Estimating the effects of trait knowledge on social perception

Wildman, Andrew; Ramsey, Richard

Quarterly Journal of Experimental Psychology

DOI:
10.1177/17470218211047447

Published: 01/05/2022

Peer reviewed version

Estimating the Effects of Trait Knowledge on Social Perception

Andrew Wildman[a] & Richard Ramsey[b,*]

[a] Wales Institute for Cognitive Neuroscience, School of Psychology, Bangor University, Bangor, Gwynedd, Wales, LL57 2AS, United Kingdom

[b] Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia.

Word count: TBC

* Corresponding author

E-mail: richard.ramsey@mq.edu.au

Keywords: social cognition; person perception; priming; trait inference; body size

Abstract

Research in social cognition has predominantly investigated perceptual and inferential processes separately, however real-world social interactions usually involve integration between person inferences (e.g., generous, selfish) and the perception of physical appearance (e.g., thin, tall). Therefore, in the current work, we investigated the integration of different person-relevant signals, by estimating the extent to which bias in one social information processing system influences another. Following an initial stimulus-validation experiment (Experiment 1, N=55), two further pre-registered experiments (Experiments 2, N=55 & 3; N=123) employed a priming paradigm to measure the effects of extraversion-diagnostic information on subsequent health and body-size judgements of a target body. The results were consistent across both priming experiments and supported our predictions: compared to trait-neutral control statements, extraversion-diagnostic statements increased judgements of health and decreased those of body size. As such, we show that trait-based knowledge does not only influence mappings towards similar types of person judgments, such as health judgments. Rather, even a brief re-configuration of trait-space alters mappings towards non-trait judgments, which are based on body size and shape. The results complement prior neuroimaging findings that showed functional interactions between the body-selective brain regions in the ventral visual stream and the theory of mind network when forming impressions of others. Therefore, we provide a functional signature of how distinct information processing units exchange signals and integrate information in order to form impressions. Overall, the current study underscores the value of behavioural work in complementing neuroscience when investigating the role and properties of functional integration during impression formation. Additionally, it stresses the potential limitations of an over-reliance on studying separate systems in isolation.

1. Introduction

When interacting with another person, we combine many distinct features and recognise that these belong to a single entity. For example, physical features, such as what someone looks like (e.g., tall and slim) are integrated with judgments regarding their character (e.g., outgoing and friendly). Such integrated person representations coordinate social behaviour by signalling who to approach and avoid, as well as how and when to interact with others. Although distinct person features must clearly be integrated, researchers studying the neurobiological underpinnings of social cognition typically address perceptual and inferential processes separately. Consequently, the nature of interplay between perceptual and inferential person representations is largely unknown. For example, it is unclear to what extent holding a particular social judgement about someone (e.g., friendly) might bias how we "see" them in a perceptual sense (e.g., slimmer). In the current study, therefore, we estimate the impact of drawing trait inferences on person perception. By doing so, we aim to build new links between two sub-disciplines of social cognition and assess the hypothesis that a holistic person representation in part comprises reciprocally connected person feature representations.

Research in social cognition and social neuroscience has largely focussed on understanding how separate sub-systems operate during social information processing, which span perceptual, cognitive and emotional processes (Adolphs, 2009; Frith & Frith, 2012). For example, person perception research aims to understand systems whose roles include detecting the presence and appearance of others (Kanwisher, 2010). In contrast, person inference research is focussed on investigating the systems that enable one to reason and make inferences about other people's "hidden" mental states and trait-based character (Frith & Frith, 1999; Saxe & Kanwisher, 2003; van Overwalle, 2009). These sub-disciplines of social cognition research have made significant advances to understanding social information processing, whilst largely remaining separate research entities that operate in their own silos with little communication.

In everyday life, however, we integrate multiple sources of information to form complete person representations, which are likely to encompass the interaction of perceptual and inferential processes. For example, the identification of another person's face or body often leads to spontaneous person inferences, whereby trait-based character impressions are formed on limited or incomplete social cues (Ambady & Rosenthal, 1992; Todorov et al., 2015). Indeed, one of the most studied aspects of impression formation concerns traits imbued by facets of a person's visual appearance, including facial expressions, body shape, gestures and posture (Naumann et al., 2009; Oosterhof & Todorov, 2008; Puhl & Heuer, 2009). However, we do not solely rely on visual appearance to form judgments of people's character. Trait-diagnostic information can be extracted from the perception of others' behaviour, whether observed directly or learned about indirectly, such as when talking with a friend or when reading a book (Mitchell, 2009; Mitchell et al., 2006). Furthermore, accurate visual representations of body shape can be derived from verbal descriptions, which shows a close link between verbal and visual body representations (Hill, Streuber, Hahn, Black, & O'Toole, 2016). Ultimately, therefore, disparate modalities of person-specific information (visual percept, written or spoken word), are integrated to form a single holistic person-representation. Therefore, without studying perceptual and inferential processes together, it seems difficult to build a more complete understanding of how holistic person representations manifest.

To date, the study of person perception has been dominated by research on faces (e.g. Kanwisher et al., 1997; Todorov et al., 2015). Bodies, however, also signal important social information (de Gelder, 2006; de Gelder et al., 2010; Hu et al., 2018), and at times express unique information that faces conceal (Aviezer et al., 2012). Moreover, given globally increasing obesity rates (Wang et al., 2011), body weight is becoming an ever more salient dimension along which people can vary, which is likely to elevate the social consequences of inferences based on body shape (Puhl & Heuer, 2009). Indeed, from a public health perspective, the nature and content of trait judgements that arise from perceptions of weight have been shown to have negative health consequences for those individuals perceived as being overweight (Daly et al., 2019). Understanding the role of body perception in social cognition, therefore, has downstream implications for understanding and remediating the processes which may lead to potentially damaging prejudice and stigmatisation.

The separation of research specialisations into perceptual and inferential processes is mirrored by a focus within these sub-disciplines on largely non-overlapping brain circuits. Indeed, two largely separate neural circuits have been associated with body perception and person inference. In terms of body perception, brain regions along the ventral visual stream in the Extrastriate Body Area (EBA; Downing, Jiang, Shuman, & Kanwisher, 2001) and Fusiform Body Area (FBA; Peelen & Downing, 2007; Schwarzlose, Baker, & Kanwisher, 2005) show greater activation in response to bodies or body parts in comparison to control stimuli such as chairs and cars (Downing & Peelen, 2011). Together, it has been argued that EBA and FBA are primarily sensitive to body shape and posture processing, rather than more elaborate cognitive processes such as emotion or identity processing (Downing & Peelen, 2011; Kemmerer, 2011).

The second neural system of relevance to the current work is one associated with person inferences. The mentalising or theory of mind network is a system of regions which engage when mental states such as beliefs, desires and attitudes are ascribed to others (Frith & Frith, 1999). The theory of mind network spans the temporo-parietal junction (TPJ), medial prefrontal cortex (mPFC), anterior cingulate cortex (ACC), temporal poles (TPs), Precuneus (PreC) and superior temporal sulcus (STS) (Saxe & Kanwisher, 2003; van

Overwalle, 2009). The theory of mind network is thought to be responsible for generating inferences about people on the basis of learned or observed behaviour, such as whether they are outgoing or friendly, and as such, it is has been associated impression formation (Mitchell et al., 2005, 2006).

Much like human neuroscience in general, social neuroscience research has primarily identified the function of segregated brain networks, which span perceptual, cognitive and affective processes (functional segregation; see Adolphs, 2009; Kanwisher, 2010). Less research has investigated the function of interplay between multiple systems (functional integration; Bullmore & Sporns, 2009; Park & Friston, 2013). Newer research in social neuroscience is beginning to emerge, however, which places greater emphasis on understanding functional integration between component processing units. For instance, with regard to body perception and trait inference research, neuroimaging studies have demonstrated functional coupling between body perception and theory of mind regions during impression formation when participants are presented with trait-diagnostic information alongside an image of a person's body (Ramsey, 2018). Such functional integration between neural circuits associated with person perception and person inference have been shown to be involved when forming impressions (Greven et al., 2016), as well as when recalling stored social knowledge (Greven & Ramsey, 2017a) and evaluating ingroup versus outgroup members (Greven & Ramsey, 2017b). Therefore, these studies are beginning to demonstrate that for a more complete understanding of social information processing during body perception, functional integration must be considered alongside functional segregation (Quadflieg et al., 2011; Ramsey, 2018; Ramsey et al., 2011).

The demonstration of functional coupling between distinct neural networks when forming impressions is only a starting point, however. The functional relevance of this interplay is still poorly understood. Indeed, neuroscience research needs behavioural research to help provide a relevant context to interpret brain-based findings (Krakauer et al., 2017). Key questions remain unanswered concerning the nature and structure of links between 'trait space' and 'face/body space' when forming impressions (Over & Cook, 2018). How and when are distinct person features bound together? What are the functional consequences of reconfiguring 'trait space' for judgments that rely on 'face/body space'? Indeed, the consequences of delivering mutually relevant person information to two separate systems has not received much attention. Historically, much more research has investigated how multiple features within a single modality are weighted to produce an overall percept or judgment state (Anderson, 1962; Asch, 1946; Hendrick et al., 1975).

The current behavioural work, therefore, seeks to address this gap in understanding by estimating the extent to which a trait-based person inference can influence other types of person inference and person perception. A considerable amount of prior work has investigated how images of faces and bodies trigger spontaneous trait inferences (Greven et al., 2019; Naumann et al., 2009; Puhl & Heuer, 2009; Sutherland et al., 2013; Todorov et al., 2015). Here, we test the opposite flow of information by hypothesising that person inferences generated in the theory of mind network can influence other person inferences, as well as person perception processes in the ventral visual stream. More specifically, we hypothesise that forming a person inference based on trait knowledge (e.g., extraversion) will have functional consequences for related person inferences (e.g., health), as well as purely shape-based body judgments (e.g., size and shape). Such findings would suggest that re-structuring 'trait space' with new person information can generalise and bias judgments of other types of person inference that place similar demands on person inference systems (e.g., health judgments), as well as judgments that place low demands on person inference systems and that largely rely on visual feature processing along the ventral visual stream (e.g., body-size judgments).

Investigating the relationships between distinct types of person knowledge is important for several reasons. First, in terms of understanding basic cognitive and brain systems, the findings illuminate how and when separate social information processing systems integrate information across ‘trait space’ and ‘face/body space’ (Over & Cook, 2018). This is important due to the lack of research that focusses on understanding functional integration in general (Park & Friston, 2013) and in social perception research (Kanwisher, 2010; Ramsey, 2018). By doing so, the current work will provide a functional description of the links between perceptual and inferential processes during body perception, and thus build new links between sub-disciplines of psychology and neuroscience that typically do not overlap. Second, on a more societal and social level, given the health consequences for those individuals who are perceived as being overweight (Daly et al., 2019), as well as the growing obesity rates globally (Wang et al., 2011), understanding the mechanisms that might mediate such perceptions could have important longer-term consequences for society.

