Trait Judgements and Gender Differences of Multimodal Emotional Cues

Taylor Barton

University of Adelaide

This thesis is submitted in partial fulfilment of the Honours degree of Bachelor of Psychological

Science (Honours)

Word Count: 9291

Table of Contents

List of Figures

Abstract

Judgements of certain traits have been linked with certain emotional expressions, such that trustworthiness is linked with happy expressions and dominance with anger. There are also differences in trait judgements made of male and female posers. Studies into trait judgement often use isolated facial expressions. Few studies have examined judgement in context with body movement and vocalisations. This study aims to examine the influence of multiple cues on trait judgements. Participants (N=158) judged the trustworthiness and dominance of multimodal expressions. Videoclips contained either congruent emotional expressions (all cues expressing the same emotion) or incongruent expressions (two cues expressing same emotion, one cue expressing a different emotion). Main analyses were conducted by examining emotions that were the majority emotion (two cues showed the same emotion) or the minority emotion (only one cue showed the emotion). As predicted, expressions that showed mostly happiness were rated high on trustworthiness and expressions showing mostly anger were rated as high on dominance. Additionally, there were significant gender differences. As predicted, the female actor was rated higher on trustworthiness, however the female actor was also rated higher on dominance, when literature typically finds males as higher on dominance. Results for minority cue judgement were not consistent with previous research into emotional perception of multimodal cues. This suggests that the cues that people use when judging traits are different then when evaluating emotion. Overall, results suggest that the addition of multiple cues influences the trait judgements people make.

Declaration

This thesis contains no material which has been accepted for the award of any other degree of diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time.

Contribution Statement

In writing this thesis, my supervisor, Dr Nicole Nelson, Hai Linh Tran and I collaborated to generate research questions of interest and design the appropriate methodology. Dr Nelson provided stimuli used in research design. Hai Linh Tran and I completed the ethics application and were responsible for all participant recruitment and testing. Some hypotheses were generated in collaboration with Hai Linh Tran, and some hypotheses were generated by me and were unique to my thesis. All hypotheses were preregistered on the Open Science Framework. I completed my own data coding and analyses, using Jamovi software. I also wrote up all aspects of the thesis.

Acknowledgements

I would like to thank my parents and brother for their continuous support throughout my entire education journey and especially this year, even through the busyness in their own lives.

My best friends Angie and Moira for making sure I take a night off every now and then and recharging my brain.

My 'lab sibling' Linda, our shared understand of our projects provided invaluable support and working together was a true pleasure.

My supervisor Dr Nicole Nelson, for the ongoing support and guidance, I feel I have learnt a lot this year and her kind support made the experience a lot more enjoyable.

Trait Judgements and Gender Differences of Multimodal Emotional Cues

People often make snap judgements about others based on how they look, and there is consistent agreement on how people judge others. Judgements of facial traits such as trustworthiness and dominance have been known to influence perceptions of criminality and sentencing decisions, and can even influence death penalty decisions (Flowe, 2012; Wilson & Rule, 2015). Facial appearance can also influence leadership selection and election outcomes. Facial judgement of leaders can affect how leaders are treated by others, the expectations of them in their role and even salary decisions (Antonakis & Eubanks, 2017; Lin et al., 2017).

Background

We often form impressions of others quickly and spontaneously. The amygdala automatically responds to facial cues and will assess trustworthiness without our conscious awareness (Engell et al., 2007; Freeman et al., 2014). There is argument that these judgements have an innate origin with evolutionary benefit (Over & Cook, 2018). However, the evolutionary argument for the existence of trait judgments is called into question as our spontaneous judgements are consistently inaccurate (Todorov et al., 2015; Zebrowitz & Montepare, 2008). For example, judgments of aggressiveness from faces has been demonstrated to be unrelated to with actual aggressive behaviour (Todorov et al., 2015). There is evidence that attribution of traits based on faces can be learnt from personal experiences (Todorov et al., 2015) and develop over time (Over & Cook, 2018). This is also evident from the influence of learnt stereotypes on trait judgements (Chua & Freeman, 2021). While our judgements may have developed from many different contextual and innate origins, certain trait judgements, such as threat, may have more connections to survival (Oosterhof & Todorov, 2008). It also appears that our judgements are able to be unlearnt; Chua and Freeman (2021), found that counter-stereotype training

changed participants judgments so that they were unaffected by facial trustworthiness compared to control subjects.

Dimensional Theories

Oosterhof and Todorov (2008), proposed that face evaluations existed on two trait dimensions. Using a principal components analysis, 13 different traits were narrowed down to 2 dimensions, which they described as representing valence and dominance. Valence is closely related to trustworthiness and describes the perceived intention of harm. The dominance dimension describes the perceived ability to harm. The model is relevant for perception of others where there is no additional context provided, for example when making snap judgments about strangers (Oosterhof & Todorov, 2008). This has been a prominent theory of trait judgement over the last decade, mostly tested in western cultures (Jones et al., 2021). Other dimensional theories have been proposed, including a three-dimensional theory suggested by Sutherland et al. (2013), which included: approachability, dominance and youthful-attractiveness. While approachability and dominance are in line with the dominance-valence dimension proposed by Oosterhof & Todorov (2008), the youthful-attractiveness dimension was suggested to have emerged due to the larger variety of ages and levels of perceived attractiveness used in the study (Sutherland et al., 2013). A recent four-dimensional theory was also explored (Lin et al., 2021), which also had youth as a factor, alongside warmth, competence, and femininity. This theory was able to single out femininity as its own dimension while the other theories did not. However, there were significant differences between gender of the posers or perceived femininity/masculinity when tested for the other dimensional theories (Oosterhof & Todorov, 2008; Sutherland et al., 2013).

Gender Differences in Trait Judgement

The gender of perceived faces consistently affects ratings of trait judgements. Additionally, the perception of femininity or masculinity for any gender can influence judgment in similar ways. Generally, more feminine or female faces are rated as more trustworthy and less dominant than more masculine or male faces (Oh et al., 2020; Oosterhof & Todorov, 2009; Sutherland et al., 2015). In addition, findings for competence show a bias towards males as being more competent (Oh et al., 2019). When evaluated for various trait judgements, females had stronger correlation between trait judgements then males did, leading to less variation of overall positive/negative valence and suggesting a more simplified evaluation of women (Oh et al., 2020). This simplification in perception of women could lead to the harsher evaluations found for women who are regarded as counter-stereotypical (Sutherland et al., 2015). Specifically, dominant looking female faces were rated as more negatively valanced than male faces and feminine female faces. It was suggested that this may be due to the rarity of seeing counter-stereotypical faces. However, the authors noted that this does not explain why there was little evidence for differences in valence between masculine and feminine male faces (Sutherland et al., 2015).

Oh et al. (2020) created models of expressions based on the same facial changes for both male and female posers. By comparing trait judgments between the gender models, the authors were able to determine that participants rated male and female posers based on similar visual information and changes, for example: the resemblance to emotional expressions was a key factor for the similarities in how participants rated male and female posers. While there are similarities in the visual information used to form impressions, there are differences in how these

impressions are evaluated between male and female posers. Thus, why there are clear influences from gender stereotypes on trait judgement (Oh et al., 2020).

Judgement of Emotional Expressions

Research has consistently found links from emotional expressions to trait ratings. This is applicable even when faces are neutral but have slight similarity to certain emotional expressions (Calvo et al., 2019; Oosterhof & Todorov, 2009; Said et al., 2009; Todorov, 2008). Trait ratings of emotional-looking neutral faces parallel ratings of emotional expressing faces (Montepare & Dobish, 2003). This is often described as an overgeneralization of emotion recognition and can affect interactions between cultures due to the overgeneralization of subtle emotional resemblance in faces of different groups (Zebrowitz et al., 2010).

Approach/Avoidance Behaviours

Happiness and anger signal approach and avoidance behaviours respectively more than any other emotional expression (Engell et al., 2010; Todorov, 2008). This behaviour relies on the evaluation of trustworthiness, which stems from the same neural mechanism as emotion perception (Engell et al., 2010). Evaluations of other emotions, such as fear, did not have as strong a response as happiness and anger (Engell et al., 2010). Hence why happy and angry expressions are the most studied in this area.

Happiness

Expressed happiness or slightly happy neutral facial expressions are typically judged as having positive traits, mainly trustworthiness (Calvo et al., 2019; Said et al., 2009). In fact, happy faces are judged higher on trustworthiness than any other facial expression. Additionally, the relationship between happy facial expressions and traits is bi-directional - trustworthy faces are rated as happier than non-trustworthy faces (Oosterhof & Todorov, 2009).

Anger

Angry expressions are typically linked with judgements of dominance and threat, and lower ratings of trustworthiness and affiliation (Montepare & Dobish, 2003; Oosterhof & Todorov, 2009; Said et al., 2009). With anger typically associated with aggression, this type of expression sends signals in the brain in response to a potential threat (Hess et al., 2000; Said et al., 2009). Similarly to happy expressions, there is a bi-directional relationship between anger and trustworthiness, when an untrustworthy face is expressing anger they are seen as angrier than a trustworthy face expressing anger (Oosterhof & Todorov, 2009).

