3 per cent. *ABC News*. <https://www.abc.net.au/news/2021-03-18/unemployment-rate-falls-to-5.8-per-cent/13258872>
- Innocence Project. (2021, August 25). *Eyewitness identification reform*. <https://innocenceproject.org/eyewitness-identification-reform/>

- Janda, M., & Pupazzoni, R. (2021, February 24). Wages growth stuck at record low as public sector pay freezes, award delays bite. *ABC News*. <https://www.abc.net.au/news/2021-02-24/wage-price-index-abs/13187102>
- Kwang, T., Crockett, E. E., Sanchez, D. T., & Swann, W. B. (2013). Men seek social standing, women seek companionship: Sex differences in deriving self-worth from relationships. *Psychological Science*, 24(7), 1142-1150.  
<https://doi.org/10.1177/0956797612467466>
- Lin, Y., Ding, H., & Zhang, Y. (2021). Gender differences in identifying facial, prosodic, and semantic emotions show category- and channel-specific effects. *Journal of Speech, Language, and Hearing Research*, 64(8), 2941-2955.  
[https://doi.org/10.1044/2021\\_JSLHR-20-00553](https://doi.org/10.1044/2021_JSLHR-20-00553)
- Maitland, S. B., Herlitz, A., Nyberg, L., Backman, L., & Nilsson, L. (2004). Selective sex differences in declarative memory. *Memory & Cognition*, 32(7), 1160-1169.  
<https://doi.org/10.3758/BF03196889>
- Malloy, T. E., DiPietro, C., DeSimone, B., Curley, C., Chau, S., & Silva, C. (2021). Facial attractiveness, social status, and face recognition. *Visual Cognition*, 29(3), 158-179.  
<https://doi.org/10.1080/13506285.2021.1884630>
- Man, T. W., & Hills, P. J. (2016). Eye-tracking the own-gender bias in face recognition: Other-gender faces are viewed differently to own-gender faces. *Visual Cognition*, 24(9-10), 447-458. <https://doi.org/10.1080/13506285.2017.1301614>
- Marquis, A. R., & Sugden, N. A. (2019). Meta-analytic review of infants' preferential attention to familiar and unfamiliar face types based on gender and race. *Developmental Review*, 53, Article 100868. <https://doi.org/10.1016/j.dr.2019.100868>

- McBride, D. M., Beckner, J. K., & Abney, D. H. (2011). Effects of delay of prospective memory cues in an ongoing task on prospective memory task performance. *Memory & Cognition*, 39, 1222-1231. <https://doi.org/10.3758/s13421-011-0105-0>
- Megreya, A. M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: Evidence from matching tasks. *Acta Psychologica*, 137(1), 83-89. <https://doi.org/10.1016/j.actpsy.2011.03.003>
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3-35. <https://doi.org/10.1037//1076-8971.7.1.3>
- Mishra, M. V., Likitlersuang, J., Wilmer, J. B., Cohan, S., Germine, L., & DeGutis, J. M. (2019). Gender differences in familiar face recognition and the influence of sociocultural gender inequality. *Scientific Reports*, 9, Article 17884. <https://doi.org/10.1038/s41598-019-54074-5>
- Mjaavatn, P. E., Frostad, P., & Pijl, S. J. (2016). Adolescents: Differences in friendship patterns related to gender. *Issues in Educational Research*, 26(1), 45-64. <https://search.informit.org/doi/epdf/10.3316/ielapa.020662409350941>
- Naveh-Benjamin, M., & Brubaker, M. S. (2019). Are the effects of divided attention on memory encoding processes due to the disruption of deep-level elaborative processes? Evidence from cued- and free-recall tasks. *Journal of Memory and Language*, 106, 108-117. <https://doi.org/10.1016/j.jml.2019.02.007>
- Ng, A. H., Steele, J. R., Sasaki, J. Y., & George, M. (2020). How robust is the own-group face recognition bias? Evidence from first- and second-generation East Asian Canadians. *PLoS ONE*, 15(5), Article e0233758. <https://doi.org/10.1371/journal.pone.0233758>