The current paper comprises three experiments. The first experiment was primarily focussed on developing stimuli to make sure that we select bodies that cue the required type of person inferences. The two subsequent experiments then use these bodies to test if trait inferences regarding a person’s character bias judgments based on body shape. We chose extraversion as an example of a trait inference to test our general question of interest, but other dimensions and features would also have addressed the same basic question.

2. Experiment 1 – Stimuli Development

2.1 Introduction

Although prior research has established how clearly distinct body shape exemplars (e.g., muscly versus obese) impact trait inferences (Greven et al., 2019), the nature of trait attributions across small intervals of body shape/size dimensions remain largely unknown. Experiment 1, therefore, sought to establish the relationship between intervals of body size (from low to high body fat) and trait judgements and thereby validate which body stimuli would ultimately be used in subsequent priming studies. Computer-generated body images were created using *MakeHuman* (version 1.1.1; www.makehumancommunity.org), a python-based program for creating anatomically realistic 3D human models (*toons*). The basic model was adjusted to produce both a slim and overweight archetype, which were then saved as target meshes so that toons could be created procedurally across increments of these extremes.

This preparatory experiment sought to establish the relative change in visual and trait-based ratings of bodies across increasing increments of body size, by asking participants to make judgements about a series of 15 body-sizes in response to four questions: “How outgoing?”, “How attractive?”, “How healthy?” and “How heavy?”. The purpose was to map out the responses to each and establish their independence from each other; it was expected that the incremental changes in response to the body size question (How heavy?) would be roughly the inverse of those observed in response to the others. Attractiveness was included on the basis that this may otherwise be used as a heuristic for the other trait ratings.

2.2 Method

Our first experiment sought to identify the judgements made about a series of newly developed body stimuli. We sought to collect judgments across 50 participants in order to provide a reasonable index of typical responses to our dependent variables. Given that the results in this initial experiment were expected to be relatively clear, a target of 50 participants was judged to be sufficiently powerful for the purposes of estimating the average rating for each body size increment. In addition, a total of 50 per cell of a design is increasingly considered the minimum sample size for conventional psychological research given the reduced ability of smaller sample sizes to produce robust estimates of effect sizes (Simmons et al., 2018).

2.2.1 Participants

Fifty-five participants took part in the study in exchange for monetary compensation or course credit (13 males, $Mean_{age}$=24.15, SD_{age}=5.05, age range = 18 to 38). All participants provided informed consent before completing the task. Participants were excluded from a given cell of the design if their mean response for that combination of factors (15 body size increment and 4 questions) was 2.5 standard deviations from the mean of that cell. This criterion excluded 1% of data points, and the minimum number of participants within a cell was 53. Thirty-three of the cells included all 55 participants. All procedures were approved by the Research Ethics and Governance Committee of the School of Psychology at Bangor University.

2.2.2 Materials

A short script was produced using the coding utility in *MakeHuman*, to create and render JPG images of toons ranging from low to high body fat. Four different identities (2 male and 2 female) which differed in skin texture were created at 15 body size increments, resulting in a total of 60 bodies. Basic clothing assets (white underwear) were downloaded from the *makehumancommunity.org* forum and added to the toons before they were screen grabbed as 1523x882 PNG images. These were then cropped to 785x774 and had their faces obscured with a solid black square (see Figure 1.).

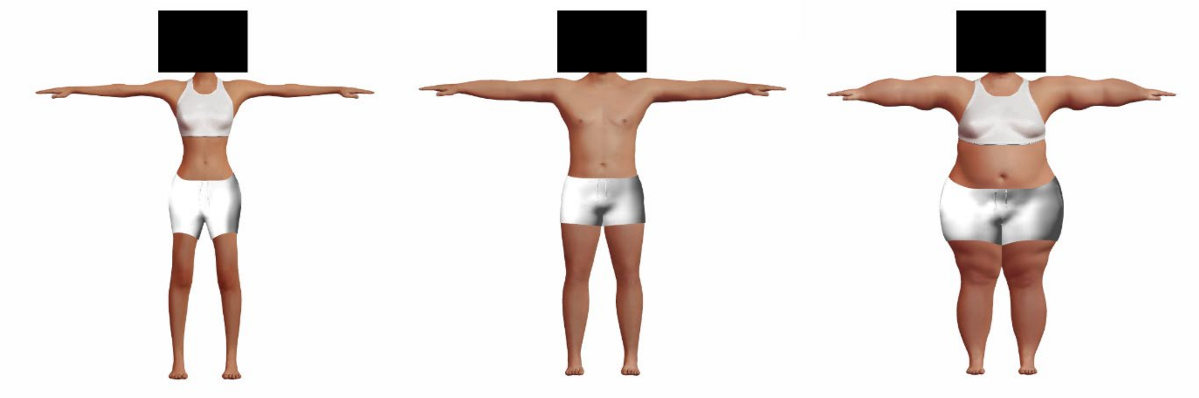

Figure 1. Example stimuli. Body sizes 1 (left), 8 (middle) and 15 (right)

2.2.3 Task and Procedure

A body-rating task was produced and implemented in *MATLAB* 2015b, using Psychtoolbox 3 (www.psychtoolbox.org). On each trial participants were presented with a body and a question, which they had to respond to with a 1-9 key press within 6 seconds (see Figure 2.). Participants were advised that they could take a break after every 40 trials, and press space to resume the task. In total the task had 240 trials.

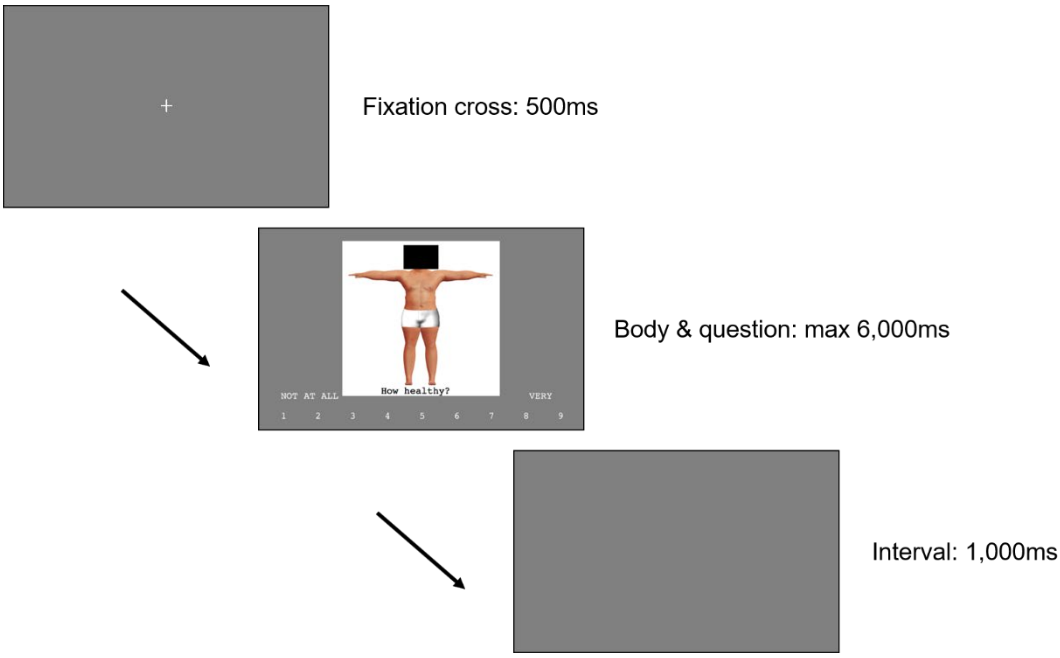

Figure 2. Trial of body-rating task

2.3 Results and Discussion

Means and 95% confidence intervals were calculated and plotted for each combination of body size increment and dependent variable (Figure 3). With the exception of more “extreme” body sizes at the thin end of the range, increasing increments of body size generally brought about lower ratings of health, extraversion and attractiveness on an incremental basis. In contrast, body size judgements generally increased across increasing increments of size (see Figure 3.). Cronbach’s Alphas were also calculated for each dependent variable of interest showing high consistency and agreement across measures (see Supplementary Table 9). Previous studies investigating size judgements of incrementally increasing body sizes of real or computer-generated bodies have observed a sigmoid curve as ratings of body sizes at the extreme ends of the scale are less noticeable (Weber's law) (Alexi et al., 2019; Cornelissen et al., 2016). It is possible that this is caused in part by participants' tendency to avoid the extreme ends of a finite Likert scale, and that measurement error can only occur in one direction once the end of the scale is reached. As such, it is not clear whether impeded size estimation at the extreme ends of a stimuli set is an artefact of testing methods or a genuine property of the psychophysics of body perception.

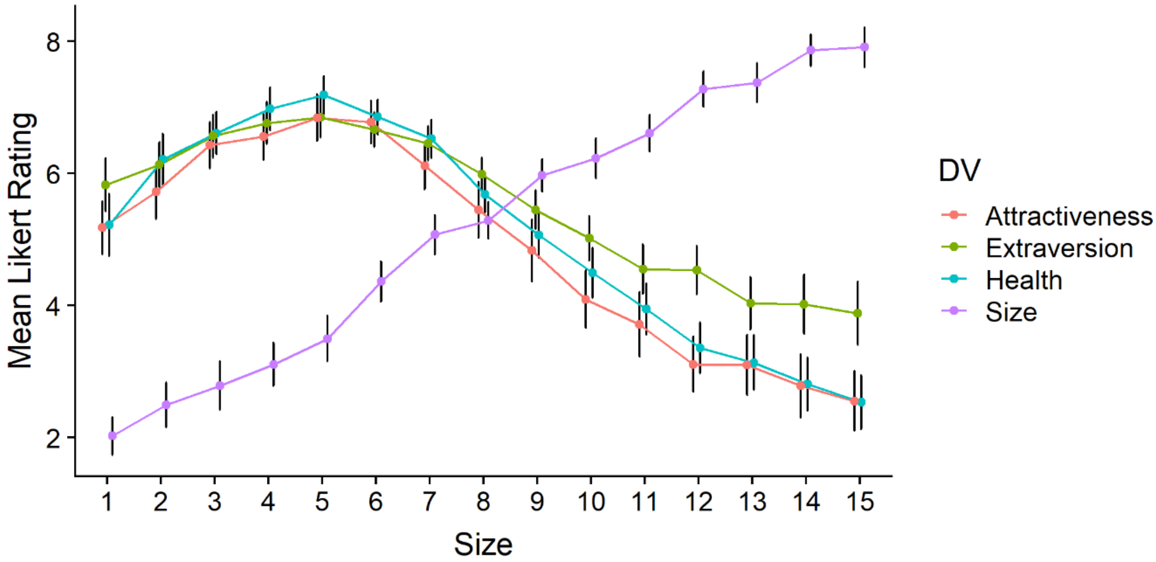

Figure 3. Mean Likert ratings of each variable by body size.

A grand mean and pooled standard deviation were calculated for body sizes 5-12 (as to exclude bodies without a clear change in judgements between increments), and the distances from the grand mean in standard deviation units are shown below (see Table 1.).