Sadness and Fear

Sad or fearful looking neutral faces did not tend to be overgeneralized to trait judgements (Montepare & Dobish, 2003). While emotions such as fear and disgust are not usually threatening, they are still considered to be unexpected in social situations. This leads to low ratings of valence, and avoidance behaviours (Said et al., 2009). Interestingly, unlike anger, fear is considered low on dominance, while both are an avoidance expression (Montepare & Dobish, 2003). Although, as fear is low on trustworthiness (Said et al., 2009), this shows that dominance is not necessarily an avoidance trait. This also adds to the evidence that approach/avoidance behaviours are more related to ratings of trustworthiness.

Trait judgements related to intelligence, had a negative correlation to fear. It was suggested that this correlation was due to the overgeneralization in perceiving mental illness. Interestingly, sadness does not correlate strongly with intelligence, dominance or trustworthiness (Said et al., 2009). Generally, sadness, along with fear, as a signal of distress often result in avoidance behaviour (Montepare & Dobish, 2003).

Gender and Emotion Judgement

There are gender differences in how emotional expressions are perceived, and this perception is affected by gender stereotypes. Sad and happy expressions were perceived faster in female actors than in male actors, while anger was more accurately and sooner perceived for male actors (Becker et al., 2007; Parmley & Cunningham, 2014). Participants also had a greater startle response when viewing a happy female, this was also seen when viewing an angry male, indicating that people have differing immediate reactions to emotional expressions when expressed by a male or female (Åsli & Øvervoll, 2020). It has been suggested that when viewing faces that were more ambiguous in emotion expression, participants relied on gender stereotypes to guide their emotion perception (Parmley & Cunningham, 2014). On the other hand, these differences could be as a result of men's facial structure more generally resembling angry-looking faces. However, the effects of learnt stereotypes on judgement cannot be denied when there are found gender variations for faces that do not resemble certain emotions (Zebrowitz et al., 2010).

The label 'resting bitch face' is often used to describe women whose neutral faces have a structure that makes them look angry. Previous research has described this as perceived resting negative emotion, as some people have resting facial expressions that resemble negative emotions, such as anger. Both men and women perceived as having a resting negative emotion are rated higher on threat. However, women with resting negative emotion are also rated as less attractive, while men are not (Hester, 2019). Women expressing approach related emotions were rated higher on dominance then their neutral expressions. While men are perceived as more dominant overall, their emotional expressions were not significantly higher on dominance than their neutral expressions (Hareli et al., 2009).

Body and Voice Research

The majority of research into the link between emotion and trait judgement is through the study of faces. Fewer studies have examined the link between emotional body expression, vocalisation, and trait judgement. In everyday interactions we do not see only facial expressions, and research has found that people make trait judgements based on body and voice cues (Aung & Puts, 2020; Koppensteiner, 2013).

Body Expression

Just as we would find in face research, the trait judgements we make for bodies appear to happen quickly, inaccurately and with high consensus (Koppensteiner, 2013; Thoresen et al., 2012), with effects on electoral decisions, especially in election rallies and speeches, where subjects are viewed from a greater distance (Koppensteiner, 2013). Additionally, when viewing bodies expressing certain emotions, trait judgements were similar to those made for faces, for example: angry bodies are rated higher on dominance (Van Der Zant et al., 2021). Research using stick-figure motion, based on politician speech movements, revealed that highly expansive movement and high amounts of movement were linked to high judgments of dominance and low judgments of trustworthiness. Additionally, dominance was perceived more often and negatively predicted trustworthiness, especially for male speakers (Koppensteiner et al., 2016).

Research into gait movement found that participants are consistent in their ratings of walkers on personality traits. Judgments could be broken up into two components for high and low amplitude of gait movement. High amplitude of gait movement correlated with traits such as trustworthiness, adventurousness, warmth, and extraversion (Koppensteiner, 2013; Thoresen et al., 2012). This is in contrast to trustworthiness judgements found in the study using stick-figure motion based on speech movements, where more movement was less trustworthy, suggesting

that the type of movement or the context may influence how people judge bodies. Furthermore, gait judgements were mediated by emotion, masculinity, and attractiveness, where attractiveness predicted ratings of approachability and masculinity predicted ratings of emotional stability (Thoresen et al., 2012). Indeed, it was found that people can accurately identify emotion based on gait, particularly happiness, sadness and anger, although only fear was equally identified in female walkers (Halovic & Kroos, 2018). Thus, people are able to identify traits and emotions based on body movement and there is a perceived difference in gender for these judgements.

Vocalisations

Vocalisation research often analyses vocal pitch as a main factor in perception and judgement. Impressions of lower pitch voices are often linked with ratings of masculinity, dominance, and larger physicality. Vocal pitch is often modified in certain social settings, for example, men will raise their pitch when they view another as more dominant, and lower their pitch when they view themselves as more dominant (Aung & Puts, 2020). Lower pitch was also related to higher trustworthiness (Elkins & Derrick, 2013), but the context of the vocalization was important. Specifically, lower pitch was judged as more trustworthy for men in economic trust settings but the opposite was found in mating related contexts (Schild et al., 2019). Research often looks at the effect of pitch on the perception of men, however vocal pitch of women is also important, as it can influence the perception of success, although the authors noted that more research is needed in that area. Additionally, these results are often strong in lab settings but less so in contextual settings (Aung & Puts, 2020).

Cue Interaction

It is clear in research that expression in context - whether that be environment or integration of cues such as facial or postural expressions - influences perception and judgement

of others (Aviezer et al., 2012; Elkins & Derrick, 2013; Stoesz & Jakobson, 2014). When multiple cues are available to a viewer, it appears that people take all cues into account when making a judgment (Koppensteiner, 2013). Having said that, there may be a preference for certain cues. When voice cues were paired with a smiling face they were higher on trustworthiness than when paired alone (Elkins & Derrick, 2013) and there is evidence that adults will make judgements based more on visual information when paired with audio cues (Zupan et al., 2015). Although there is some ambiguity when introducing cues with different expressions at once, described as incongruent emotion cues. Comparing incongruent cues can determine which cue or emotion are preferred in trait judgement. In one study, dominance and anger were attributed more to body expressions than the face, and happiness and trust were attributed to facial expressions rather than the body (Van Der Zant et al., 2021). In a second study, Casey et al. (2021) researched emotion perception for three integrated cues, revealing patterns in perception that are expected to mirror trait judgements for multimodal cues. For instance, their results revealed that participants had preference for facial cues over body and voice cues and that emotions were most easily perceived when expressed by the majority of cues.

The Current Study

This study aims to build on research involving multimodal cues and draw attention to the literature gap in trait judgement of multimodal cues. Research by Casey et al. (2021) aimed to study emotion perception of multimodal cues. Stimuli comprised of congruent (3 matching expressions), incongruent (2 vs. 1 emotion expressions) and isolated dynamic emotional cues. Results included the analysis of ratings when a certain emotion is in the majority (majority emotion) and when a cue is expressing the 1 different emotion (minority cue). Using the same stimuli, the current study tested the influence of combining facial cues, body expressions and

vocalisations on judgments of traits. Through studying incongruent stimuli we were able to test which emotions and which cue types have a bigger impact on judgements of trustworthiness and dominance and whether the influence varies with gender of the poser.

Hypotheses

Hypotheses 1 and 2 draw on literature connecting trustworthiness and dominance with emotional expressions. These analyses focus on incongruent stimuli where the target emotion is expressed in 2 cues (majority emotion).

1. Emotion. Happy cues will be higher on trustworthiness than any other emotion, and low on dominance. Anger cues will be higher on dominance than any other emotion and low on trustworthiness. Scared cues will be lower on dominance than any other emotion and low on trustworthiness.

2. Gender. Overall, judgements of dominance will be higher for the male poser and trustworthiness will be higher for the female poser.

2a. As female faces are judged more harshly when they look dominant, I predicted that the female poser will be rated as less trustworthy when anger is the majority emotion.

Hypotheses 3 and 4 draw on literature involving the preference of certain cue types and emotion perception and judgement, by using the incongruent stimuli where a certain emotion is expressed by 1 cue type (minority cue)

3. Cue type. There will be a preference for facial expressions when expressing the minority emotion, thus trait judgements will be stronger when the face is the minority emotion, in the expected direction. For example, when the face is happy and other cues are sad, the trait judgement will be higher on trustworthiness, compared to when happy is a different minority cue.

4. Emotion. As certain emotions are more easily recognisable when cues are incongruent, I predicted this would lead to differences in judgement when an emotion is in the minority compared to congruent emotion expressions.