- Palmer, M. A., Brewer, N., & Horry, R. (2013). Understanding gender bias in face recognition: Effects of divided attention at encoding. *Acta Psychologica*, 142, 362-369. <http://dx.doi.org/10.1016/j.actpsy.2013.01.009>
- Papadelos, P. (2021). "Greeks are different to Australians": Understanding identity formation among third-generation Australians of Greek heritage. *Ethnic and Racial Studies*, 44(11), 1975-1994. <https://doi.org/10.1080/01419870.2020.1813318>
- Pavlova, M. A., Mayer, A., Hosel, F., & Sokolov, A. N. (2016). Faces on her and his mind: Female and likable. *PLoS ONE*, 11(6), Article e0157636. <http://doi.org/10.1371/journal.pone.0157636>
- Pisoni, A., Turi, Z., Raithel, A., Ambrus, G. G., Alekseichuk, I., Schacht, A., Paulus, W., & Antal, A. (2015). Separating recognition processes of declarative memory via anodal tDCS: Boosting old item recognition by temporal and new item recognition by parietal stimulation. *PLoS ONE*, 10(3), Article e0123085. <https://doi.org/10.1371/journal.pone.0123085>
- Polivanova, K. (2018). Modern parenthood as a subject of research. *Russian Education & Society*, 60(4), 334-347. <https://doi.org/10.1080/10609393.2018.1473695>
- Proverbio, A. M. (2017). Sex differences in social cognition: The case of face processing. *Journal of Neuroscience Research*, 95(1-2), 222-234. <https://doi.org/10.1002/jnr.23817>
- Rehman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta Psychologica*, 124(3), 344-355. <https://doi.org/10.1016/j.actpsy.2006.04.004>
- 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. <https://doi.org/10.1037/a0025750>

- Risse, L., Farrell, L., & Fry, T. R. (2018). Personality and pay: Do gender gaps in confidence explain gender gaps in wages?. *Oxford Economic Papers*, 70(4), 919-949.  
<https://doi.org/10.1093/oep/gpy021>
- Roberts, L., Aubrey, J. S., Teran, L., Dajches, L., & Ward, L. M. (2021). The super man: Examining associations between childhood superhero imaginative play and wishful identification and emerging adult men's body image and gender beliefs. *Psychology of Men & Masculinities*, 22(2), 391-400. <https://doi.org/10.1037/men0000335>
- Rutishauser, U. (2019). Testing models of human declarative memory at the single-neuron level. *Trends in Cognitive Sciences*, 23(6), 510-524.  
<https://doi.org/10.1016/j.tics.2019.03.006>
- Seeley, E. A., Gardner, W. L., Pennington, W. L., & Gabriel, S. (2003). Circle of friends or members of a group? Sex differences in relational and collective attachment to groups. *Group Processes & Intergroup Relations*, 6(3), 251-263.  
<https://doi.org/10.1177/13684302030063003>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359-1366.  
<https://doi.org/10.1177/0956797611417632>
- Smarlarz, L., Yang, Y., & Wells, G. L. (2021). Eyewitnesses' free-report verbal confidence statements are diagnostic of accuracy. *Law and Human Behavior*, 45(2), 138-151.  
<https://doi.org/10.1037/lhb0000444>
- Sommer, W., Hildebrandt, A., Kunina-Habenicht, O., Schacht, A., & Wilhelm, O. (2013). Sex differences in face cognition. *Acta Psychologica*, 142(1), 62-73.  
<https://doi.org/10.1016/j.actpsy.2012.11.001>

- Stankov, L. (2019). Applied metacognition and separation of confidence and accuracy in correlational studies. *Metacognition and Learning*, 14, 509-516.  
<https://doi.org/10.1007/s11409-019-09212-9>
- Steffens, M. C., Landmann, S., & Mecklenbrauker, S. (2013). Participant sexual orientation matters: New evidence on the gender bias in face recognition. *Experimental Psychology*, 60(5), 362-367. <https://doi.org/10.1027/1618-3169/a000209>
- Sumter, S. R., Cingel, D., & Hollander, L. (2021). Navigating a muscular and sexualized Instagram feed: An experimental study examining how Instagram affects both heterosexual and nonheterosexual men's body image. *Psychology of Popular Media*. Advanced online publication. <https://doi.org/10.1037/ppm0000355>
- Thompson, M. N., Her, P., Fetter, A. K., & Perez-Chavez, J. (2019). College student psychological distress: Relationship to self-esteem and career decision self-efficacy beliefs. *The Career Development Quarterly*, 67(4), 282-297.  
<https://doi.org/10.1002/cdq.12199>
- Tian, Y., Li, L., Yin, H., & Hunag, X. (2019). Gender differences in the effect of facial attractiveness on perception of time. *Frontiers in Psychology*, 10, Article 1292.  
<https://doi.org/10.3389/fpsyg.2019.01292>
- Van der Auwera, S., Terock, J., Teumer, A., Schomerus, G., Homuth, G., & Grabe, H. J. (2021). Sex effects for the interaction of dopamine related genetic variants for COMT and BDNF on declarative memory performance. *Genes, Brain and Behavior*, 20(5), Article e12737. <https://doi.org/10.1111/gbb.12737>
- Wan, L., & Crookes, K. (2017). Face-blind for other-race faces: Individual differences in other-race recognition impairments. *Journal of Experimental Psychology*, 146(1), 102-122. <https://doi.org/10.1037/xge0000249>