	Body Size							
	5	6	7	8	9	10	11	12
How Outgoing?	1.24	1.04	0.82	0.32	-0.26	-0.72	-1.22	-1.24
How Attractive?	1.22	1.18	0.71	0.23	-0.20	-0.72	-0.99	-1.42
How Healthy?	1.28	1.04	0.81	0.21	-0.23	-0.64	-1.03	-1.44
How Heavy?	-1.65	-0.95	-0.38	-0.20	0.35	0.56	0.87	1.41

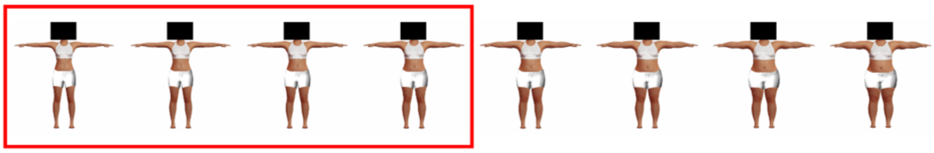

Table 1. Distances from grand mean of bodies 5-12 in standard deviation units. Bodies selected for experiments 2 and 3 are highlighted.

Bodies 5, 6, 7 and 8 were selected to be used in the two subsequent priming experiments. These bodies were selected because they showed relatively large increases in judgments of size, as well as relatively large decreases in the other judgements. As we expect the impact of trait-inference priming on judgments of size to be relatively small, we chose bodies that we thought would maximise our sensitivity to detect a change in judgments of size after manipulating trait judgments.

3. Experiment 2

3.1 Introduction

To investigate how person inferences influence subsequent body-perception, we presented participants with two separate pieces of information about target persons before asking them to make judgements about them. First, we gave participants a statement, which either primed extraversion or was trait-neutral, and then secondly, we showed participants a body image that varied in size and identity across trials. Bodies were subsequently judged on one of three possible dimensions: extraversion ("How outgoing?"), health ("How healthy?"), and body size ("How heavy?").

Extraversion ratings were included as a 'positive control', as these judgements would be expected to increase on prime trials relative to neutral ones. The inclusion of a positive control ensured that participant's judgements of a target's extraversion were affected by our priming stimuli, and that our design was sensitive to effects of priming in general. It also served as a reference point for interpreting effect sizes given that the change in extraversion ratings between conditions would likely be the largest and clearest. Health ratings were included to test whether primed extraversion content would generalise to other person-inferences. The final condition, size ratings, assessed whether the imputed trait information would yield effects on person perception. It was predicted that priming with statements diagnostic of extraversion would increase subsequent judgements of health and decrease those of body-size.

3.2 Method

3.2.1 Participants

Sixty-five Bangor University students were recruited through Bangor University's student participation panel in exchange for course credit (11 males, M_{age}=19.98, SD_{age}=3.27). Our sample size was determined by an a priori power analysis using *G*Power* (Faul et al., 2007), which indicated that a sample of 52 would give 80% power to detect a Cohen's *d* of 0.35 with a one-tailed paired-samples t-test for each of our two dependent variables of interest (health and body-size). This would conventionally be considered a small-to-medium effect (Cohen, 1988). Our stopping rule was therefore to have 52 useable observations by the cessation of data collection. As separate t-tests were used for each dependent variable, final sample sizes differed for each analysis. Following data pre-processing and outlier removal, final sample sizes for each dependent variable were 57 for Extraversion, 57 for Health and 60 for Body Size. Our predetermined experimental design, sample size and analysis approach were pre-registered online (https://aspredicted.org/blind.php?x=65ye4c).

3.2.2 Materials

The four body sizes selected from Experiment 1 (sizes 5, 6, 7 and 8) were used for the priming task. Three body identities were created at these sizes, with minor adjustments made to skin tone and subtle characteristics such as naval position and proportions. All of the stimuli in the experiment were female to reduce the number of permutations required throughout the experiment and thus avoid participant fatigue.

A series of 20 extraversion-diagnostic (prime) statements were produced to reflect five of the adjectives comprising the taxonomy of extraversion as defined by the Big Five Inventory (BFI) (John & Srivastava, 1999). The trait adjective "energetic" was omitted due to its close affiliation with our health dependent measure. Several trait-neutral statements were taken from Mitchell et al. (2006) and supplemented with newly generated ones making a total of 40 statements (20 prime and 20 neutral). These statements were validated by a sample of 15 participants recruited online, who were asked to rate the extent to which each statement reflected behaviour typical of openness, conscientiousness, extraversion, agreeableness, neuroticism, and health. Participants responded on a 5-point Likert scale ranging from "Not at all", to "Extremely". 15 statements with the highest mean extraversion ratings were selected as the priming stimuli, and the 15 lowest were selected as the neutral counterparts (see Supplementary Data 1). The difference between statements was confirmed using a t-test comparing the mean extraversion ratings attributed to each set of 15 statements, which revealed a large difference between statements in the two conditions, $t(14)=10.72$, $p<.001$, $d=2.77$ [1.63, 3.89], mean difference = 2.38 [1.90, 2.85] (square brackets denote 95% confidence intervals for all statistics in the article). Numerically, the priming statements received an average extraversion rating of 4.28, while the neutral ones received an average rating of 1.90.

Although the survey did not explicitly measure introversion, it is possible that low extraversion ratings could reflect a judgement of high introversion. Although this would not greatly affect our predictions, as we expect low levels of extraversion to be associated with a heavier body shape and lower health, there are implications for the interpretation of effect size estimates, which we address later (see General Discussion). Importantly, it should be noted that statements we designate as 'trait-neutral' are likely to contain some trait-diagnostic information, and our experimental conditions could equally be thought of as 'high-extraversion' and 'low-extraversion'.

3.2.3 Task

The experimental task was produced in MATLAB (2015b) using PsychToolbox (version 3; www.psychtoolbox.org). The task involved four body sizes, 30 statements (15 extraversion-diagnostic [prime], 15 trait-neutral [neutral]) and three questions ("How outgoing?" [positive control], "How healthy?", "How heavy?"), all of which were presented in every possible permutation in a single randomised experimental block, giving a total of 360 experimental trials (body identity was selected randomly on each trial). Each trial would commence with a statement appearing on-screen until the participant pressed the space key (e.g. "She went on an exciting road trip across the USA"). A fixation cross was then presented for 500ms, followed by the target body stimulus for 100ms. The body stimulus was then backward-masked for 400ms to reduce the visual after-effect of the image. Finally, one of the three questions appeared and remained on-screen until the participant's response or up to a maximum of 4,000ms (see Figure 4.).

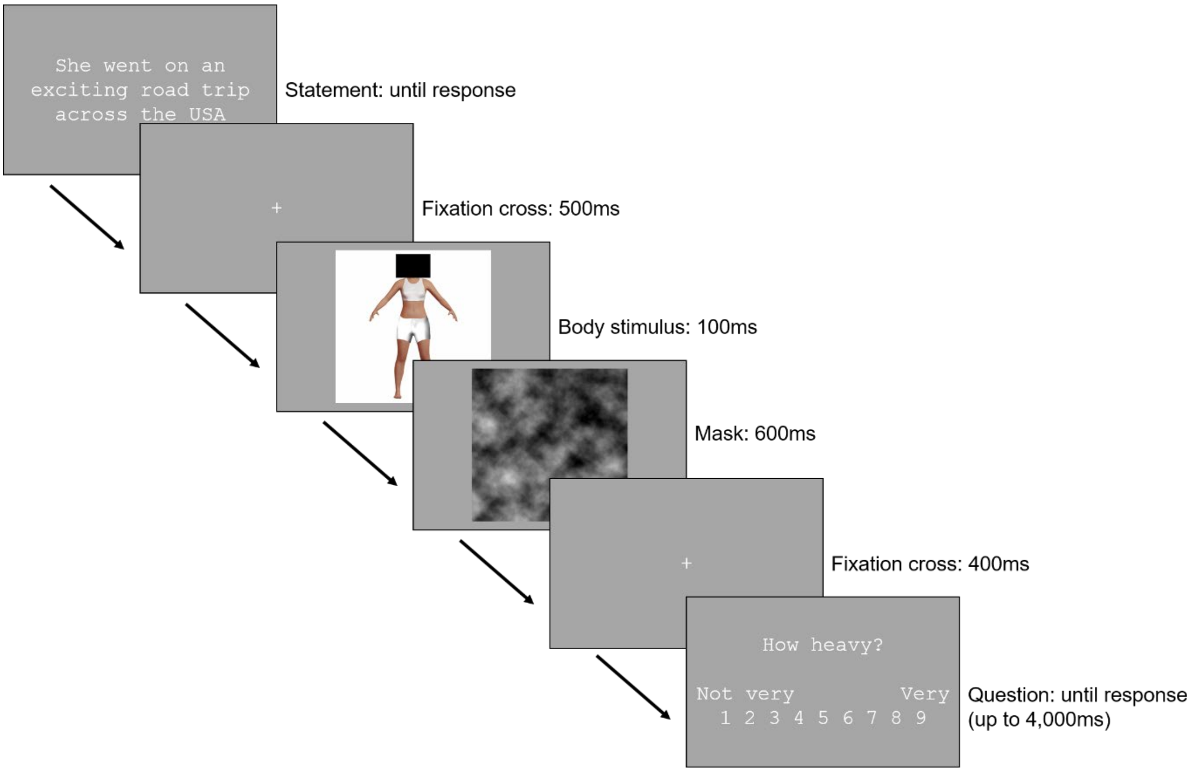

Figure 4. Example of experimental trial in priming task

Every 24 trials, a catch-trial would be initiated. Catch trials began with the usual statement and fixation cross, however instead of a body stimulus, participants were instead presented with a second 'true or false' statement and were asked to press either 1 (false) or 9 (true) with regards to whether the second statement concorded with the first. For example, the extraversion-diagnostic statement: "She spoke to her friend on Skype for an hour", could be followed by "She spoke to her father on Skype", alongside "False" and "True" in place of the "Not very" and "Very" cues.

In addition to the main task, participants filled out a questionnaire measuring basic demographic information and the short Need for Cognition Scale (sNCS; Cacioppo et al., 1984). This sNCS scale was included as part of an exploratory set of analyses due to its historic relevance to phenomena of social cognition (Petty et al., 2008; Wolf et al., 2017).

Data analysis

All data analyses and plots were produced in *R* (R Core Team, 2020). First, trials without a response were removed, and four participants who scored below chance on the catch-trials (< 8 out of 15) were removed. Secondly, trials with a reaction time of <= 100ms were filtered out of the dataset; this step removed less than one percent of the data remaining after initial exclusions. One participant had fewer than 360 experimental trials before any filtering, indicating that the computer had crashed and exited the experiment early. Following filtering, however, this participant had a roughly equal number of trials for each condition as the other participants, and therefore they were kept in the sample. The minimum percentage of trials completed by any participant was 80%; 36 participants completed over 99% of trials. Data were then split into the three respective outcome measures (extraversion [positive control], health and body size) to be processed and analysed separately.