For this study, I pre-registered five hypotheses, which can be found on the Open Science Framework (https://osf.io/xtu7v/?view_only=90e82de00028427fbb80bbb0215958ba)[1]

Methods

Participants

There were 204 participants recruited for this study. Of these participants, 46 were excluded for failing to complete the task or not following instructions, leaving a total of 158 included in the study (49 males, 109 females). Participants were first year psychology students enrolled at the University of Adelaide who participated in exchange for course credit. The participants were aged between 17 and 55 years (mean = 20.96) and had lived in Australia between 0 and 55 years (mean = 16.10). An a priori power analysis was conducted using G*Power (Faul et al., 2007) for a 2(gender) x 4(emotion) x 2(trait) repeated measures ANOVA, to detect a small effect size (f = 0.10) at 85% power and alpha of 0.05. The required sample size to power this study was determined to be 102 participants. This study was approved for ethics through the University of Adelaide.

---

[1] This study was part of a larger project, and not all pre-registered hypotheses are covered here. My 4 hypotheses correspond to Hypothesis 1, 2, 3 and 5 on the OSF page.

Materials

Participants were presented with a total of 104 videoclips, which ranged from 1 to 4 seconds in length. Videoclips featured a male or female poser expressing: congruent cues where all cues were expressing the same emotion, incongruent cues with two cues expressing the same emotion and one cue expressing a different emotion (Figure 1), or isolated cues (just one cue presented). Each cue was presented through face, body, and voice expressions, conveying one of the four emotions: happiness, sadness, anger or fear. All stimuli used were sourced from (Casey et al., 2021).

Emotional Expressions

Videoclips showed the poser starting with neutral posture and facial expression and moving to the intended position and expression for the given emotion (Refer to appendix 1 for all images of expressions). The male and female poser each posed all emotional expressions and cue combinations.

Happy. Expressions of happiness included smiling for the facial cues, upright body posture with swinging arms, and laughter vocalisations.

Sad. Expressions of sadness included a lower lip pout and raising the inner corner of the eyebrows for the facial cues, slouching body posture, and crying vocalisations.

Angry. Expressions of anger included pursed lips and furrowed brows for the facial cues, upright body posture with fists slamming down to the sides, and growling vocalisations.

Fearful. Expressions of fear included raised eyebrows, widened eyes and fear grimace for the facial cues, leaning back posture with hands defensively raised in front of the torso, and shrieks or screaming vocalisations.

Figure 1

Incongruent Stimuli

Note. Male incongruent stimuli example: happy face, scared body and happy vocalisation.

Female incongruent stimuli example: happy face, sad body and happy vocalisation.

Multimodal Stimuli

There were 80 multimodal clips containing 8 congruent and 72 incongruent emotion cue combinations. Half of the clips depicted the male poser and half depicted the female poser. For the congruent stimuli, the face, body, and voice all showed the same emotion. Emotion was then manipulated amongst the cue types for the incongruent stimuli, with one emotion always being in the majority (e.g. happy face, happy body, sad voice).

Isolated Stimuli

There were also 24 isolated clips, which depicted only one emotional cue at a time. There were 8 clips each for face, body, and voice expressions. Half of the expressions were posed by the male and the female posers. The videoclips included either a face or body expression, with

the other edited out and no sound. Voice expressions were presented with the audio only. For the face expressions, only the head and neck were visible. For the body expressions, the body from the shoulders down were visible (refer to Appendix 1).

Stimuli were validated by the source study authors and reported as the accuracy of the emotion recognition. Out the expressions chosen for use, accuracy was above 74% for facial expressions and above 80% for body expressions (Nelson & Mondloch, 2017). Vocalisations were reported as high in accuracy (>90%, Hawk et al., 2009).

Procedure

Participants enrolled via SONA and the study was presented as a Qualtrics survey. Participants first provided demographic information (age, gender, and how many years they had lived in Australia). Next, participants completed a practice trial, so they could get used to the question format, where they viewed a videoclip and rated it on how trustworthy, dominant and competent they thought it was. These traits were presented in a matrix with a 5 point Likert scale (1: Not at all to 5: Completely, see the Appendix 2 for further details). They then moved on to the main survey where they viewed and were asked to rate the multimodal clips in randomised order, followed by the isolated clips, which were also randomised. This data was collected for use in a larger study, but for the purposes of this study, the competence data was not analysed.

Analysis Plan

Analysis of trustworthiness and dominance was conducted via Jamovi statistical software. Descriptive statistics were used for participant demographic information. Five ANOVAs were used to provide outcomes on trait judgement for majority emotion (the emotion that is expressed by 2 cues), minority cue (the cue that expresses the 1 different emotion), gender

differences, and congruent stimuli. Significant effects were followed up with Bonferroni-corrected post-hoc tests and significant results were reported.

To test hypotheses 1 and 2, a 2 (gender of poser) x 2 (trait) x 4 (emotion) repeated measures ANOVA was used to examine the majority emotion variables. Hypothesis 1 was tested using the emotion x trait interaction, to analyse majority emotion association with trustworthiness and dominance. Hypothesis 2 was tested using the gender x trait interaction, to analyse overall gender differences for trustworthiness and dominance. Hypothesis 2a was tested using the gender x emotion x trait interaction, to analyse if anger is less trustworthy for the female poser.

To test hypothesis 3, a 3 (cue type) x 4 (emotion) x 2 (trait) x 2(gender of poser) repeated measures ANOVA was used to examine the minority emotion variables. This analysis was broken up into 2 ANOVAs for each trait. One 3 (cue type) x 4 (emotion) x 2(gender of poser) ANOVA for trustworthiness and one for dominance. This provided an analysis of which minority cues were rated higher for each trait and poser gender differences. These results were compared to majority analysis results to check for directionality.

To test hypothesis 4 a 4 (emotion) x 2 (trait) x 2 (multimodal type: congruent, minority) repeated measures ANOVA was used to analyse which minority emotions were closer to the congruent stimuli rating of the same emotion.

As an exploratory analysis, Bonferroni-corrected t-tests were conducted to compare averaged isolated stimuli and congruent stimuli for each emotion.

The differences in trait judgements between the male and female participants were also analysed via ANOVA in an exploratory analysis.

Results

Majority Emotion

A 4 (emotion) x 2 (trait) x 2 (gender) repeated measures ANOVA was conducted to determine how participants rated stimuli when a particular emotion is in the majority and whether these ratings varied with poser gender. There was a main effect of majority emotion $F(3, 471) = 74.55$, $p < .001$, $\eta p2 = .322$, and significant interaction between majority emotion and trait, $F(3, 471) = 134$, $p < .001$, $\eta p2 = .46$. Follow-up tests for the emotion x trait interaction compared emotion nested within trait, resulting in 12 t-tests. These tests were Bonferroni-corrected with the alpha set at $0.05/12=.004$.

Dominance

Majority angry cues (M = 2.65, SD = 0.81) were rated significantly higher than any other emotion on dominance (ps < .001, Figure 2). This is consistent with hypothesis 1 for majority angry cues. Majority happy cues (M = 2.04, SD = 0.66) were rated significantly higher than both scared and sad majority ratings (ps < .001). There was no significant difference between scared (M = 1.83, SD = 0.55) and sad (M = 1.76, SD = 0.57) majority ratings. There was partial support for the scared cue prediction in hypothesis 1 as ratings were lower than happiness and anger but not significantly lower than sadness.

Trustworthiness

Majority happy cues (M = 2.30, SD = 0.71) were rated significantly higher than any other emotion on trustworthiness (ps < .001, Figure 2). This is consistent with predictions for majority happy ratings in hypothesis 1. Majority angry cues (M = 2.00, SD = 0.61) were rated lower on trustworthiness than any other emotion (ps < .001). This is also consistent with hypothesis 1.

Again, there was no significant difference between scared (M = 2.12, SD = 0.58) and sad (M = 2.14, SD = 0.61) majority ratings.

Hypothesis 1 predicted that happiness would be higher on trustworthiness than all other emotions, anger would be higher on dominance than all other emotions and scared cues would be lower on dominance than all other emotions. These predictions were supported for anger and happiness ratings. Scared ratings were partially supported in that mean ratings were considered low on the scale especially for dominance, however ratings of scared expressions were not differentiated from ratings of sad expressions.

Gender of Poser

In addition, there was a main effect for gender of the poser $F(1, 157) = 30.95$, $p < .001$, $\eta p2 = .16$ (Figure 2). Bonferroni-corrected t-tests ($\alpha = 0.05/2 = .025$) were conducted to compare overall ratings for male and female majority emotion per trait. These tests revealed that the female poser was rated significantly higher on dominance $t(157) = 4.27$, $p < .001$ and trustworthiness $t(157) = 4.73$, $p < .001$ than the male poser. This partially supports hypothesis 2 as the female ratings were higher on trustworthiness, however the male was not higher on dominance as was predicted. There was also a significant interaction between emotion, trait and gender of poser, $F(3, 471) = 10.05$, $p < .001$, $\eta p2 = .06$. Bonferroni-corrected t-tests ($\alpha = 0.05/8 = .006$) were conducted to compare the ratings of the male and female poser for each emotion and trait.

Happy. The female poser was rated significantly higher on dominance and trustworthiness for majority happy cues ($p < .001$).

Sad. The female poser was rated significantly higher on trustworthiness for sad majority cues $t(157) = 4.31$, $p < .001$. There was no significant difference on dominance ratings.

