

Highlights

- Information from torso edges, not the central body, drives self-estimates of body size in women
- Information extraction was independent of bubble size in the bubble masking task used
- Normal eye fixations up and down the central torso remained despite the bubble mask
- Eye movements and diagnostic regions for self-estimates of body size are not necessarily equivalent

Title: **The visual cues that drive the self-assessment of body size: dissociation between fixation patterns and the key areas of the body for accurate judgement**

Authors: Kamila R. Irvine^a, Kristofor McCarty^a, Thomas V. Pollet^a, Katri K. Cornelissen^a,
Martin J. Tovée^b, Piers L. Cornelissen^a

^aDepartment of Psychology, Faculty of Health & Life Sciences, Northumbria University,
Newcastle upon Tyne, NE1 8ST, UK

^bSchool of Psychology, College of Social Science, University of Lincoln, Lincolnshire, LN6
7TS, UK

CORRESPONDENCE TO:

Piers Cornelissen: piers.cornelissen@northumbria.ac.uk

Declarations of interest: none

60
61
62 **Abstract**
63
64

65 4 A modified version of the bubbles masking paradigm was used in three experiments to
66
67 5 determine the key areas of the body that are used in self-estimates of body size. In this
68
69 6 paradigm, parts of the stimuli are revealed by several randomly allocated Gaussian “windows”
70
71 7 forcing judgements to be made based on this partial information. Over multiple trials, all
72
73 8 potential cues are sampled, and the effectiveness of each window at predicting the judgement
74
75 9 is determined. The modified bubbles strategy emphasises the distinction between central versus
76
77 10 edge cues and localises the visual features used in judging one’s own body size. In addition,
78
79 11 eye-movements were measured in conjunction with the bubbles paradigm and the results
80
81 12 mapped onto a common reference space. This shows that although observers fixate centrally
82
83 13 on the torso, they are actually directing their visual attention to the edges of the torso to gauge
84
85 14 body width as an index of body size. The central fixations are simply the most efficient way of
86
87 15 positioning the eye to make this estimation. Inaccurate observers are less precise in their central
88
89 16 fixations and do not evenly allocate their attention to both sides of the torso’s edge, illustrating
90
91 17 the importance of efficiently sampling the key information.
92
93
94
95
96
97

98 19 **Key words:** BMI, self-estimates, body size estimation, eye-movements, bubbles masking
99
100 20 technique, visual cues.
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118

119
120
121 **1. Introduction**
122
123

124 22 It is well established that people who suffer from anorexia nervosa or bulimia nervosa
125
126 23 overestimate their own body size (e.g., Cornelissen, Johns, & Tovée 2013; Gardner &
127
128 24 Bokenkamp, 1996; Probst, Vandereycken, Van Coppenolle, & Pieters, 1998; Slade & Russell,
129
130 25 1973; Tovée, Benson, Emery, Mason, & Cohen-Tovée, 2003), although the magnitude of this
131
132 26 overestimation may also depend on a person's body mass index (BMI; Cornelissen et al., 2015,
133
134 27 2017). Body size overestimation is one of the most persistent of all the eating disorder
135
136 28 symptoms, the severity of which predicts the long-term outcome of treatment (Fairburn,
137
138 29 Cooper, & Shafran, 2003; Junne et al., 2019; Pike, 1998), and its persistence predicts the
139
140 30 likelihood of relapse, which occurs at high rates (Berkman, Lohr, & Bulik, 2007; Castro, Gila,
141
142 31 Puig, Rodriguez, & Toro, 2004; Channon & DeSilva, 1985; Herzog et al., 1999; Keel, Dorer,
143
144 32 Franko, Jackson, & Herzog, 2005; Slade & Russell, 1973). It is therefore important that self-
145
146 33 estimates of body size can be made accurately, that we understand how these judgements are
147
148 34 made, how they may go awry, and to develop techniques to ameliorate this.
149
150
151