Mean Likert scale responses were computed per-participant for both priming conditions (prime and neutral), first averaging across body size and identity. For the purposes of analysis, participants identified as ± 2.5 SD from the group mean in either priming condition were excluded in accordance with our preregistered analysis pipeline. Those with difference scores ± 2.5 SD from the group mean difference (prime - neutral) were kept in our main analyses, however all analyses were repeated with them excluded to provide alternative effect size estimates (see Supplementary Tables 1-6.). Shapiro-Wilk statistics were also calculated to highlight cases in which these extreme scores introduced skewness to cells of our analyses, and therefore indicate where the alternative analysis may offer a more accurate effect size estimate.

We report one-tailed t-tests as our main confirmatory hypothesis tests based on our directional predictions. We do not use inferential statistics to assess any other hypotheses, such as effects in the opposite direction to that predicted, however we include descriptive statistics and exploratory analyses, which would highlight any additional or unexpected patterns in the data (McBee & Field, 2017). Such exploratory analyses and freely available raw data can add value by serving to motivate hypothesis-testing strategies in future research (Scheel et al., 2021; Tong, 2019).

3.3 Results

3.3.1 Extraversion Ratings (positive control)

Mean extraversion ratings for the prime and neutral conditions were compared with a one-tailed paired samples t-test to establish the effectiveness of our priming manipulation in increasing participant's judgements of target's levels of extraversion. Mean extraversion ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 5.).

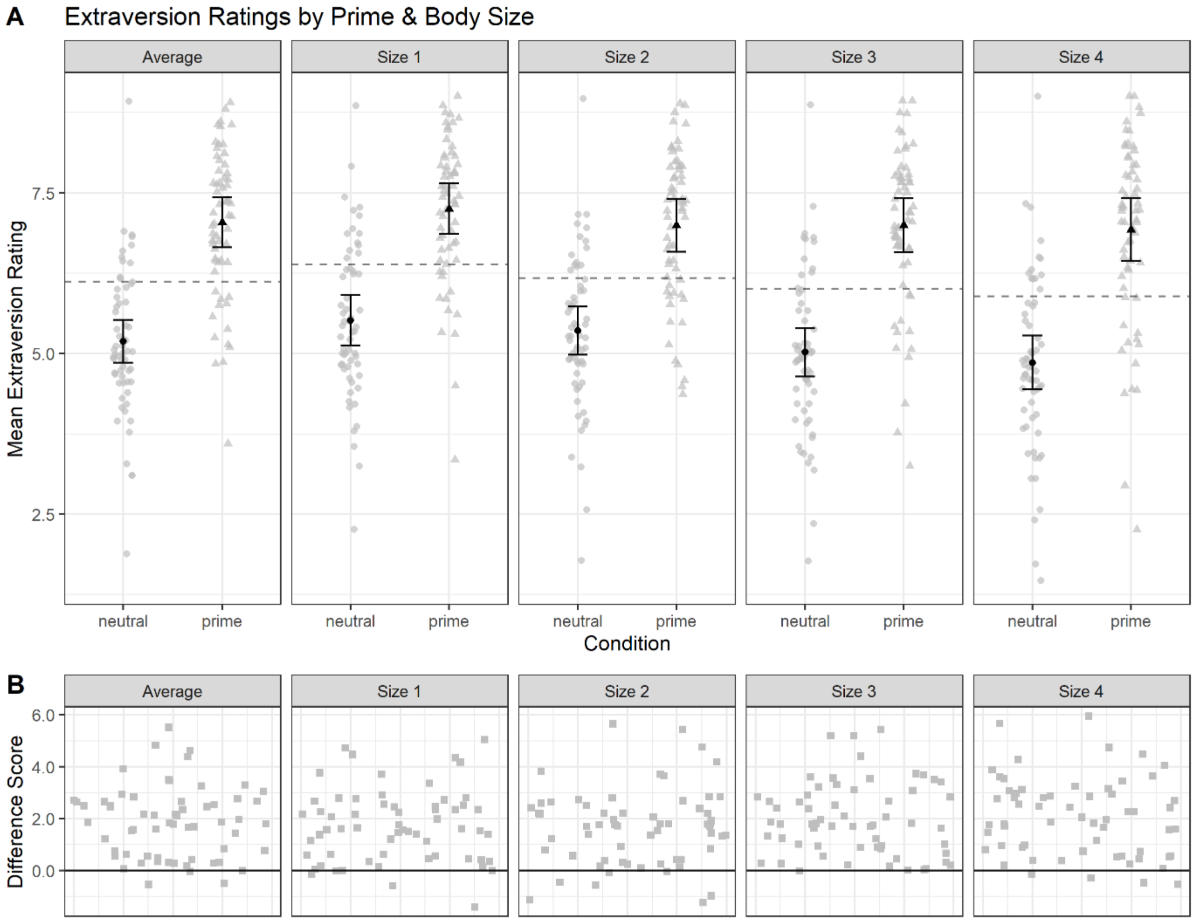

Figure 5. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

As expected, a paired samples t-test indicated a clear effect of extraversion primes on subsequent extraversion judgements, $t(56)=10.76$, $p<.001$, $d=1.41$ [1.10, ∞]. Mean difference = 1.85 [1.57, ∞]. The average difference score across participants and body sizes approached two points on the scale and was consistent in terms of direction with nearly all participants between zero and 4 points of difference.

3.3.2 Health Ratings

Mean health ratings for the prime and neutral conditions were compared with a one-tailed paired samples t-test. Mean health ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 6.).

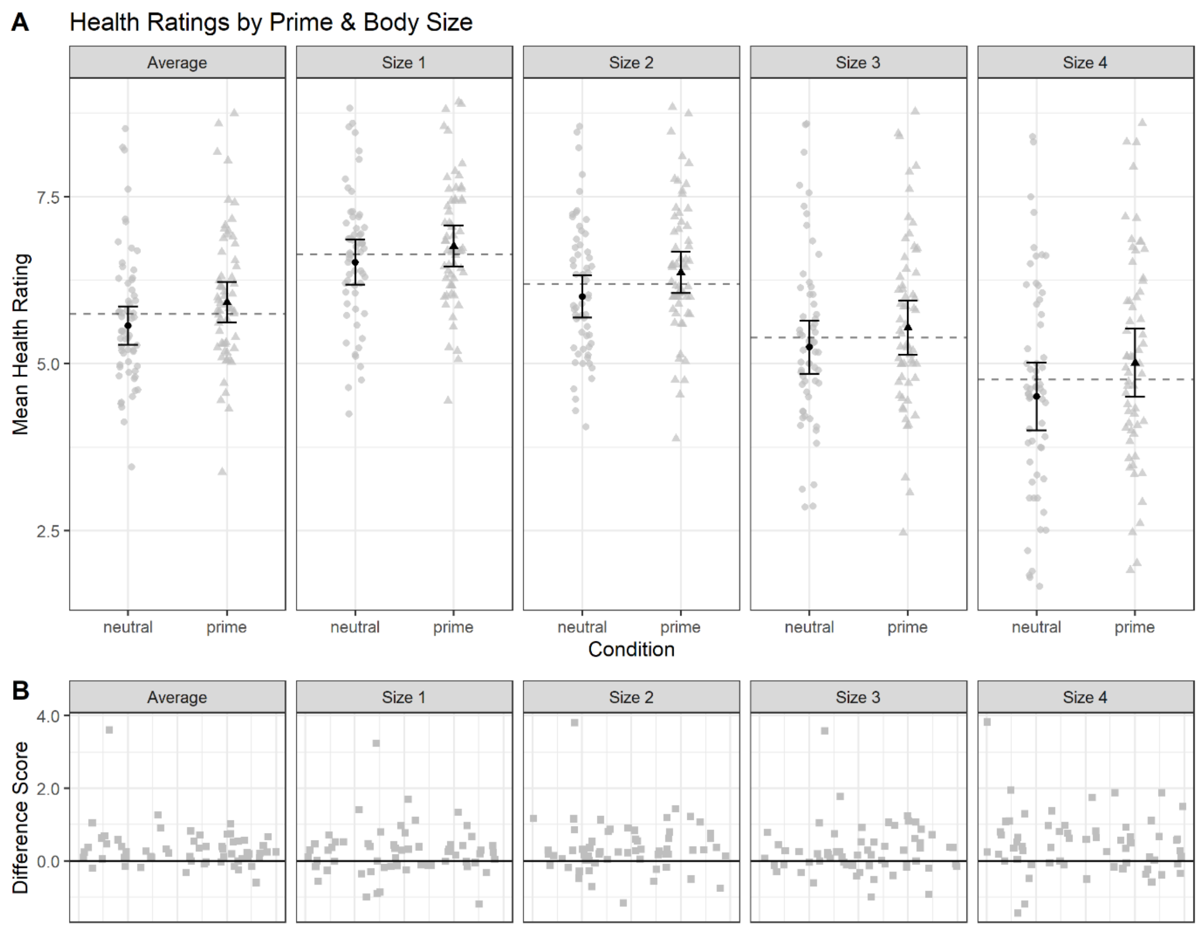

Figure 6. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

In line with our prediction, a paired samples t-test indicated that extraversion primes influenced subsequent judgements of health in the expected direction, $t(56)=4.61$, $p<.001$, $d=0.61$ [0.37, ∞]. Mean difference = 0.35 [0.22, ∞]. The average difference across participants and body sizes was approximately a third of a point on the scale and it was relatively consistent in terms of direction with most participants above zero but below 1 point of difference.

3.3.3 Body Size Ratings

Mean body size ratings for the prime and neutral conditions were compared with a one-tailed paired samples t-test. Mean size ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 7.).