Angry. There was no significant difference between male and female ratings for either trustworthiness or dominance.

Scared. The female poser was rated significantly higher on dominance for scared majority cues $t(157) = 3.51$, $p < .001$. There was no significant difference for trustworthiness ratings.

Hypothesis 2a predicted that there would be differences in trustworthiness ratings for majority angry cues, which was not supported as there were no significant gender differences for angry majority cues. However, happy majority cues had differences for both traits, and the female was rated as more trustworthy when sad but more dominant when scared, which was not predicted.

Figure 2

Majority Emotion x Trait x Gender

Note. Mean dominance and trustworthiness ratings by emotion and gender of the poser. Error bars represent standard errors.

Minority Emotion

An initial 4 (minority emotion) x 3 (minority cue type) x 2(gender) x 2(trait) repeated-measures ANOVA found a four-way interaction, $F(6, 942) = 5.42$, $p < .001$, $\eta p2 = .033$. To more closely examine this interaction, this analysis was divided into two separate 4 (minority emotion) x 3 (minority cue type) x 2(gender) ANOVAs, one for trustworthiness and one for dominance.

Dominance

A 4 (minority emotion) x 3 (minority cue type) x 2(gender) repeated measures ANOVA was conducted to test how participants rated dominance when emotions were expressed by the minority cue. There was a main effect of minority cue type $F(2, 314) = 11.07$, $p < .001$, $\eta p2 = .066$ and significant interaction between minority cue type and minority emotion $F(6, 942) = 10.98$, $p < .001$, $\eta p2 = .065$ (Figure 3). To compare cue type for each emotion, Bonferroni-corrected post-hoc t-tests ($\alpha = 0.05/12 = .004$) were conducted.

Happy. When the voice was the happy minority cue, participants rated it significantly less dominant than when the face was the minority cue, $t(157) = 6.44$, $p < .001$, or when the body was the minority cue, $t(157) = 3.76$, $p < .001$.

Sad. When the voice was the sad minority cue, participants rated it significantly less dominant than when the face was the minority cue, $t(157) = 3.77$, $p < .001$, or when the body was the minority cue, $t(157) = 3.75$, $p < .001$.

Angry. When the face was the angry minority cue, participants rated it significantly less dominant than when the body was the minority cue, $t(157) = -4.90$, $p < .001$, or when the voice was the minority cue, $t(157) = -3.86$, $p < .001$.

Scared. There were no differences between the cues for dominance ratings.

Figure 3

Dominance Ratings of Minority Emotion x Cue Type

Note. Mean dominance ratings for each minority emotion by cue type. Error bars represent standard errors.

There was also a main effect of gender, $F(1, 157) = 18.29$, $p < .001$, $\eta p2 = .10$ and a significant interaction between cue type, minority emotion and gender, $F(6, 942) = 2.71$, $p = .013$, $\eta p2 = .017$ (Figure 4).

T-tests using a Bonferroni-corrected alpha ($\alpha = 0.05/12 = .004$) were conducted to compare male and female ratings for cue type and emotion.

Face. There were no gender differences for dominance ratings when the face was the minority cue.

Body. There were no gender differences for dominance ratings when the body was the minority cue.

Voice. The female poser was significantly higher on dominance than the male poser when the voice was the scared minority cue, $t(157) = 5.07$, $p < .001$.

Figure 4

Dominance Ratings of Minority Emotion x Cue Type x Gender
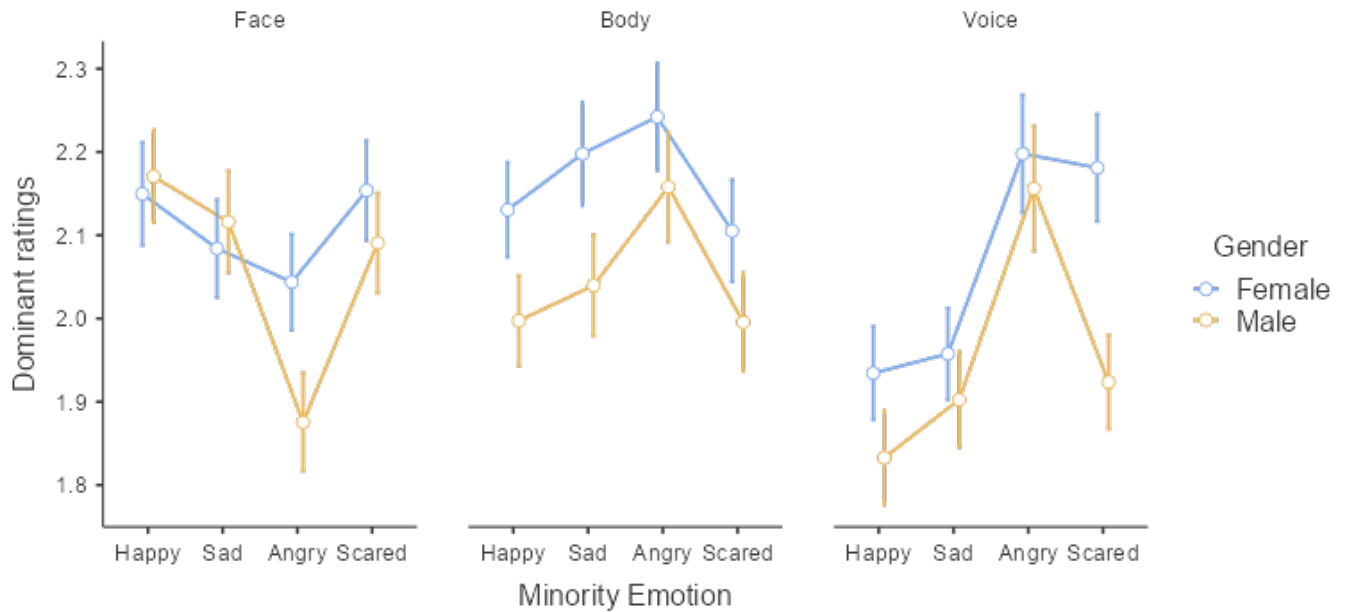
Note. Mean dominance ratings for each minority emotion by cue type and gender. Error bars represent standard errors.

Trustworthiness

A 4 (emotion) x 3 (cue type) x 2(gender) repeated measures ANOVA was conducted to test how participants rated trustworthiness when emotions are expressed by the minority cue. There was a main effect of cue type, $F(2, 314) = 34.80$, $p < .001$, $\eta_p^2 = .18$, a significant interaction between cue type and minority emotion, $F(6, 942) = 13.19$, $p < .001$, $\eta_p^2 = .07$ and a significant interaction between cue type, minority emotion and gender, $F(6, 942) = 3.38$, $p = .003$, $\eta_p^2 = .02$ (Figure 5). To compare cue type for each emotion, Bonferroni-corrected post-hoc t-tests ($\alpha = 0.05/12 = .004$) were conducted.

Happy. When the face was the happy minority cue, participants rated it significantly more trustworthy than when the voice was the minority cue, $t(157) = 5.03$, $p < .001$.

Sad. When the body was the sad minority cue, participants rated it significantly more trustworthy than when the face was the minority cue, $t(157) = -4.23$, $p < .001$.

Angry. When the body was the angry minority cue, participants rated it significantly more trustworthy than when the face was the minority cue, $t(157) = -4.41$, $p < .001$, or when the voice was the minority cue, $t(157) = 5.26$, $p < .001$.

Scared. When the body was the scared minority cue, participants rated it significantly more trustworthy than when the face was the minority cue, $t(157) = -8.62$, $p < .001$, or when the voice was the minority cue, $t(157) = 9.99$, $p < .001$. When the face was the minority cue, participants rated it significantly more trustworthy than when the voice was the minority cue, $t(157) = 3.26$, $p = .001$.