- Wang, L. A. L., Herrington, J. D., Tunc, B., & Schultz, R. T. (2020). Bayesian regression-based developmental norms for the Benton Facial Recognition Test in males and females. *Behaviour Research Methods*, 52, 1516-1527. <https://doi.org/10.3758/s13428-019-01331-0>
- 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. <https://doi.org/10.1016/j.biopsycho.2013.07.001>
- Wolff, N., Kemter, K., Schweinberger, S. R., & Wiese, H. (2014). What drives social in-group biases in face recognition memory? ERP evidence from own-gender bias. *Social Cognitive and Affective Neuroscience*, 9(5), 580-590. <https://doi.org/10.1093/scan/nst024>
- Workplace Gender Equality Agency. (2020). *Women in leadership*. Commonwealth Government of Australia, <https://www.wgea.gov.au/women-in-leadership>
- World Economic Forum. (2021). *Global gender gap report: Insight report: March 2021*, <https://www.weforum.org/reports/ab6795a1-960c-42b2-b3d5-587eccda6023>
- Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, 114(1), 101-114. [https://doi.org/10.1016/S0001-6918\(03\)00052-0](https://doi.org/10.1016/S0001-6918(03)00052-0)
- Yu, S. (2018). Uncovering the hidden impacts of inequality on mental health: A global study. *Translational Psychiatry*, 8(1), Article 98. <https://doi.org/10.1038/s41398-018-0148-0>

## **Appendix A**

### **Experimental Condition Article**

#### **Women outnumber men at university but still earn less after leaving**

In 1987 women made up most university enrolments for the first time — now, they make up 55.5%.

While women value education more (and see it as providing financial security) men still better women, after they graduate, in both salary and position.

#### **Why women outnumber men at university**

For every 100 women enrolled in an Australian university, there are 72 men. And once there, men are more likely to drop out.

The drivers behind the increase of women in higher education are a combination of social, cultural, and financial factors.

Looking at the past five decades would point to the rise in feminism and changes in attitudes about women's role in the home.

#### **And yet, women remain worse off**

It's agreed that personal and social benefits come from a degree: higher salaries, better health outcomes, stronger levels of community engagement and lower levels of criminal behaviours, to name a few.

And yet, female university graduates are expected to earn 27% less than men (averaging \$750,000) over their career.

This demonstrates a persistent gender pay gap and men moving up the career ladder more quickly than women, even in female lead fields such as health care and education.

Why is it women fail to make the most their higher-level education compared to men?

Half of all female employments started each year are in female dominated, lower-paid, fields (e.g., teaching, nursing, and childcare) while men outnumber women in two fields only — engineering and IT.

Then there are the issues of how careers are valued (childcare pays poorly but construction well), recruitment practices, and corporate cultures.

And there's the fact more women leave full time work to bring up children. While the number of women staying in the workforce has increased in recent years, at the age of 35 80% of men are employed in the workforce full-time while only 40% of women are.

It is not until their 50s that 50% of women are back in the workforce full time. And this is too late for most to make enough of their own wealth to see them through their retirement years (should their marriage go bust).

What that also means is there is a significant percentage of older women who are part time, unemployed, or underemployed.

Interestingly, the planned changes to tuition fee aids have attracted media attention, in part because they look set to benefit men while negatively impacting women.

That this is an intentional form of policy preference to improve higher education participation among men is unlikely. However, it brings up the question of whether men should be considered a disadvantaged group.

The answer, for the time being at least, is a strong no. Firstly, men are not being refused university places because there are more women — they are making choices based on the options available to them.