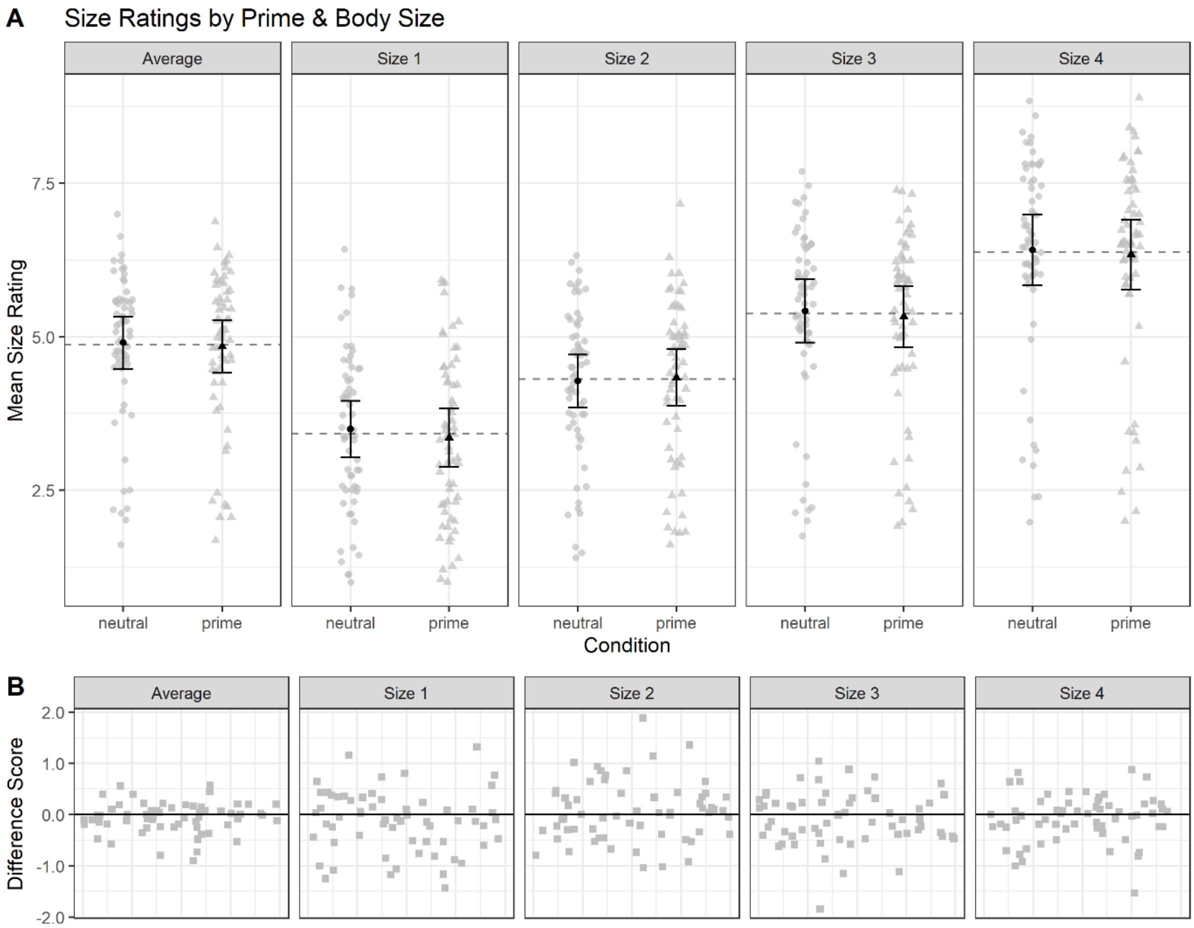

Figure 7. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

A paired samples t-test failed to give clear statistical support to our hypothesis that extraversion primes would decrease subsequent judgements of body size, $t(59)$=-1.66, p=.052, d=0.21 [-∞, 0.002]. Mean difference = -0.06 [∞, 0.0006]. However, the results of the test were in the expected direction, but the effect size was smaller than our design was powered to detect within the pre-determined confidence level. The average difference score across participants and body sizes was small (less than 0.1 point on the scale, Cohen's d =

0.21) and the direction of effect was variable around zero, with some participants showing a small positive effect (which was opposite to the direction that we predicted).

3.3.4 Exploratory Analyses

No associations were found between Need for Cognition and mean difference of ratings (prime – neutral) for any of our dependent measures (see Supplementary Table 7 and Supplementary Figures 2-5). Two sets of Cronbach's Alphas were calculated per dependent measure to test both inter-item consistency and inter-rater agreement. These show moderate inter-item consistency and high inter-rater agreement (see Supplementary Table 9). All data are made available for the pursuit of alternate exploratory hypotheses (https://osf.io/z9ds8/).

3.4 Discussion

The results from Experiment 2 provided clear evidence for the predicted effect of extraversion-diagnostic information on judgements of extraversion and health. Therefore, we are confident that the extraversion prime was working as expected and that priming extraversion generalises to person inferences associated with health. However, there was not the same level of support for judgments of body-size, although the effect was in the expected direction. Given the small effect on body size judgments (*d*=0.21) and recent widespread suggestions to increase rigour and credibility in psychological science (Munafò et al., 2017; Ramsey, 2020; Simmons et al., 2011, 2018; Vazire, 2018), we decided to replicate the procedure with a more sensitive dependent measure and a larger sample size.

4. Experiment 3

4.1 Introduction

Experiment 3 served to replicate experiment 2 and confirm the presence and magnitude of the observed effects. Given the small effect on our body-size dependent measure, we decided to approximately double our sample size for Experiment 3. In addition to this, given that the mean difference for both of our dependent measures was within a single point of the likert scale used, we increased the sensitivity of our dependent variable measure by using a 0-100 visual analogue scale (VAS). Finally, the sNCS was removed, and instead a questionnaire measuring big five personality dimensions was included (Big Five Aspect Scales [BFAS]; DeYoung et al., 2007). The inclusion of the BFAS was used as an exploratory measure to test whether difference scores for our dependent variables were associated with trait dimensions in conventional personality space.

All hypotheses, procedures, materials and data analysis protocols were otherwise identical to Experiment 2, and the experimental details were pre-registered in the same manner also (https://aspredicted.org/blind.php?x=ck3zf9). Our preregistered stopping rule for this experiment was defined as the point at which we had 110 useable participant datasets. A sensitivity analysis in G*Power indicated that N=110 would give us 80% power to detect an effect of $d = 0.23$, slightly higher than our computed effect size in Experiment 2, but feasible considering the resources available.

4.2 Method

4.2.1 Participants

One-hundred-and-twenty-three Bangor University students were recruited through Bangor University's student participation panel in exchange for course credit (22 males, 1 not specified, M_{age}=20.9, SD_{age}=4.16). Following data pre-processing and outlier removal, final sample sizes for each dependent variable were 109 for Extraversion, 108 for Health and 106 for Body Size.

4.2.2 Visual Analogue Scale

Our replication used a VAS in place of the likert scale used in Experiment 2. During the response phase of a trial participants chose a position on this scale by moving the mouse left and right, before clicking to record the response. This was then stored as a number from 0-100, but participants could not see the number itself (see Figure 8.).

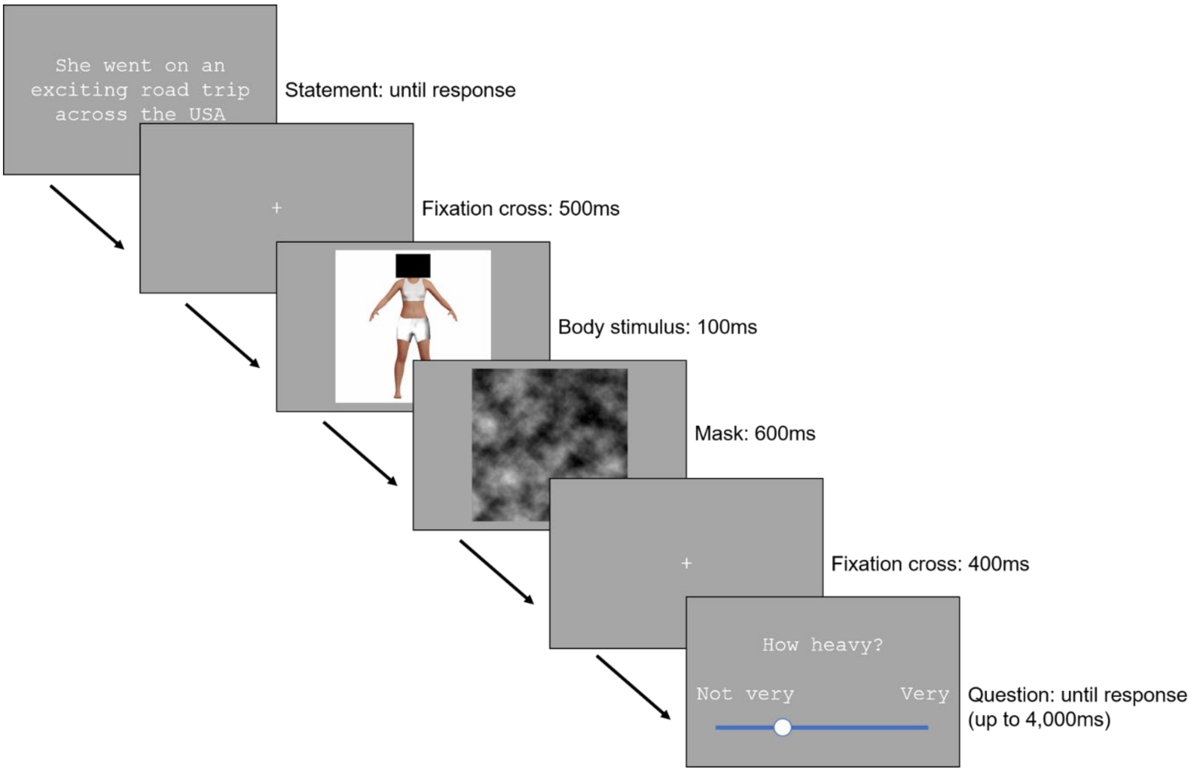

Figure 8. Example of experimental trial with VAS

4.2.3 Data analysis

Data pre-processing protocols were identical to those used in Experiment 2. Less than half a percent of data was discarded based on the reaction time threshold of 100ms. The minimum percentage of trials completed by any participant was 77%; 61 participants completed over 99% of trials.

4.3 Results

4.3.1 Extraversion Ratings (positive control)

Mean extraversion ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 9.).