Figure 5

Trustworthiness Ratings of Minority Emotion x Cue Type
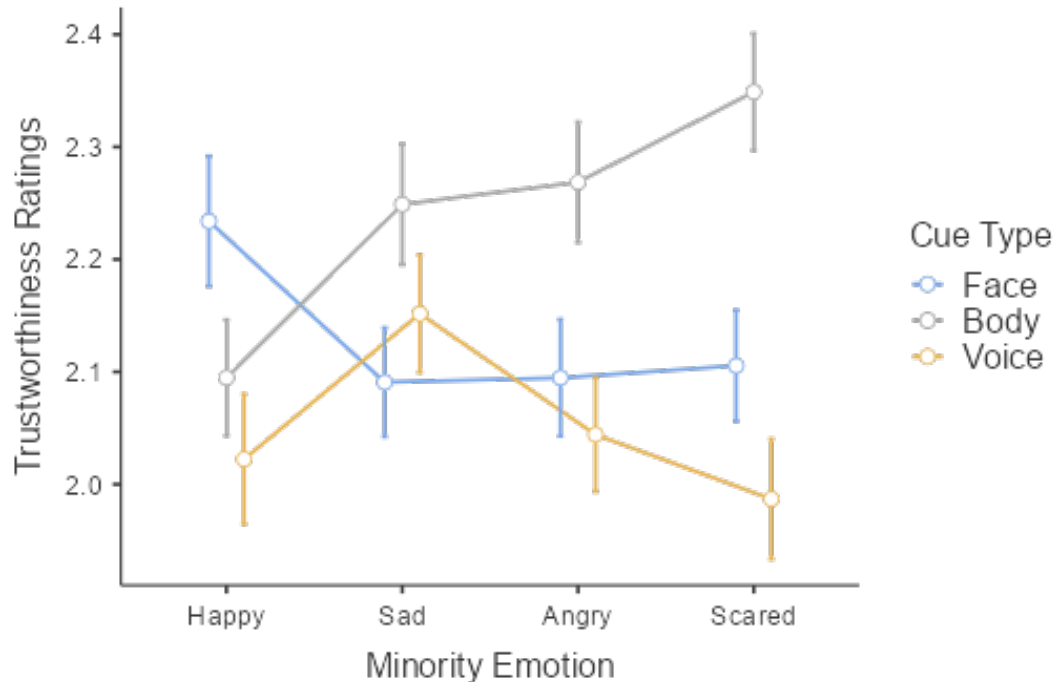
Note. Mean trustworthiness rating for each minority emotion by cue type. Error bars represent standard errors.

There was also a main effect of gender $F(1, 157) = 22.39$, $p < .001$, $\eta_p^2 = .12$ and a significant interaction between cue type and gender $F(1, 157) = 14.16$, $p < .001$, $\eta_p^2 = .08$ (Figure 6). Bonferroni-corrected t-tests ($\alpha = 0.05/12 = .004$) were conducted to compare male and female ratings for cue type and emotion.

Face. The male poser was significantly higher on trustworthiness ratings than the female poser when the face was the scared minority cue, $t(157) = -3.16$, $p = .002$.

Body. The female poser was significantly higher on trustworthiness ratings than the male poser when the body was the happy minority cue, $t(157) = 3.77$, $p < .001$, or the angry minority cue, $t(157) = 3.97$, $p < .001$ or the scared minority cue, $t(157) = 3.85$, $p < .001$.
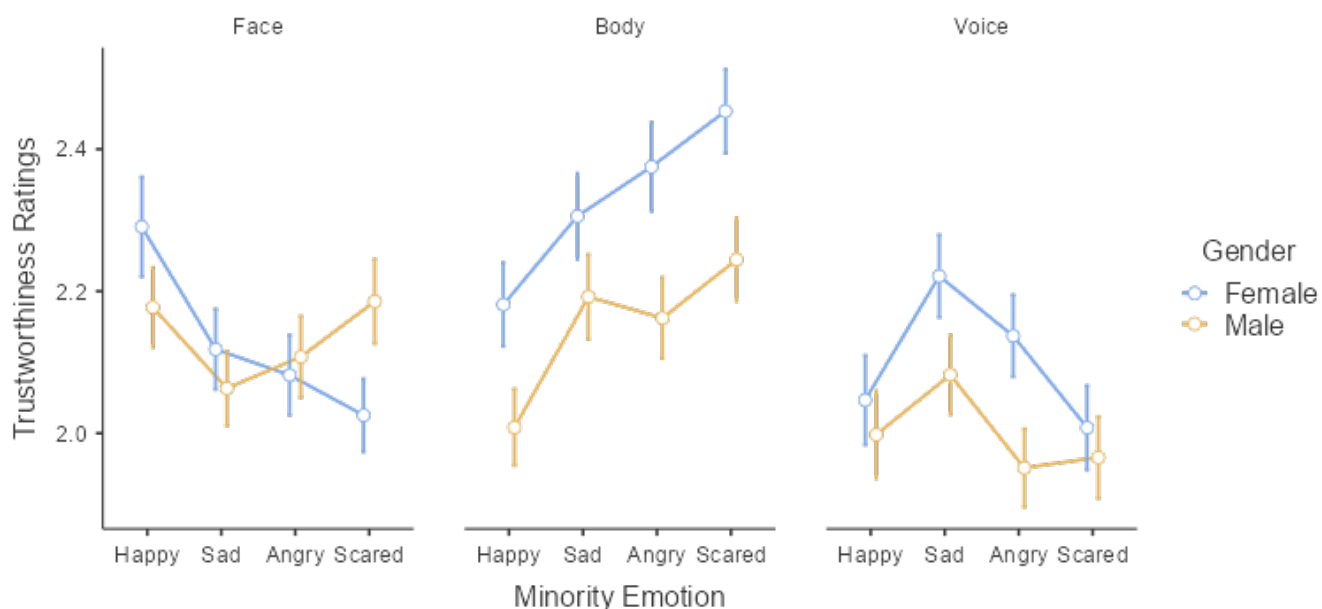
Voice. The female poser was significantly higher on trustworthiness ratings than the male poser when the voice was the sad minority cue, $t(157) = 3.04$, $p = .003$ or the angry minority cue, $t(157) = 3.76$, $p < .001$.

Figure 6

Trustworthiness Ratings of Minority Emotion x Cue Type x Gender

Note. Mean trustworthiness ratings for each minority emotion by cue type and gender. Error bars represent standard errors.

Facial expressions as the minority emotion do not appear to have the highest ratings on trustworthiness and dominance, contrary to what was predicted in hypothesis 3. It appears that the body expressing the minority emotion had higher trait ratings. Especially for trustworthiness ratings when angry and scared are the minority emotions, as the minority body expressions were significantly more trustworthy than both the face and voice expressions.

Scared expressions being rated as more trustworthy is unexpected as scared majority analysis revealed low ratings of trustworthiness. It was also unexpected that minority angry

expressions in the body were rated as trustworthy. Only for dominance ratings when the voice is expressing anger was voice rated significantly higher than the face, although, there was no significant difference to the body. Neither face nor voice minority ratings were significantly higher than the body ratings, even when a higher trait rating does not match the emotion as found in majority analysis.

The male poser was not significantly higher than the female poser on ratings of trustworthiness and dominance for any majority emotion. However, when the male face was the minority scared emotion, trustworthiness ratings were higher, so only scared expressions influenced the minority face ratings. Additionally, the angry female body and voice minority cues were rated higher on trustworthiness, which is also inconsistent with majority results.

When the scared female voice was the minority, ratings were also higher on dominance, which is consistent with the majority rating that scared cues were higher on dominance for the female poser. Sad minority cues were also consistent with majority analysis as the female poser was rated higher on trustworthiness.

Congruent and Minority Emotion

To test Hypothesis 4, which predicted that minority ratings would be significantly different from congruent ratings, a 4 (emotion) x 2 (trait) x 2 (Multimodal type: congruent, minority cue) repeated measures ANOVA was conducted. There was a main effect of multimodal type (congruent or minority), $F(1, 157) = 47.9$, $p < .001$, $\eta_p^2 = .234$, and a significant interaction between the multimodal type, emotion and trait, $F(3, 471) = 117.7$, $p < .001$, $\eta_p^2 = .429$ (Figure 7).

Bonferroni-corrected t-tests ($\alpha = 0.05/8 = .006$) were conducted to compare congruent and minority ratings. There were no significant differences between dominant ratings when happy

was the congruent or minority emotion and there were no significant differences between trustworthy ratings when anger was the congruent or minority emotion. Thus, happy dominant ratings and angry trustworthy ratings were inconsistent with predictions.

Other emotions had significant differences between congruent and minority ratings, however it was unclear which emotions or traits had the biggest differences. By testing the effect size of the minority and congruent comparisons, the differences can be ranked for each emotion and trait. For the comparisons that had significant differences, two of the effect sizes could be classified as large (Cohen, 1988): angry dominant ratings $t(157) = 13.17$, $p < .001$, d = 1.06; happy trustworthy $t(157) = 11.71$, $p < .001$, d = 0.95. In addition, 2 of the effect sizes could be classified as small to medium: sad dominant $t(157) = -8.56$, $p < .001$, d = 0.65; scared dominant $t(157) = -7.66$, $p < .001$, d = 0.59. Finally, two of the effect sizes could be classified as small: sad trustworthy $t(157) = 4.53$, $p < .001$, d = 0.33; scared trustworthy $t(157) = 4.22$, $p < .001$, d = 0.31.

This would suggest that, for happy and angry expressions for trustworthiness and dominance respectively, the combination of congruent cues generated a larger effect on ratings than was observed for the other emotions. It would be expected that angry cues would be rated significantly lower on trustworthiness, as seen in majority analysis, however there were no significant differences between congruent and minority anger cues for trustworthiness ratings. Additionally, there were also bigger differences for dominant ratings, except for happy cues, where trustworthiness had a bigger difference. This would suggest that while congruent anger and happiness expressions had higher ratings, dominance ratings are affected more by congruency than trustworthiness.