And men largely have access to more well-paying career paths that do not need a university degree. Trades, for example, continue to be male dominated and can be well-rewarded, unlike similar jobs for women.

Women also have to compete with the gender pay gap, interrupted careers, and fewer opportunities to enter leadership positions. Because they make the “choice” to be the primary carer, women almost never make it back financially when they return to work.

Workplaces, homes, and institutions continue to hold attitudes towards women that “devalues and discriminates against” them, aggravating a gender pay gap that is unlikely to close for another 26 years (according to a report on workplace equality published on Friday).

“A quarter of the century to close the gender pay gap for full time workers is a long time, but we can get there faster if we start paying more attention to how women and men are paid and supported in our workplaces,”

The report found that community and personal service workers are some of the lowest paid workers in Australia. Even in these low-paid roles women working full-time can expect to earn around 10% less than men.

“Employers in this field need to be more aware of how they are recruiting and rewarding women and men for their work. We also need to reassess the value of these roles more broadly and whether the pay reflects this.”

## **Appendix B**

### **Control Condition Article**

#### **Wages growth stuck at record low as public sector pay freezes**

A stronger than expected increase in private sector pay over the last three months of 2020 has kept wage growth above economists' expectations, as many employers reversed COVID pay cuts or freezes.

Australians' pay packets rose an average of 0.6% over the last three months of last year, double the 0.3% expected.

But many Australians, particularly those relying on awards for pay increases and public servants whose wages have been frozen by governments, are falling behind.

#### **Wage freezes still commonplace**

Wage growth over the year remained at a record low of 1.4%, with an ABS representative advising that was despite several one-off factors lifting the outcome.

"December quarter's moderate growth was impacted by businesses rolling back short-term wage reductions, returning wages to pre-COVID levels."

"The inclusion of the Fair Work Commission annual wage review also had a small positive impact on wages."

Other information from businesses indicates that substantial pay rises are a long way off.

"Although we saw wage cut reversals really boost growth in this quarter more than 60% of businesses either currently have a wage freeze in place or expect to put one in place."

Regions and fields that were recovering from pay cuts had the strongest December quarter figures, with professional, scientific and technical workers booking a 1.2% increase having suffered a half a per cent fall in wages in the June quarter.

### **Recruiters say pay offers improving**

Angela Franks, who recruits for HR roles in Sydney, said she did not see any real wage growth in 2020.

“A lot of people in the sector worked reduced hours or had their pay cut,” she said.

Ms Franks saw about a 50% fall in recruitment for most of last year, which only began reversing in December and into this year.

That fall in the number of jobs on offer has pushed the authority to set wages further into the hands of employers, who are trying to save on costs to offset a dismal year.

However, Ms Franks is expecting the next quarter to show a better lift in wages.

“We’re seeing a lot more recruitment roles coming online and being advertised, which is a really good sign that the job market is going to start heating up, which will help drive wage growth,” she observed.

Workers in the mining sector have the upper hand when it comes to negotiating their wages.

“There’s big demand for technical and professional people and we’re constrained by the supply, so naturally that has pressure on rates,” a Perth recruiter said.

His resources recruiting firm started 2020 with 500 contractors on its books.

By the end of the year, that had grown to 850.

An extra 88,700 people found jobs in February, pushing the number of Australians with jobs above 13 million for the first time in 11 months.

It made Australia's unemployment rate fall to 5.8%, down from 6.3%.

It was a surprisingly large fall and provided evidence the labour market was improving much faster than expected.

Total employment is now just 1,800 below pre-pandemic levels, in seasonally adjusted figures.

The surge in employment in February was well above the consensus estimate of 30,000.

“The jobs lost in the early months of the coronavirus pandemic have now been fully replaced.”

ABS Head of Labour Statistics said the data showed the recovery in Australia's labour market was continuing.

In February, full time jobs increased by 89,100, but part-time jobs fell by 500 positions.

Problematically, the underemployment rate — referring to people with jobs who would prefer more hours — rose from 8.1% to 8.5% in February.

However, Ms Franks is already seeing her clients revise the salaries they are prepared to pay.

A small tourism business she recently recruited for realised they needed to offer more money than they originally wanted, to hire a new recruit.

“They’ve had to increase the salary they were offering by 25% to get that person,” she said.

Despite the lift, it was still a pay cut for the successful applicant.

“The candidate has actually taken about a 20% pay cut to take that role because they love the business and it’s close to their home,” she explained.