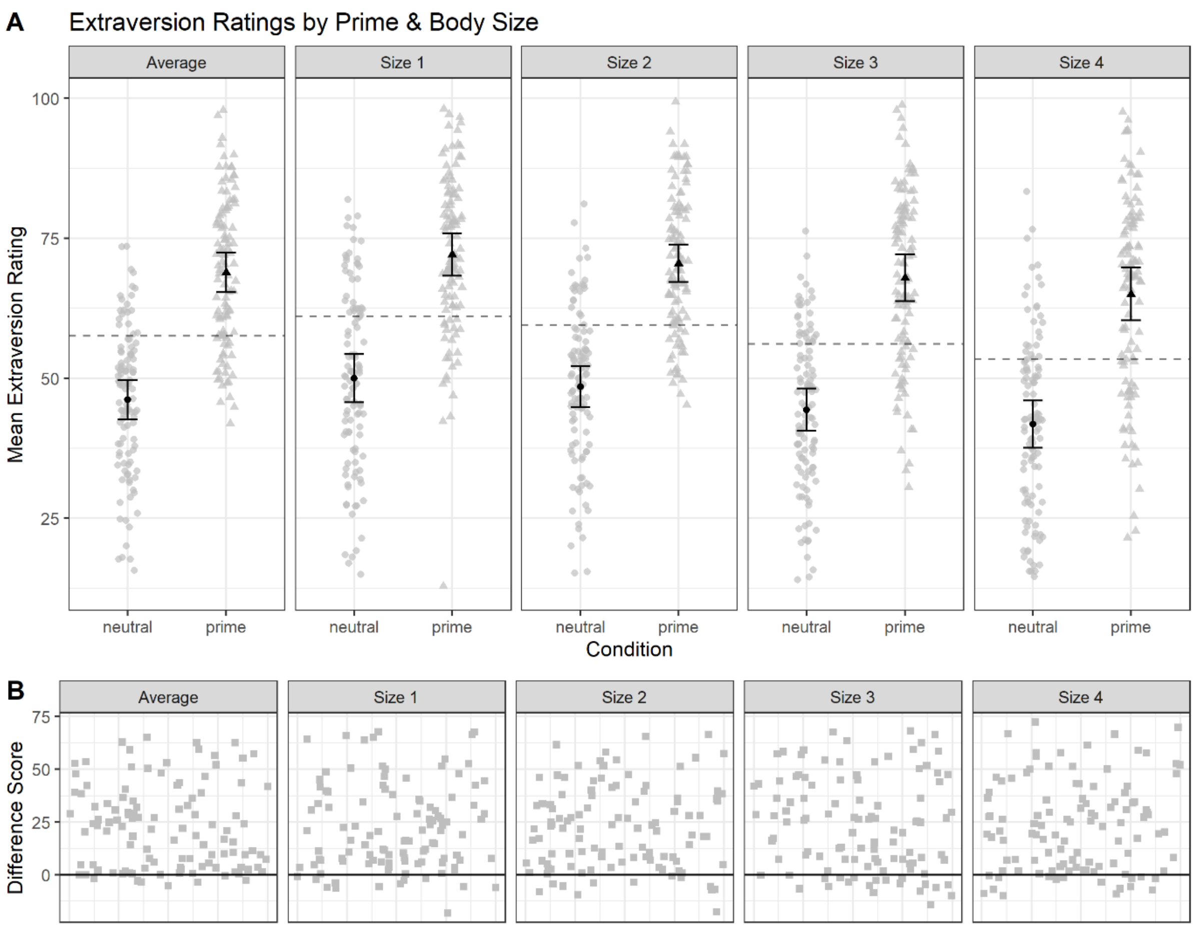

Figure 9. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

A paired samples t-test indicated a clear effect of extraversion primes on subsequent extraversion judgements in the expected direction, $t(108)=12.01$, $p<.001$, $d=1.15$ [0.95, ∞]. Mean difference = 22.73 [19.59, ∞]. The mean difference on the VAS was over 20 with nearly every participant above zero and many participants ranging up to 60 points in difference.

4.3.2 Health Ratings

Mean health ratings for the prime and neutral conditions were compared with a one-tailed paired samples t-test. Mean health ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 10.).

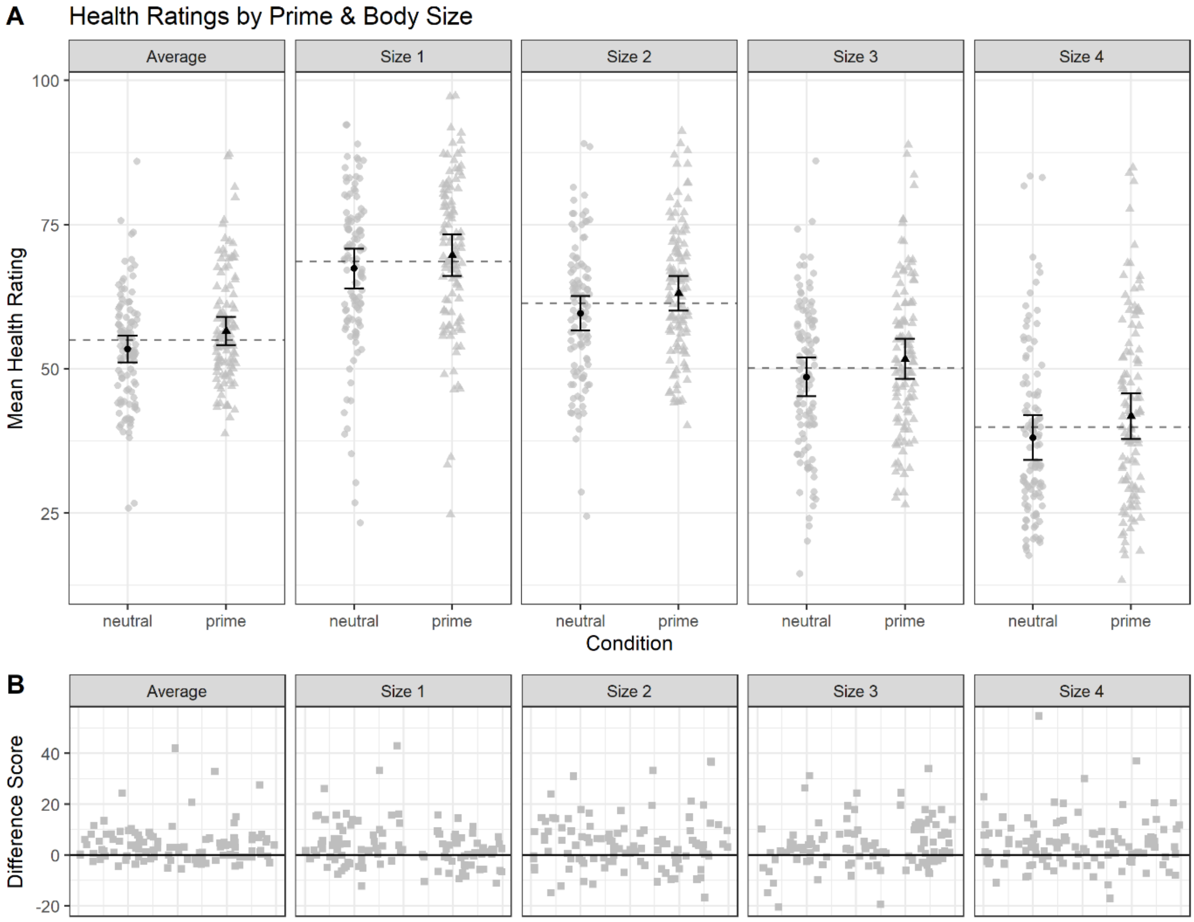

Figure 10. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

In line with our prediction, a paired samples t-test indicated extraversion primes influenced subsequent judgements of health in the expected direction, $t(107)=5.65$, $p<.001$, $d=0.54$ [0.37, ∞]. Mean difference = 3.980 [2.81, ∞]. The mean difference was approximately four points on the VAS with most participants between zero and 10 points. A number of participants did show a small negative difference, however.

4.3.3 Body Size Ratings

Mean body size ratings for the prime and neutral conditions were compared with a one-tailed paired samples t-test. Mean size ratings, both average and broken down by the four body sizes, are shown below alongside difference scores (prime - neutral) showing the distribution around zero (see Figure 11.).

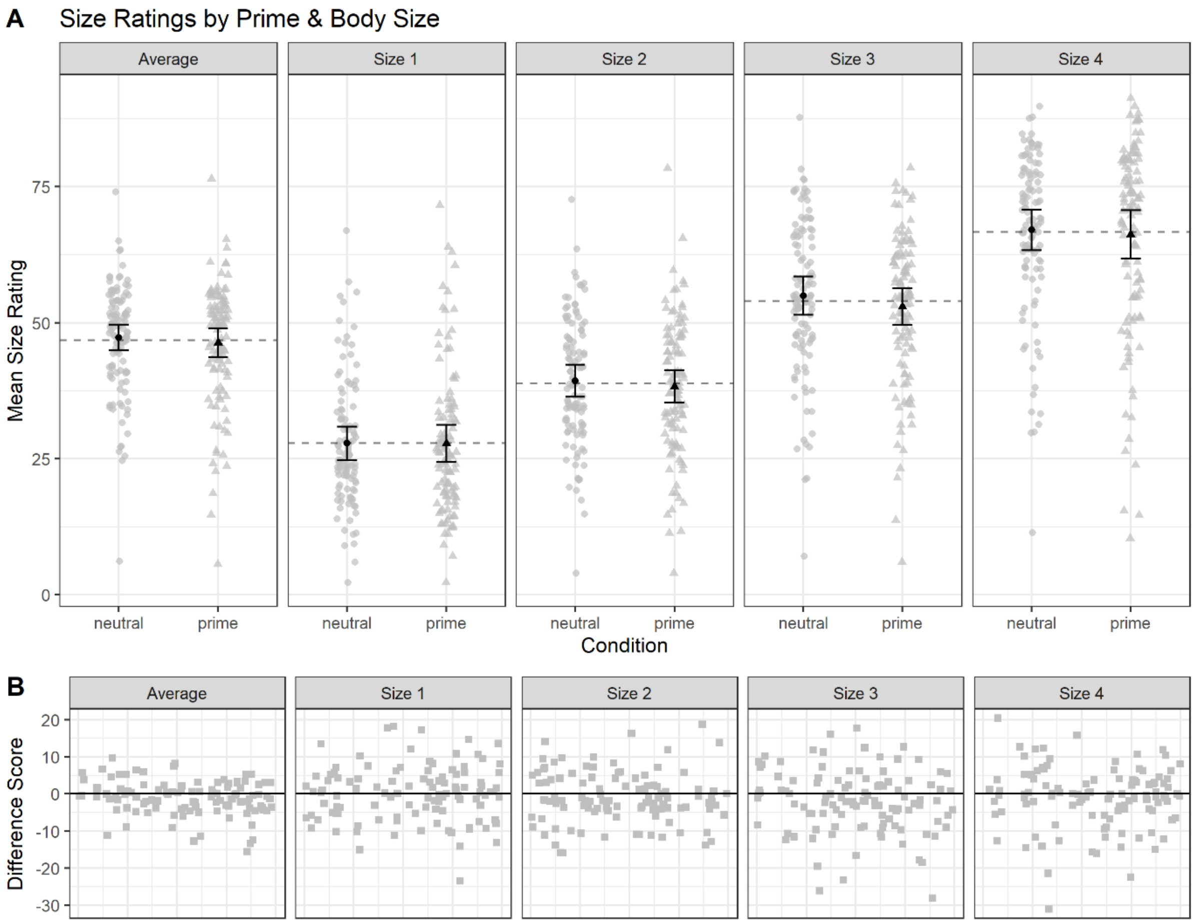

Figure 11. (A) Participants' mean ratings for each condition and aggregated mean scores. Error bars show 95% CIs. Dashed lines indicate the average of prime and neutral. (B) Participants' difference scores, showing distribution around zero (null).

A paired samples t-test revealed a difference in the same direction as Experiment 2 and the effect size was of a very similar magnitude, $t(105)=-2.20$, $p=.015$, $d=-0.214$ $[-\infty, -0.05]$. Mean difference = -0.969 $[-\infty, -0.24]$. Like Experiment 2, the average effect across the group is small and in the expected direction (1 point on the VAS, Cohen's d = -0.21). Many individual participants show an effect on or above zero, which demonstrates considerable variability across participants.

4.3.4 Exploratory Analyses

Pearson's correlations between the difference scores for each dependent variable and each sub-facet of the are reported in Supplementary Table 8 and visualised in Supplementary Figures 6-9. No clearly meaningful patterns of data emerged from these exploratory correlations. Two sets of Cronbach's Alphas were calculated per dependent measure to test both inter-item consistency and inter-rater agreement. These show mixed inter-item consistency and high inter-rater agreement (see Supplementary Table 9). The variability in inter-item consistency appears to be driven by differences in the scale values used to reflect participants' lowest and highest responses, which sometimes drive negative correlations between scores for the smallest and largest body sizes. Additionally, a series of 2x4 factorial ANOVAs were carried out to evaluate potential interactions between our priming effects and target body sizes for Experiments 2 and 3. These analyses did not support the presence of an interaction between prime condition and body size (see Supplementary Analyses) All data are made available for the pursuit of alternate exploratory hypotheses (https://osf.io/z9ds8/).