Figure 7

Multimodal Type x Trait x Emotion
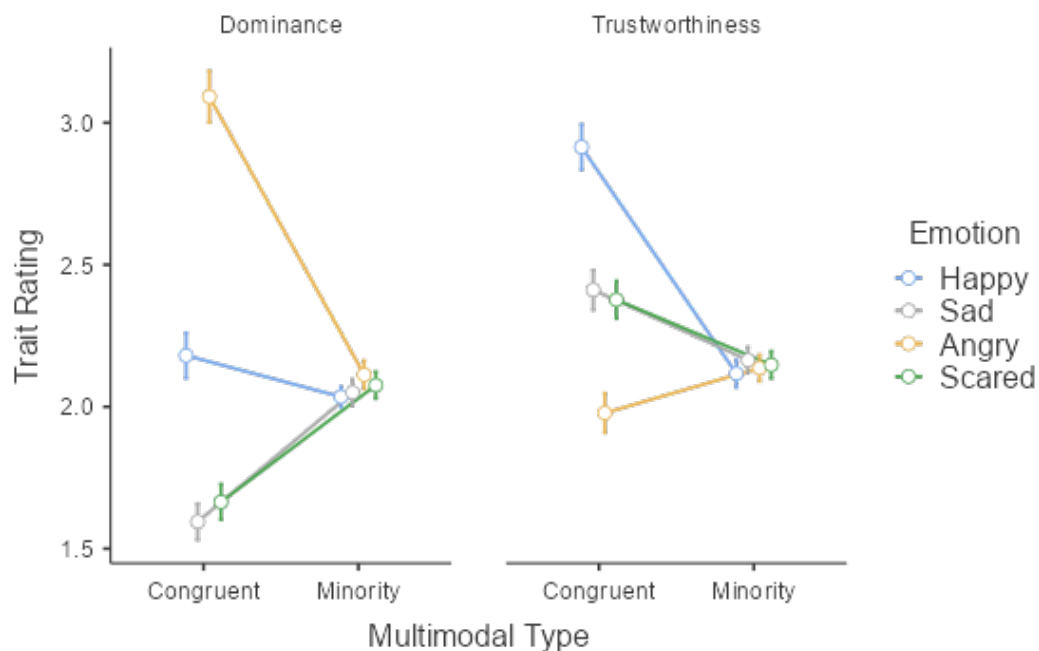
Note. Mean dominance and trustworthiness ratings for each multimodal type by emotion. Error bars represent standard errors.

Exploratory Analysis

Isolated and Congruent Stimuli

To test if there were differences between when cues are isolated to when they are combined in the same stimuli, ratings were compared for isolated and congruent stimuli. Bonferroni-corrected post-hoc t-tests ($\alpha = 0.05/8 = .006$) were conducted between congruent stimuli and the isolated stimuli face, body, and voice cues for each emotion. There was one significant difference in trustworthiness ratings, with congruent scared stimuli rating significantly higher than isolated stimuli $t(936) = -2.9384$, $p = .003$. This would suggest that the scared cues in isolation are considered less trustworthy.

Gender of the Participant

Whether the gender of the participant influenced stimuli ratings was analysed using repeated measures ANOVAs that parallel those reported in earlier analyses examining the gender of the poser. These included a 4 (emotion) x 2 (trait) x 2 (poser gender) ANOVA, a 4 (emotion) x 2 (trait) x 3 (cue type) x 2(poser gender) ANOVA and a 4 (emotion) x 2 (trait) x 2 (congruent, minority) ANOVA with participant gender as a between subjects factor. There were no significant main effects or interactions related to participant gender for any ANOVA ($ps > .05$).

Discussion

The purpose of this study was to gain an understanding of how multiple cues effect trait judgements and how results differ between male and female posers. The results of this study demonstrate that emotions expressed by the majority of cues tend to be rated as predicted and there are differences in judgements based on the cue expressing the minority emotion (emotion expressed by one cue). Happy and angry expressions were rated high on trustworthiness and dominance respectively, with no significant differences in ratings between sad and scared expressions.

A key finding when analysing the minority cues was that when the body was expressing anger and fear as the minority, ratings were significantly higher on trustworthiness than the face or voice minority cues. Furthermore, face and voice minority cues were not significantly higher than body minority cues for any emotion, this went against predictions that the face would be the prominent influence on trait judgement.

Another key finding related to gender analysis was that the female poser was rated higher on trustworthiness and dominance, this was inconsistent with predictions. Additionally,

predictions involving the female angry expression being rated as less trustworthy were also not found in the analysis.

Finally, congruent ratings of trustworthiness for happy expressions and dominant ratings for angry expressions had the biggest difference from when they were expressed in the minority.

Preference for Majority Emotion

Results obtained for majority emotion analysis provide support for hypothesis 1 for happy and anger expressions. Mostly happy expressions were rated higher on trustworthiness than any other emotion and anger was higher on dominance than any other emotion, this is consistent with previous research for facial expressions (Calvo et al., 2019; Montepare & Dobish, 2003; Said et al., 2009). There is partial support for scared expression predictions, in that ratings were generally low on both traits. Although, mostly scared and sad cues were not rated differently. They were both lower on dominance then other emotions and in the middle of happiness and anger for trustworthiness ratings. Previous research has not found strong support for sad and scared facial expressions being overgeneralised to trait judgements (Montepare & Dobish, 2003). Thus, these results are consistent with previous research.

Previous research into combining multiple cues were primarily focused on emotion perception. Casey et al. (2021) found that participants accurately perceived and had preference for the majority emotion when three cues were presented. As trait judgments have links to overgeneralisation of emotion perception, it would be expected to mirror emotion perception patterns. Trait judgement had the same pattern of perception for majority emotion as there was a preference for the majority emotion when there were incongruent expressions. Thus, the integration of three cues did not disrupt the expected trait judgements for majority emotion. This is consistent with research suggesting that when making a trait judgments, participants will take

all available cues into account (Koppensteiner, 2013). The consistency and clear pattern of trait judgment of the majority emotion, despite which cues are expressing such emotions, provides evidence for integration of all emotional cues.

Gender of the Poser and Majority Emotion

Differences between the male and female posers were examined and there were found differences between ratings for emotions expressed by the majority of cues. Hypothesis 2 was partially supported as the female poser was higher on trustworthiness than the male poser, however the male poser was not higher on dominance. This is inconsistent with previous research that consistently finds males or masculine faces as higher on dominance (Oh et al., 2020; Oosterhof & Todorov, 2009; Sutherland et al., 2015). Research into gender stereotypes shows that counter-stereotypical faces are judged more harshly for female faces (Sutherland et al., 2015). It was suggested that this was due to the less variation of overall positive/negative valence, leading to a stronger response when the poser is counter-stereotypical (Oh et al., 2020). This study's integration of multimodal cues adds an extra element that may have affected results. Specifically, incongruent expressions could have been viewed as unexpected if the more simplified view of women extends to expressions that are mis-matched. This could be a reason for why the female poser was rated more strongly on dominance. However, dominance is not necessarily regarded as a negative trait (Oosterhof & Todorov, 2008) and harsher evaluations of women tend to relate more to valence or trustworthiness (Sutherland et al., 2015). As trustworthiness was expected to be higher for women in this study, results found for dominance may be unrelated to counter-stereotypical effects.

Hypothesis 2a predicted that the female poser would be judged as less trustworthy when expressing anger. Predictions were made based on counter-stereotype research examining

situations when a female poser is dominant or masculine looking (Sutherland et al., 2015). As seen in past research and in majority emotion analysis in this study, dominance is often rated higher on anger. Therefore, it was predicted that anger would also lead to negative ratings, in the same way a dominant female face would. However, the results revealed that there were no gender differences for anger expressions. An interpretation of this is that while dominant and masculine looking women are viewed as less trustworthy, this may not be applicable to angry looking women. Indeed, there is evidence that female facial expressions of anger are rated lower on attractiveness, when this was not the case for male posers (Hester, 2019). Perhaps gender differences for anger expressions are related more to ratings of attractiveness and the initial prediction that ratings for dominant looking women would be reflected for angry looking women was incorrect.

The female poser was higher on dominance and trustworthiness for happy expressions and higher on trustworthiness for sad expressions. One explanation for this relates to emotion perception, in that participants perceived sad and happy expressions in female posers quicker than male posers (Becker et al., 2007; Parmley & Cunningham, 2014). It could be that participants perceive and overgeneralise these expressions to traits quicker and to a higher degree in female posers. Finally, this study found that scared expressions were higher on dominance for the female poser. This finding was unexpected as scared expressions do not tend to result in strong dominance ratings overall. Analysing if specific combinations of emotions in stimuli have stronger effects for female or male posers was beyond the scope of this study, but could be a route for future research.

Minority Cue Preferences

The body expressing the minority emotion led to stronger trustworthiness ratings for anger and fear. This is inconsistent with hypothesis 3, as it was predicted that when the face as the minority emotion would have the strongest ratings, and that ratings would be in the predicted direction. Analysing the minority emotional expression provided insight into which cues are preferred when the other cues are different. Past research on emotion perception revealed a preference for facial expressions overall, and that vocalisations had greater preference than body expressions (Casey et al., 2021).