“The problem we have is the data doesn’t tell those stories.”

## **Appendix C**

### **Experimental Condition Manipulation Check Questions**

Answer the following true or false questions according to the article:

Question 1: Workers in childcare get paid more than construction workers.

Answer: False

Question 2: There are more female students than male students in Australian universities.

Answer: True

Question 3: It is predicted that the gender pay gap will take twenty-six (26) years to close.

Answer: True

Question 4: The same percentage of males and females are working at the age of thirty-five (35).

Answer: False

Question 5: Community and personal service workers are some of the lowest paid in the workforce.

Answer: True

## **Appendix D**

### **Control Condition Manipulation Check Questions**

Answer the following true or false questions according to the article:

Question 1: Australian workers have had no salary increase due to COVID-19.

Answer: False

Question 2: Salaries averaged an increase double what was expected over the last 12 months.

Answer: True

Question 3: It is suggested that the issue is with a rise of underemployment levels.

Answer: True

Question 4: On average wage growth was of a record high over the last year.

Answer: False

Question 5: Many people needed to work reduced hours or taken a pay cut because of the pandemic.

Answer: True

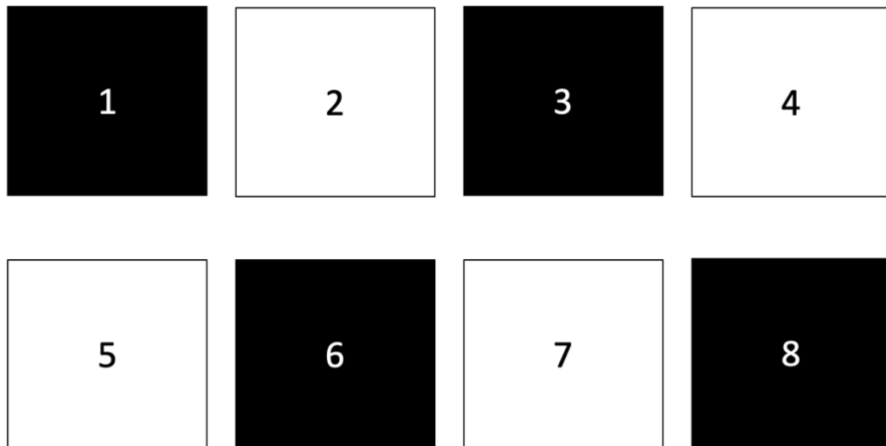
**Appendix E**  
**Examples of Faces Used**



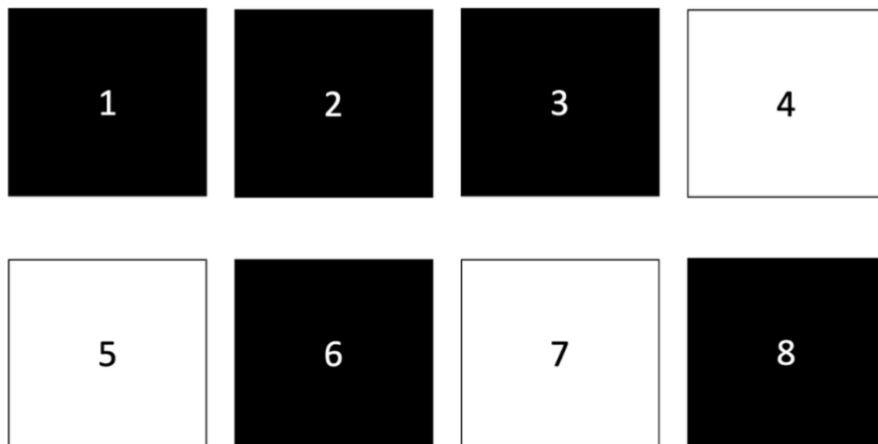
## Appendix F

### Example of Filler Task Activity

#### Learn Phase:



#### Test Phase:



## Appendix G

### Ethics Approval Letter



#### Ethics Approval Letter

14/07/2021

To: Dr Sauer

Project ID: 18677

Project Title: Confidence, Response Time, and Accuracy in Decision Making (H0018677)

The amendment received in support of the above named project has been approved by the University of Tasmania Human Research Ethics Committee on 14 July 2021.