4.3.5 Meta-analysis across experiments

A meta-analysis of the effect sizes measured in Experiments 2 and 3 was performed to calculate a pooled effect size. This was conducted in ESCI using unbiased estimates of population effect sizes (Cumming, 2013) (see Figure 12.). Exact values for all statistics presented in the article can be found in Supplementary Tables 2-6.

Figure 12. Summary of effect sizes and pooled effect size estimates

4.4 Discussion

The findings from Experiment 3, as well as the meta-analysis, confirm our prior findings. First, the extraversion and health judgments showed a clear and consistent increase between neutral and prime trials, with the impact on extraversion judgments being approximately twice as large as the impact on health judgments. As such, the effect on extraversion judgments served as a useful 'positive control' and manipulation check by showing that the extraversion prime was operating on person judgments in a manner that we intended. By contrast, the effect of extraversion-diagnostic primes on health judgments demonstrates the generalisability of this trait inference to person inferences that extend beyond the initial personality construct.

A second finding that Experiment 3 replicates, but with greater precision in the estimated effect size, is that there is a small negative effect of trait knowledge on body size judgments. The impact on body size judgments operates in a predictable direction on average across participants. The effect also varies between individuals with some not showing the effect (i.e., some participants show an effect close to zero or a small positive effect). Therefore, the effect of trait knowledge on social perception may manifest as an individual difference, whereby only a subset of the population shows the effect. Alternatively, the lack of consistent effect across participants may reflect limits to the sensitivity of the perceptual measure, which future research would have to establish by developing different measures that exhibit greater sensitivity.

5. General Discussion

We show that making a trait inference about a person generalises sufficiently to influence other similar person inferences (health), as well as distinctly different judgments that rely more heavily on visual person representations (body size). These results deepen our understanding of the relationship between 'trait space' and 'body space' when forming impressions of other people by providing a behavioural characterisation of the function of the interplay between distinct pieces of person knowledge. In contrast to much prior work, which focussed on person inferences prompted by face or body images, here we show that a brief trait-inference can bias judgements that are based more on perceptual representations of body shape. Therefore, a relatively transient person inference can provide a small change in the way one "sees" other people. We suggest that the primary value of this work is that it underscores why it is important to link neuroscience research with behavioural research (Krakauer et al., 2017). Indeed, by providing a relevant functional description of the links between trait knowledge and perceptual processes, we aid the interpretation of prior neuroimaging studies, which used functional connectivity measures and showed links between body-part processing and theory of mind networks (Ramsey, 2018).

5.1 Implications

Our findings deepen understanding of the mapping between 'trait space' and 'face/body space' when forming impressions of other people (Over & Cook, 2018). We show that reconfiguring trait space via a person inference, subsequently alters links between other person inferences, as well as judgments that rely on a distinct system that is sensitive to body size judgments. As such, we show that trait-based knowledge does not just influence mappings towards similar types of person judgments, such as health judgments. Rather, re-configuring trait-space alters mappings towards non-trait judgments, which are based on body size and shape. The strength of re-mapping is not the same in all cases, however. Modifying trait space has a much stronger influence on similar rather than dissimilar judgment types. This suggests that the mapping of within trait space is stricter than between trait space and body space, as one may intuitively expect. Taken together with prior work, which showed that facial or body features can prompt trait inferences (Greven et al., 2019; Todorov et al., 2015), we suggest that judgments of body size and person inferences are reciprocally linked and mutually reinforce each other.

The results complement prior neuroimaging findings that showed functional interactions between the body-selective brain regions in the ventral visual stream and the theory of mind network when forming impressions of others (Ramsey, 2018). During impression formation, distinct information processing units do not operate in isolation; instead, they exchange signals to integrate information and bias the overall judgment space. We feel that the general approach taken here, as well as in previous papers (Greven et al., 2016; Over & Cook, 2018; Ramsey, 2018), underscores the value of considering the integration of different signals when forming an overall impression, rather than the modal approach in social cognition that studies perceptual and inferential processes separately. Furthermore, we believe that the use of behavioural research aids in characterising the functional qualities of integration between these neural networks, where prior neuroscience alone has focussed more on establishing the presence of such links. Our findings, therefore, add to recent proposals that highlight how considering behavioural and neural data sets together can help adjudicate between competing mechanistic models and place useful constraints on mechanism discovery in the human brain (Kaplan & Hewitson, 2020;

Krakauer et al., 2017; Niv, 2020). We hope that links between sub-disciplines of social cognition and neuroscience will continue to emerge, because a piecemeal approach to understanding any aspect of cognitive and brain function is limited (Churchland, 2013). Whether one typically focusses on inferences common in theory of mind research (e.g. Frith & Frith, 1999; Saxe & Kanwisher, 2003; van Overwalle, 2009), or the sensitivity within the visual system to features of another person (e.g. Kanwisher, 2010), we feel that both endeavours work better when they are considered together, and not only separately. This suggestion appears especially relevant when one considers a typical social exchange, which requires one to fuse together physical features of a person with knowledge about their trait-based character.

5.2 Strengths and limitations

One possible limitation of the results concerns the general difficulties that are associated with interpreting small effect sizes. The effects of trait inferences on body-size judgments were small (approximately Cohen's d = 0.2) and many participants did not show an effect in the predicted direction. One could argue, therefore, that it is difficult to interpret such findings because they are more likely to reflect sampling error and chance variation. We point to several factors within our approach that make sampling error an unlikely explanation of our findings. First, all of our predicted effects, which comprised three separate dependent variables per priming experiment, were in the expected direction and consistent across both priming experiments, including the high-power replication experiment. Second, given that the effects on dependent variables were not all in the same direction, it is unlikely that the body-size effect can be accounted for by an artefact of the experimental paradigm or a simple response rule (i.e. always responding higher on prime trials). Indeed, any explanation of our findings needs to account for why the same prime systematically biases different judgments in different ways. Moreover, it is one reason why we included a 'positive control' condition to serve as a manipulation check – a condition where we have good reason to expect a particular effect, which can guide the interpretation of other results. Finally, as expected, within-modality priming effects were considerably larger than cross-modal priming effects, which should also be expected from prior priming and adaptation studies (Burton et al., 1990; Hills et al., 2010; Watson et al., 2014). As such, although small effects, we feel that they have been precisely estimated and a cautious interpretation is therefore justified.

Other factors also provide relevant context when interpreting small effects. First, the aims of the present research matter. We were concerned with performing basic science research that tested a model system of the structure of social cognition and person perception. We were not aiming to provide results that served an immediate practical benefit. As such, we feel that small but relatively precise estimates of effect size license a judgment about the target systems of interest. A second aspect of relevant context is the potential for effects to aggregate over time (Funder & Ozer, 2019). That is, in a one-off trial or over the course of an experiment, any given small effect may be inconsequential in practical terms. But, in real life, if that effect – say making a trait-inference about a colleague at work – happens 20 times a day, five days a week, then the effects may become cumulative and be stronger than the current experiment can demonstrate. Of course, this is an empirical question, which would need demonstrating using a different research design, but we nonetheless feel that it provides an important consideration when interpreting effect sizes.

Earlier we noted the possibility of our trait-neutral statements leading to judgements of high introversion rather than being truly trait-neutral. Indeed, it is likely that some trait-diagnostic information can be extracted from an ostensibly neutral statement, particularly in the context of a task that demands some form of social evaluation. This is an important consideration for the interpretation of our effect sizes because, assuming a wholly linear relationship between body size and the introverted-extraverted axis, one would expect the true difference between extraverted and neutral judgements to be approximately half the size of the true difference between extraverted and introverted judgments. We address this possibility in the service of informing future work using similar paradigms, and in the interests of accurately characterising our effect size estimates: Each dataset that involved a judgement of extraversion (both priming experiments and the survey validating our trait statements), reflects a task in which participants judged the extent to which the person in question was extraverted, so it is difficult to interpret whether responses in the lower half of the scale reflected a judgement of 'neutral' or 'introverted'. However, in the 'positive control' condition of each priming experiment, the average extraversion rating was close to or above the centre of the response scale, making it unlikely that participants were judging the target to be highly introverted. Furthermore, when judging extraversion, variability in participants' responses was far better explained by priming condition than the size of the associated body, further supporting that this rating reflected their interpretation of the prime more so than the body. Therefore, we argue that it is unlikely that we've greatly overestimated the difference between extraverted and trait-neutral character in influencing judgements, however acknowledge and highlight that all 'neutral' statements are likely to possess some trait-diagnostic information which may influence judgements in these types of experiments.

Finally, it is important to address the extent to which our design can support specific claims about cross-modal influences on perceptual processes. We recognise that our findings could reflect a general "halo effect", where trait characteristics generally deemed positive lead people to judge other aspects of a person in a way that is culturally and/or subjectively favoured (i.e., thin-ideal). This is difficult to fully disentangle from perceptual processes, as we would expect judgements of body size to be biased whether individuals are forming a body size judgement in real-time, from memory, or even about an imagined body. This is because generating a judgement of someone's size is likely to require body shape to be internally represented in some way, and trait inferences would be expected to influence judgements which are then based off this body shape representation. As such, it is possible that the effect arises from the perceptual component during encoding, the memory component during recall, or a combination of the two processes. We note that the designs used in the current experiments are unable to clearly separate the role of perceptual versus memory processes, and we suggest that a valuable future direction would be to probe this question. What we can conclude from this series of experiments, however, is that when participants formed a single judgement about a target identity, whether based on extraversion, health or body size, this judgement reflected the influence of both the visual percept and imbued trait character of that target identity. That is, when averaging across prime condition, all types of judgements vary as a function of body size, and when averaging across body size, all types of judgement vary as a function of prime condition (see Supplementary Analyses).

5.3 Constraints on Generality

In terms of our theoretical interpretation of findings, we acknowledge that the present work says nothing about the accuracy of links between trait inferences and body shape representations (that is, the extent to which they reflect true correlations between traits and body shapes in the real world). We therefore remain largely agnostic to the possible functional benefits of these inferences to guide social interactions or predict how someone is likely to behave, as the ways in which character judgements are linked to physical appearance are often found to reflect idiosyncratic and culturally-acquired stereotypes. Whether these stereotypes serve an adaptive function despite being largely inaccurate (e.g., heuristics for anticipating the maximal bounds of probable behaviour), or reflect a once functional system now biased by a heavily skewed 'perceptual diet', is a separate empirical question which remains untested by the current study.