Evidence would suggest that emotions expressed by the body are linked to trait judgements in similar ways as faces (Van der Zant et al., 2021). This would suggest that, even when there is a preference for body expressions, ratings would be similar to those found for facial expressions. Anger and scared facial expressions were not considered high on trustworthiness in the majority ratings in this study, or in previous research (Montepare & Dobish, 2003; Said et al., 2009). For ratings of trustworthiness for happy expressions, and other emotions where we would expect a higher rating, the body as the minority was usually rated higher than the voice. Although not always rated higher than when the face was the minority. Therefore, even in the expected direction the minority face expressions did not lead to higher ratings than the body.

One explanation for this discrepancy between emotion perception and trait judgement is that the way people perceive emotions is different to the way people judge others for multiple cues. It may indicate that when participants are asked to specify emotion, they will prefer to assess this information from the face, while when evaluating traits, participants will attend to the body or most obvious expression.

Body expression research would suggest that highly expansive movement is judged as more dominant (Koppensteiner, 2013), while trustworthiness ratings depended on the nature of movement (Koppensteiner et al., 2016; Thoresen et al., 2012). The anger and scared body expressions in this study could be considered as having higher movement amplitude, as only these expressions featured raised arms and had more noticeable movement. Thus, these expressions may have led to a higher trustworthiness rating when in the minority. However, it would be expected that this would result in higher ratings for dominance as well, which was not the case. While it is unexpected that certain emotions were rated higher for the body, it is clear that trait judgements have a different mechanism for people than emotion perception. Especially given the current study's use of the same stimuli as the referenced study (Casey et al., 2021).

When the voice was the minority cue, ratings were mostly lower than when the face or body were the minority. This is consistent with previous research that found that participants pay more attention to visual cues when audio cues are also present (Zupan et al., 2015). While the voice as the minority was not rated as significantly higher than the body for any trait, anger as the minority voice cue was higher than the face on dominance ratings. This is not unexpected especially due to the high ratings of dominance when anger is the majority emotion. This would imply that anger as a voice cue specifically stands out as an expression and effects the judgement of other cues in the stimuli, at least more than angry facial expressions did.

Gender of the Poser and Minority Cue

An analysis of gender differences for the minority cue stimuli revealed that there were significant differences found for each cue, emotion, and trait. First, the male poser with a scared minority cue was rated higher on trustworthiness. Given results when emotions were in the majority, it is unexpected that the male poser would be higher in ratings than the female poser.

These results imply that when a male poser looks scared in the face, differently from other cues, they look more trustworthy than the female with the same expressions.

For the majority emotion analysis, it was predicted that the female would be rated lower on trustworthiness than the male poser for angry expressions. While this was not found for anger in either majority or minority analysis, the current results suggest that participants rated scared facial expressions differently than other emotions. As previously discussed, dominant looking female faces are rated more negatively (Sutherland et al., 2015) and angry expressions in females are rated low on attractiveness (Hester, 2019). As far as we know, this counter-stereotypical effect is not found for scared expressions in females. However, our stimuli utilised incongruent expressions and these results specifically refer to when scared facial expressions are the minority. It could be the case that a scared face in combination with other expressions appears less trustworthy in female posers. It is interesting then that the female poser's minority scared expressions in the body were rated as more trustworthy than the male poser. These findings suggest that participants integrated the minority facial and body expressions differently for each cue and each gender.

The only differences in dominance ratings were when the female had a minority scared voice cue. This is consistent with majority analysis, where mostly scared expressions were rated higher on dominance for the female poser. Although, typically scared expressions do not rate highly on dominance (Said et al., 2009). An interpretation of this finding is that in combination with other emotional cues, the scared voice sounds more dominant. As the scared vocalisation was a scream, this could be interpreted as more aggressive and angry, especially when paired with angry expressions. This together with the other findings related to scared expressions

suggest that for multimodal cues, scared expressions are interpreted differently between the male and female posers for both traits tested in this study.

Congruent vs. Minority Stimuli

Results comparing congruent emotional expressions with the minority emotion revealed which emotions had the strongest ratings when all cues had the same expression. Hypothesis 4 predicted that there would be differences in ratings between the minority emotion stimuli and the congruent stimuli. There is partial support for this hypothesis as there were differences in ratings between congruent and minority stimuli for some emotions.

The emotional expressions with the biggest differences between congruent and minority emotion were anger for dominant ratings and happy for trustworthy ratings. Both were higher on ratings for congruent stimuli. Happiness and anger expressions are often rated higher on traits than other emotions for facial expressions, (Oosterhof & Todorov, 2009; Said et al., 2009) and they signal approach and avoidance behaviours more than other emotions (Engell et al., 2010; Todorov, 2008). Therefore, it is not unexpected that these expressions would lead to higher ratings. Anger and happiness being higher on traits would suggest that when there is no ambiguity in expression for happiness and anger, people will have stronger judgments.

Inconsistent with hypothesis 4 were results for happy expressions for dominance ratings and angry expressions for trustworthiness ratings, where there were no significant differences between ratings for the minority emotion stimuli compared to congruent stimuli. It was expected that happiness would not be rated low or high on dominance (Said et al., 2009). Results would suggest that participants would not judge happy expressions strongly on dominance whether it was in the minority or completely congruent. It is interesting then that there were no differences in trustworthiness for anger expressions. Anger is often rated lower on trustworthiness and

signals avoidance behaviours in participants (Engell et al., 2010). If congruent anger expressions are rated significantly higher on dominance then when in the minority, we may also expect congruent anger expressions to be rated significantly lower on trustworthiness than when in the minority, however this was not the case. In fact, trustworthiness ratings overall had smaller differences between congruent and minority emotion, except for happiness. Thus, emotions linked to avoidance behaviour - anger, sadness and fear (Engell et al., 2010; Todorov, 2008) - were judged as higher in dominance when cues were congruent. Perhaps the stronger ratings are not specifically influenced by anger and happiness, but have more to do with general avoidance emotions invoking a higher dominance rating, and approach emotions leading to higher ratings of trustworthiness when 3 cues are present.

Limitations and future directions

While the present study provides research for a literature gap in trait judgement, there are some limitations. Firstly, the stimuli in this study were originally created for use in research into emotion perception (Nelson & Mondloch, 2017). While the stimuli were validated for accurate emotion perception, there may be limitations for use in trait judgement research. This study provides evidence that incorporating face, body and voice cues has effects on how participants rated posers. Therefore, future research should further examine this with multimodal stimuli that are created and validated for trait judgement research.

Secondly, a potential limitation was that most ratings on trustworthiness and dominance were lower than 3 on the 5 point Likert scale (1: *Not at all* to 5: *Completely,* Appendix B*)*. Meaning that participants didn't rate stimuli as highly trustworthy or dominant. While analysis mainly involved comparison between other cues, emotions and traits, it is worth noting that this

comparison used ratings from a small section of the scale. This could be addressed by revising the scale in future multimodal research designs.

A key factor in this study involved differences in how participants rated the male and female poser. Another potential limitation was that the stimuli had only one male and female poser. While there were differences in how participants rated the posers, the present research cannot rule out the possibility of differences unrelated to gender stereotypes effecting results. Factors such as, attractiveness and facial typicality have effects on trait ratings, and these factors also have implications for gender stereotypes (Hester, 2019; Sofer et al., 2015). Future research designs should include multiple female and male posers to further test the implications of multimodal stimuli on gender stereotypes and subsequent trait judgements.

Further suggestion for future research would be to analyse ambiguous or androgynous posers. This idea stems from prior research that found that participants viewing emotionally ambiguous faces relied more on stereotypes to guide their judgements (Parmley & Cunningham, 2014). It would be interesting to apply ambiguous cues to body and voice expressions, along with multimodal stimuli. The use of more androgynous looking posers would provide insight into how this effects trait ratings for multimodal stimuli. This would add to research suggesting that analysis of androgynous faces provides novel insights on the nature femininity and masculinity, while contributing to the inclusion of more gender diverse research (Hester et al., 2020).

Conclusion

This study provides the first analysis of trait judgement for multimodal cues. Consistent with emotion perception research, participants have preference for the emotion that is expressed by the majority when making judgments of trustworthiness and dominance. Examining ratings

for emotions expressed by the minority revealed that participants attended to the body expressions more often than face or voice. This was inconsistent with previous research where there was preference for facial expressions. This provides evidence that participants use different information when making trait judgements than when perceiving emotion.

The study provides evidence for gender differences in how participants view multimodal stimuli and the potential influences of gender stereotypes. The female poser in this study was generally rated as higher on trustworthiness and dominance. However, the analysis of the minority cue revealed that there were gender differences for cue type preference, and differences in how emotions in certain cue types were perceived. The implications of this warrants the analysis of gender differences in future multimodal research. Finally, results suggested that the congruency of cues have larger effects for certain emotions, namely anger and happiness. Future research could build on this study with the inclusion of a range of stimuli, and further analysis of how specific emotion pairings effect trait judgment.