Approval has been granted for the following documentation:

| Submission Document Name                                                   | Submission Document File Name                                                   | Submission Document Type                 | Submission Document Date | Submission Document Version |
|----------------------------------------------------------------------------|---------------------------------------------------------------------------------|------------------------------------------|--------------------------|-----------------------------|
| Sauer Minimal Risk ethics 2020 signed Amended 12 July 2021 - track changes | Sauer Minimal Risk ethics 2020 signed Amended 12 July 2021 - track changes.docx | Application (Tracked)                    | 12/07/2021               | 2                           |
| amended participant information and consent June 2021 - track changes      | amended participant information and consent June 2021 - track changes.docx      | PARTICIPANT INFORMATION AND CONSENT FORM | 12/07/2021               | 3                           |
| amended participant information and consent June 2021 - clean              | amended participant information and consent June 2021 - clean.docx              | PARTICIPANT INFORMATION AND CONSENT FORM | 12/07/2021               | 3                           |
| Sauer Minimal Risk ethics 2020 signed Amended 12 July 2021 - clean         | Sauer Minimal Risk ethics 2020 signed Amended 12 July 2021 - clean.docx         | Application (Tracked)                    | 12/07/2021               | 2                           |
| new materials_text passages                                                | new materials_text passages.docx                                                | Other Documents                          | 12/07/2021               | 1                           |
| cover letter                                                               | cover letter.docx                                                               | Additional Information                   | 12/07/2021               | 1                           |

Please ensure that all investigators involved with this project have cited the approved versions of the documents listed within this letter and use only these versions in conducting this research project.

This approval constitutes ethical clearance by the University of Tasmania Human Research Ethics Committee. The decision and authority to commence the associated research may be dependent on factors beyond the remit of the ethics review process. For example, your research may need ethics clearance from other organisations or review by your research governance coordinator or Head of Department. It is your responsibility to find out if the approvals of other bodies or authorities are required. It is recommended that the proposed research should not commence until you have satisfied these requirements.

In accordance with the [National Statement on Ethical Conduct in Human Research](#), it is the responsibility of institutions and researchers to be aware of both general and specific legal requirements, wherever relevant. If researchers are uncertain they should seek legal advice to confirm that their proposed research is in compliant with the relevant laws. University of Tasmania researchers may seek legal advice from Legal Services at the University.

The University of Tasmania Human Research Ethics Committee (HREC) operates under and is required to comply with the National Statement on the Ethical Conduct in Human Research.

Therefore, the Chief Investigator's responsibility is to ensure that:

- (1) All investigators are aware of the terms of approval, and that the research is conducted in compliance with the HREC approved protocol or project description.
- (2) Modifications to the protocol do not proceed until **approval** is obtained in writing from the HREC. This includes, but is not limited to, amendments that:
  - (i) are proposed or undertaken in order to eliminate immediate risks to participants;
  - (ii) may increase the risks to participants;
  - (iii) significantly affect the conduct of the research; or
  - (iv) involve changes to investigator involvement with the project.

Please note that all requests for changes to approved documents must include a version number and date when submitted for review by the HREC.

- (3) Reports are provided to the HREC on the progress of the research and any safety reports or monitoring requirements as indicated in NHMRC guidance.

Guidance for the appropriate forms for reporting such events in relation to clinical and non-clinical trials and innovations can be located under the ERM "Help Tab" in

"Templates". All adverse events must be reported regardless of whether or not the event, in your opinion, is a direct effect of the therapeutic goods being tested.

(4) The HREC is informed as soon as possible of any new safety information, from other published or unpublished research, that may have an impact on the continued ethical acceptability of the research or that may indicate the need for modification of the project.

(5) All research participants must be provided with the current Participant Information Sheet and Consent Form, unless otherwise approved by the Committee.

(6) This study has approval for four years contingent upon annual review. A Progress Report is to be provided on the anniversary date of your approval. Your first report is due on the anniversary of your approval, and you will be sent a courtesy reminder closer to this due date. Ethical approval for this project will lapse if a Progress Report is not submitted in the time frame provided.

(7) A Final Report and a copy of the published material, either in full or abstract, must be provided at the end of the project.

(8) The HREC is advised of any complaints received or ethical issues that arise during the course of the project.

(9) The HREC is advised promptly of the emergence of circumstances where a court, law enforcement agency or regulator seeks to compel the release of findings or results. Researchers must develop a strategy for addressing this and seek advice from the HREC.