Our findings were demonstrated using computer-generated female bodies in a sample of students, where extraversion-diagnostic information was delivered through behavioural statements. Given our use of computer-generated bodies we do not expect the experimental task to have fully tapped body perception processes, nor do we expect the effects to translate 1:1 to an analogous real-world context given the constrained presentation of bodies without wider context such as the face. Rather, we argue that the presence of such an effect in a tightly controlled lab environment is indicative of the manner in which separate systems integrate information. While we would expect this integration to have real-world consequences, the precise nature and outcomes of this is likely to vary based on numerous factors including the context of other morphological characteristics of the target body, and individual differences in the structure of conceptual trait space (e.g. Stolier et al., 2018). It is also important to highlight that the current set of experiments explicitly required participants to form judgements about the stimuli, so it is unclear whether the integration demonstrated here occurs spontaneously or only in the context of explicitly forming judgements. Lastly, the specific dimensions selected for the current study, extraversion, health and body fat, may represent a special case for such an effect to occur given strong evidence for their alignment in judgements of bodies (Greven et al., 2019; Hu et al., 2018). Broader investigation would be required to establish evidence of a general pattern of integration where various dimensions of social evaluation influence various dimensions of size and shape.

Although the current work represents basic research that aimed to understand a model system of cognitive function, in the longer term, understanding the complex underpinnings of impression formation may have applied relevance. For instance, such work may provide insight into the mechanisms that support body-size-based stigma. If simply reading statements about other individuals under sanitised and socially impoverished laboratory conditions can bias estimates of observed body size, it may be no surprise that social media, advertising and healthy lifestyle messaging can be a powerful reinforcer of such stigma. A further future consideration for applied research is the relationship between perceptual and inferential processes when understanding distortions in judgments of one's own body. For example, body size distortion is a key feature of Anorexia Nervosa (Zopf et al., 2016), and young people who self-harm also have an altered body representation (Hielscher et al., 2019). Therefore, it is not difficult to see how a deeper understanding of the complex and multi-faceted bases of body image representations may ultimately have applied relevance.

References

Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*, *60*, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Alexi, J., Dommisse, K., Cleary, D., Palermo, R., Kloth, N., & Bell, J. (2019). An Assessment of Computer-Generated Stimuli for Use in Studies of Body Size Estimation and Bias. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02390

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256. https://doi.org/10.1037/0033-2909.111.2.256

Anderson, N. H. (1962). Application of an Additive Model to Impression Formation. *Science*, *138*(3542), 817–818. https://doi.org/10.1126/science.138.3542.817

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science (New York, N.Y.)*, *338*(6111), 1225–1229. https://doi.org/10.1126/science.1224313

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience*, *10*(3), 186–198. https://doi.org/10.1038/nrn2575

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81*(3), 361–380. https://doi.org/10.1111/j.2044-8295.1990.tb02367.x

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Churchland, P. M. (2013). *Matter and Consciousness*. MIT Press.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

Cornelissen, K. K., Gledhill, L. J., Cornelissen, P. L., & Tovée, M. J. (2016). Visual biases in judging body weight. *British Journal of Health Psychology*, *21*(3), 555–569. https://doi.org/10.1111/bjhp.12185

Cumming, G. (2013). *The new statistics: Estimation for better research*. www.thenewstatistics.com

Daly, M., Sutin, A. R., & Robinson, E. (2019). Perceived Weight Discrimination Mediates the Prospective Association Between Obesity and Physiological Dysregulation: Evidence From a Population-Based Cohort. *Psychological Science*, *30*(7), 1030–1039. https://doi.org/10.1177/0956797619849440

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242–249. https://doi.org/10.1038/nrn1872

de Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., & Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neuroscience & Biobehavioral Reviews*, *34*(4), 513–527. https://doi.org/10.1016/j.neubiorev.2009.10.008

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science (New York, N.Y.)*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Downing, P. E., & Peelen, M. V. (2011). How might occipitotemporal body-selective regions interact with other brain areas to support person perception? *Cognitive Neuroscience*, *2*(3–4), 216–226. https://doi.org/10.1080/17588928.2011.613987

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

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science (New York, N.Y.)*, *286*(5445), 1692–1695. https://doi.org/10.1126/science.286.5445.1692

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313. https://doi.org/10.1146/annurev-psych-120710-100449

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Greven, I. M., Downing, P. E., & Ramsey, R. (2016). Linking person perception and person knowledge in the human brain. *Social Cognitive and Affective Neuroscience*, *11*(4), 641–651. https://doi.org/10.1093/scan/nsv148

Greven, I. M., Downing, P. E., & Ramsey, R. (2019). Neural networks supporting social evaluation of bodies based on body shape. *Social Neuroscience*, *14*(3), 328–344. https://doi.org/10.1080/17470919.2018.1448888

Greven, I. M., & Ramsey, R. (2017a). Person perception involves functional integration between the extrastriate body area and temporal pole. *Neuropsychologia*, *96*, 52–60. https://doi.org/10.1016/j.neuropsychologia.2017.01.003

Greven, I. M., & Ramsey, R. (2017b). Neural network integration during the perception of in-group and out-group members. *Neuropsychologia*, *106*, 225–235. https://doi.org/10.1016/j.neuropsychologia.2017.09.036

Hendrick, C., Franz, C. M., & Hoving, K. L. (1975). How do children form impressions of persons? They average. *Memory & Cognition*, *3*(3), 325–328. https://doi.org/10.3758/BF03212919

Hielscher, E., Whitford, T. J., Scott, J. G., & Zopf, R. (2019). When the body is the target-Representations of one's own body and bodily sensations in self-harm: A systematic review. *Neuroscience and Biobehavioral Reviews*, *101*, 85–112. https://doi.org/10.1016/j.neubiorev.2019.03.007

Hills, P. J., Elward, R. L., & Lewis, M. B. (2010). Cross-modal face identity aftereffects and their relation to priming. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 876. https://doi.org/10.1037/a0018731

Hill, M. Q., Streuber, S., Hahn, C. A., Black, M. J., & O'Toole, A. J. (2016). Creating body shapes from verbal descriptions by linking similarity spaces. *Psychological Science*, *27*(11), 1486-1497. https://doi.org/10.1177/0956797616663878

Hu, Y., Parde, C. J., Hill, M. Q., Mahmood, N., & O'Toole, A. J. (2018). First Impressions of Personality Traits From Body Shapes. *Psychological Science*, *29*(12), 1969–1983. https://doi.org/10.1177/0956797618799300

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research, 2nd ed* (pp. 102–138). Guilford Press.

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, *107*(25), 11163–11170. https://doi.org/10.1073/pnas.1005062107

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kaplan, D. M. & Hewitson, C. L. (2020). Modelling Bayesian Computation in the Brain: Unification, Explanation, and Constraints. In F. Calzavarini & M. Viola, (Eds.). *Neural Mechanisms: New Challenges in Philosophy of Neuroscience.* Springer International Publishing.

Kemmerer, D. (2011). Do body-part concepts depend on the EBA/FBA? *Cognitive Neuroscience*, *2*(3–4), 204–205. https://doi.org/10.1080/17588928.2011.604718

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

McBee, M. T., & Field, S. H. (2017). Confirmatory study design, data analysis, and results that matter. In M. C. Makel & J. A. Plucker (Eds.), *Toward a more perfect psychology: Improving trust, accuracy, and transparency in research* (pp. 59–78). American Psychological Association. https://doi.org/10.1037/0000033-004i

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1309–1316. https://doi.org/10.1098/rstb.2008.0318

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, *28*(4), 757–762. https://doi.org/10.1016/j.neuroimage.2005.03.011

Mitchell, J. P., Cloutier, J., Banaji, M. R., & Macrae, C. N. (2006). Medial prefrontal dissociations during processing of trait diagnostic and nondiagnostic person information. *Social Cognitive and Affective Neuroscience*, *1*(1), 49–55. https://doi.org/10.1093/scan/nsl007

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Naumann, L., Vazire, S., Rentfrow, P., & Gosling, S. (2009). Personality Judgments Based on Physical Appearance. *Personality & Social Psychology Bulletin*, *35*, 1661–1671. https://doi.org/10.1177/0146167209346309

Niv, Y., (2020) The primacy of behavioral research for understanding the brain. PsyArXiv, https://psyarxiv.com/y8mxe/

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

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

Park, H.-J., & Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science*, *342*(6158). https://doi.org/10.1126/science.1238411

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*(8), 636–648. https://doi.org/10.1038/nrn2195

Petty, R. E., DeMarree, K. G., Briñol, P., Horcajo, J., & Strathman, A. J. (2008). Need for cognition can magnify or attenuate priming effects in social judgment. *Personality & Social Psychology Bulletin*, *34*(7), 900–912. https://doi.org/10.1177/0146167208316692

Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity (Silver Spring, Md.)*, *17*(5), 941–964. https://doi.org/10.1038/oby.2008.636

Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., & Macrae, C. N. (2011). Stereotype-based modulation of person perception. *NeuroImage*, *57*(2), 549–557. https://doi.org/10.1016/j.neuroimage.2011.05.004

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at https://www.R-project.org/

Ramsey, R. (2018). Neural Integration in Body Perception. *Journal of Cognitive Neuroscience*, *30*(10), 1442–1451. https://doi.org/10.1162/jocn_a_01299

Ramsey, R. (2020). Advocating for the credibility revolution. *Cognitive Psychology Bulletin*, *5*. https://doi.org/10.31234/osf.io/3kwnu

Ramsey, R., Schie, H. T. van, & Cross, E. S. (2011). No two are the same: Body shape is part of identifying others. *Cognitive Neuroscience*, *2*(3–4), 207–208. https://doi.org/10.1080/17588928.2011.604721

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science*, *16*(4), 744-755 https://doi.org/10.1177/1745691620966795

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(47), 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005

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

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2018). False-Positive Citations. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *13*(2), 255–259. https://doi.org/10.1177/1745691617698146

Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018). The conceptual structure of face impressions. *Proceedings of the National Academy of Sciences*, *115*(37), 9210–9215. https://doi.org/10.1073/pnas.1807222115

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

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*, 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Tong, C. (2019). Statistical Inference Enables Bad Science; Statistical Thinking Enables Good Science. *The American Statistician, 73*(sup1), 246–261. https://doi.org/10.1080/00031305.2018.1518264

van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress: *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691617751884

Wang, Y. C., McPherson, K., Marsh, T., Gortmaker, S. L., & Brown, M. (2011). Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet (London, England)*, *378*(9793), 815–825. https://doi.org/10.1016/S0140-6736(11)60814-3

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2014). Crossmodal Adaptation in Right Posterior Superior Temporal Sulcus during Face–Voice Emotional Integration. *The Journal of Neuroscience*, *34*(20), 6813–6821. https://doi.org/10.1523/JNEUROSCI.4478-13.2014

Wolf, L. J., Hecker, U. von, & Maio, G. R. (2017). Affective and cognitive orientations in group perception. *Personality and Social Psychology Bulletin*, *43*(6), 828–844. https://doi.org/10.1177/0146167217699582

Zopf, R., Contini, E., Fowler, C., Mondraty, N., & Williams, M. A. (2016). Body distortions in Anorexia Nervosa: Evidence for changed processing of multisensory bodily signals. *Psychiatry Research*, *245*, 473–481. https://doi.org/10.1016/j.psychres.2016.09.003