References

Antonakis, J., & Eubanks, D. L. (2017). Looking Leadership in the Face. *Current Directions in Psychological Science*, *26*(3), 270–275. https://doi.org/10.1177/0963721417705888

Åsli, O., & Øvervoll, M. (2020). Model Gender Interacts With Expressed Emotion to Enhance Startle: Angry Male and Happy Female Faces Produce the Greatest Potentiation. *Frontiers in Human Neuroscience*, *14*, 576544. https://doi.org/10.3389/fnhum.2020.576544

Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, *33*, 154–161. https://doi.org/10.1016/j.copsyc.2019.07.028

Aviezer, H., Trope, Y., & Todorov, A. (2012). *Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions*. *338*, 6.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Calvo, M. G., Krumhuber, E. G., & Fernández-Martín, A. (2019). Visual attention mechanisms in happiness versus trustworthiness processing of facial expressions. *Quarterly Journal of Experimental Psychology*, *72*(4), 729–741. https://doi.org/10.1177/1747021818763747

Casey, C., Yung, M., & Nelson, N. L. (2021). *Are you mad at me? A study of multimodal emotion perception.* [Manuscript in Preparation]. School of Psychology, University of Queensland.

Chua, K.-W., & Freeman, J. B. (2021). Facial Stereotype Bias Is Mitigated by Training. *Social Psychological and Personality Science*, *12*(7), 1335–1344. https://doi.org/10.1177/1948550620972550

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Elkins, A. C., & Derrick, D. C. (2013). The Sound of Trust: Voice as a Measurement of Trust During Interactions with Embodied Conversational Agents. *Group Decision and Negotiation*, *22*(5), 897–913. https://doi.org/10.1007/s10726-012-9339-x

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit Trustworthiness Decisions: Automatic Coding of Face Properties in the Human Amygdala. *Journal of Cognitive Neuroscience*, *19*(9), 1508–1519. https://doi.org/10.1162/jocn.2007.19.9.1508

Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Common Neural Mechanisms for the Evaluation of Facial Trustworthiness and Emotional Expressions as Revealed by Behavioral Adaptation. *Perception*, *39*(7), 931–941. https://doi.org/10.1068/p6633

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

Flowe, H. D. (2012). Do Characteristics of Faces That Convey Trustworthiness and Dominance Underlie Perceptions of Criminality? *PLoS ONE*, *7*(6), e37253. https://doi.org/10.1371/journal.pone.0037253

Freeman, J. B., Stolier, R. M., Ingbretsen, Z. A., & Hehman, E. A. (2014). Amygdala Responsivity to High-Level Social Information from Unseen Faces. *Journal of Neuroscience*, *34*(32), 10573–10581. https://doi.org/10.1523/JNEUROSCI.5063-13.2014

Halovic, S., & Kroos, C. (2018). Walking my way? Walker gender and display format Confounds the perception of specific emotions. *Human Movement Science*, *57*, 461–477. https://doi.org/10.1016/j.humov.2017.10.012

Hareli, S., Shomrat, N., & Hess, U. (2009). Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion*, *9*(3), 378–384. https://doi.org/10.1037/a0015958

Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). 'Worth a thousand words': Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, *9*(3), 293–305. https://doi.org/10.1037/a0015178

Hess, U., Blairy, S., & Kleck, R. E. (2000). The Influence of Facial Emotion Displays, Gender, and Ethnicity on Judgments of Dominance and Affiliation. *JOURNAL OF NONVERBAL BEHAVIOR*, 19.

Hester, N. (2019). Perceived negative emotion in neutral faces: Gender-dependent effects on attractiveness and threat. *Emotion*, *19*(8), 1490–1494. https://doi.org/10.1037/emo0000525

Hester, N., Jones, B. C., & Hehman, E. (2020). Perceived femininity and masculinity contribute independently to facial impressions. *Journal of Experimental Psychology: General.*, *Advance online publication.* https://doi.org/10.1037/xge0000989

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., Ndukaihe, I. L. G., Bloxsom, N. G., Lewis, S. C., Foroni, F., Willis, M. L., Cubillas, C. P., Vadillo, M. A., Turiegano, E., Gilead, M., Simchon, A., Saribay, S. A., Owsley, N. C., Jang, C., … Coles, N. A. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour*, *5*(1), 159–169. https://doi.org/10.1038/s41562-020-01007-2

Koppensteiner, M. (2013). Motion cues that make an impression. *Journal of Experimental Social Psychology*, *49*(6), 1137–1143. https://doi.org/10.1016/j.jesp.2013.08.002

Koppensteiner, M., Stephan, P., & Jäschke, J. P. M. (2016). Moving speeches: Dominance, trustworthiness and competence in body motion. *Personality and Individual Differences*, *94*, 101–106. https://doi.org/10.1016/j.paid.2016.01.013

Lin, C., Adolphs, R., & Alvarez, R. M. (2017). Cultural effects on the association between election outcomes and face-based trait inferences. *PLOS ONE*, *12*(7), e0180837. https://doi.org/10.1371/journal.pone.0180837

Lin, C., Keles, U., & Adolphs, R. (2021). Four dimensions characterize attributions from faces using a representative set of English trait words. *Nature Communications*, *12*(1), 5168. https://doi.org/10.1038/s41467-021-25500-y

Montepare, J. M., & Dobish, H. (2003). The Contribution of Emotion Perceptions and Their Overgeneralizations to Trait Impressions. *JOURNAL OF NONVERBAL BEHAVIOR*, 18.

Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4–6), 563–574. https://doi.org/10.1080/13506285.2017.1301615

Oh, D., Buck, E. A., & Todorov, A. (2019). Revealing Hidden Gender Biases in Competence Impressions of Faces. *Psychological Science*, *30*(1), 65–79. https://doi.org/10.1177/0956797618813092

Oh, D., Dotsch, R., Porter, J., & Todorov, A. (2020). *Gender Biases in Impressions From Faces: Empirical Studies and Computational Models*. 20.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and

    trustworthiness impressions from faces. *Emotion*, *9*(1), 128–133.

    https://doi.org/10.1037/a0014520

Over, H., & Cook, R. (2018). Where do spontaneous first impressions of faces come from?

    Cognition, *170*, 190–200. https://doi.org/10.1016/j.cognition.2017.10.002

Parmley, M., & Cunningham, J. G. (2014). She Looks Sad, But He Looks Mad: The Effects of

    Age, Gender, and Ambiguity on Emotion Perception. *The Journal of Social Psychology*,

    154(4), 323–338. https://doi.org/10.1080/00224545.2014.901287

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions

    predicts evaluation of emotionally neutral faces. *Emotion*, *9*(2), 260–264.

    https://doi.org/10.1037/a0014681

Schild, C., Stern, J., & Zettler, I. (2019). Linking men's voice pitch to actual and perceived

    trustworthiness across domains. *Behavioral Ecology*, arz173.

    https://doi.org/10.1093/beheco/arz173

Sofer, C., Dotsch, R., Wigboldus, D. H. J., & Todorov, A. (2015). What Is Typical Is Good: The

    Influence of Face Typicality on Perceived Trustworthiness. *Psychological Science*, *26*(1),

    39–47. https://doi.org/10.1177/0956797614554955

Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies

    in static and dynamic scenes. *Frontiers in Psychology*, *5*.

    https://doi.org/10.3389/fpsyg.2014.00193

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A.

    W. (2013). Social inferences from faces: Ambient images generate a three-dimensional

    model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and

    stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are

    negatively evaluated. *British Journal of Psychology*, *106*(2), 186–208.

    https://doi.org/10.1111/bjop.12085

Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive

    reliable trait judgements. *Cognition*, *124*(3), 261–271.

    https://doi.org/10.1016/j.cognition.2012.05.018

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition

    of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of*

    Sciences, *1124*(1), 208–224. https://doi.org/10.1196/annals.1440.012

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from

    Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review*

    of Psychology, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Van Der Zant, T., Reid, J., Mondloch, C. J., & Nelson, N. L. (2021). The influence of postural

    emotion cues on implicit trait judgements. *Motivation and Emotion*, *45*(5), 641–648.

    https://doi.org/10.1007/s11031-021-09889-z

Wilson, J. P., & Rule, N. O. (2015). Facial Trustworthiness Predicts Extreme Criminal-

    Sentencing Outcomes. *Psychological Science*, *26*(8), 1325–1331.

    https://doi.org/10.1177/0956797615590992

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: Group

    differences, impression effects, and race stereotypes. *Journal of Personality and Social*

    Psychology, *98*(2), 175–189. https://doi.org/10.1037/a0017990

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why Appearance Matters: Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zupan, B., Babbage, D. R., & Sussman, J. E. (2015). *Modality Dominance in the Perception of Incongruent Bimodal Emotion Expressions*. *1*(1), 12.

Appendix A: Emotional Expressions

Happy   Sad   Angry   Scared

Happy   Sad   Angry   Scared

Appendix B: Likert Scale Matrix

Please press Play on the video. Do you think this person is..

	Not at all	A little bit	Neither A little nor A lot	A lot	Completely
Trustworthy	○	○	○	○	○
Dominant	○	○	○	○	○
Competent	○	○	○	○	○

→