Kind regards,

Ethics Executive Officer



## **Appendix H**

### **PARTICIPANT INFORMATION SHEET**

#### **A Gender In-Group Effect on Facial Familiarity**

This information sheet is for University School of Psychological Sciences Undergraduate students or invited individuals.

1. You are invited to take part in a study exploring the impact of gender on facial recognition tasks. This study is being conducted in partial fulfillment of an Honours degree for Anna Bailey under the supervision of Peter Tranent.

2. The study is examining gender differences in facial recognition. It is hoped that this study will lead to a better understanding of processes affecting decision making and predictors of accuracy.

3. You have been identified through the University of Tasmania Undergraduate databases or advertisement. As this invitation has been supplied to the organising body, your contact information has not been accessible to researchers directly. Your participation in this study is completely voluntary and there are no consequences for declining the invitation.

4. You will be asked to complete the study on a desktop or laptop computer and we suggest in a quiet room to reduce distractions. The study will involve reading an article and responding to two to three comprehension questions. You will then be shown a series of faces. Following this you will complete a short and unrelated activity. Finally, you will be shown another series of faces. For each face you will be asked to indicate if it is “new” or “previously shown” and how confident (50%-100%) you are of your response. The study will

take approximately 30-60 minutes. You will, however, be given regular prompts to take a short break if you wish. The data will be completely anonymous.

5. It is hoped that the study will assist understanding of recognition processes which have practical benefits such as within eye-witness testimonies. Additionally, you may gain insight into your own facial recall abilities. You will also be provided compensation for your time either through allocation of course credit or e-gift card.

6. No foreseeable risks have been identified in this study. If you have any concerns, however, please do not hesitate to contact us through the details provided below.

7. You are free to withdraw at any time without need for explanation. At this time any data you have provided will be removed from any online or hard drive storage systems. As your data is anonymous, once you have completed the study it is not feasible to identify and destroy it so withdrawal will not be possible. You will, however, be reminded of this stipulation at the end of the study, allowing you time to reconsider and withdraw your data from the pool if you wish.

8. All data will be stored by the University of Tasmania, School of Psychological Sciences, for 5 years at which time it will be completely erased from any systems and/or devices. Until this time the data will be kept in password secured computers within the School of Psychological Sciences at the University of Tasmania and only accessible to researchers. All data stored will be completely anonymous in nature.

9. A report of the study's results will be published within an Honours degree thesis.

Additionally, data may be utilised in future publications within a scientific journal. You will not be identifiable in any/all publications. If you would like access to the results of the study, please contact us and a summary can be provided once data collection and analyses have been completed.

10. If you have questions about this study contact:

Anna Bailey [abailey3@utas.edu.au](mailto:abailey3@utas.edu.au)

Peter Tranent [peter.tranent@utas.edu.au](mailto:peter.tranent@utas.edu.au)

This study has been approved by the University of Tasmania Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, you can contact the Executive Officer of the HREC (Tasmania) Network on (03) 6226 6254 or email [human.ethics@utas.edu.au](mailto:human.ethics@utas.edu.au). The Executive Officer is the person nominated to receive complaints from research participants. You will need to quote H0018677.

You may save this information sheet in case you wish to reference it in future.

To be involved in the study, you will need to provide your consent through submission of an anonymous online survey.

Kind regards,

Anna Bailey & Peter Tranent

## Appendix I

### PARTICIPANT CONSENT FORM

- I understand that my involvement in this research will include completing a variety of decision-making tasks
- I understand that the research will include recording of my responses
- I understand that participation involves the risk of fatigue, but that I will be provided with regular opportunities for rest breaks
- Any questions that I have asked have been answered to my satisfaction
- I understand that the results of the study will be published, but that I will not be identifiable as a participant
- I understand that my participation in this research is voluntary
- I understand that I am free to withdraw at any time, without explanation or penalty
- I understand that I will not be able to withdraw my data after completing the research as data will be collected anonymously
- I understand that all study data will be deidentified and then stored on the Open Science Framework for potential further research and review
- I **agree that my study data** can be used for this specific experiment
- I **agree that my de-identified study data can be archived** on the Open Science Framework and used for future research projects in the same general area of this research
- I understand that by clicking the **continue** button I am agreeing to participate in the study

I agree that my study data can be used for this specific experiment:

- Yes – I agree

- ☐ No – I disagree

I agree that my de-identified study data can be archived:

- ☐ Yes – I agree
- ☐ No – I disagree