152 35 Two factors contribute to the estimation of one's own body size, both of which can be
153
154 36 disturbed in eating disorders (Cash & Deagle, 1997): (1) an attitudinal component which
155
156 37 captures the feelings that a person has about their body's size and shape, and (2) a perceptual
157
158 38 component that has to do with the accuracy with which a person can judge the dimensions of
159
160 39 their own physical appearance. Although more recent reviews exist, e.g., Skrzypek, Wehmeier,
161
162 40 and Remschmidt (2001), they arrive at essentially the same conclusion. Measuring the
163
164 41 attitudinal component of body image has proved to be relatively straightforward. Typically,
165
166 42 psychometric tools are used to assess such attributes as body dissatisfaction and attitudes to
167
168 43 body shape and weight (Evans & Dolan, 1993; Fairburn & Beglin, 1994). However, measuring
169
170 44 the perceptual component of body size estimation has proved more challenging. A wide variety
171
172 45 of methods have been tried, starting from image marking procedures (Askevold, 1975) and
173
174
175
176
177

178
179
180 46 moveable calliper techniques (Slade & Russell, 1973) to distorting photograph and video
181
182 47 techniques (Gardner & Moncrieff, 1988; Probst, Vandereycken, & Van Coppenolle, 1995;
183
184 48 Shafran & Fairburn, 2002). Most recently CGI (computer generated imagery) technology has
185
186 49 been used to create standard stimuli or even personalized 3D avatars that accurately reflect
187
188 50 BMI dependent body shape change (Cornelissen et al., 2015; Irvine et al., 2018; Mölbert et al.,
189
190 51 2017; Szostak, 2018). In these perceptual body size estimation tasks, participants are typically
191
192 52 presented images of either a standard model, or an avatar of themselves, usually on a PC
193
194 53 monitor. The images vary in adiposity (indexed by BMI) and the participant's task is essentially
195
196 54 to decide which image best corresponds to the body size they believe themselves to have. Our
197
198 55 question is: what visual features do participants use to make these judgements about their own
199
200 56 body size when they are viewing such stimuli?
201
202
203

204 57 **1.1. Visual Cues to Body Size Judgements**

207 58 Previous research suggests two potential sets of cues that may drive performance on
208
209 59 perceptual body size estimation tasks: firstly, the width of the body in the stimuli and secondly,
210
211 60 the cues within the body outline. The first set of cues are straightforward. Previous studies have
212
213 61 noted that the width of the torso increases with increasing body mass index, particularly around
214
215 62 the waist region (BMI) (e.g., Cornelissen et al., 2009a; Tovée et al., 1999; Tovée &
216
217 63 Cornelissen, 2001). This “thickening” of the torso could thus provide an index of body mass.
218
219 64 The second set of cues are internal to the body outline. These include the saliency of bony
220
221 65 landmarks such as the collar bones or ribs, which become more obvious as body fat declines
222
223 66 (George et al., 2011). Additionally, as the amount of body fat increases, it is deposited as rolls
224
225 67 of fat, whose size and quantity could be used to estimate total body mass. Between these
226
227 68 extremes, the pattern of texture gradients across the surface of the body can potentially provide
228
229 69 a cue to the 3D shape of the body, such as size of the stomach (Cornelissen et al., 2013; Tovée
230
231 70 et al., 2002).
232
233
234
235
236

237
238
239
240 71 In support of the first hypothesis, a principal component analysis (PCA) of images of
241
242 72 female bodies varying in BMI, but facing forward in a standard pose, found that the change in
243
244 73 torso width was described by principal component 1 (PC1), and this factor was the main
245
246 74 predictor of body judgements (Tovée et al., 2002). Additionally, when the results of this PCA
247
248 75 were used to create a set of artificial bodies, simply varying PC1 was sufficient to drive the
249
250 76 perception of body weight change without varying any of the other shape dimensions (Smith
251
252 77 et al., 2007a). This suggests that simple changes in torso width are sufficient to drive the
253
254 78 perception of body mass.

255
256
257 79 This result is also consistent with a recent study which varied body orientation relative
258
259 80 to the observer (Cornelissen et al., 2018). The observer had to discriminate between pairs of
260
261 81 bodies in a 2-alternative forced choice task, based on differences in BMI. The finest
262
263 82 discrimination occurred for the bodies presented either in profile or at 45° relative to the
264
265 83 observer, and the worst discriminations occurred when the bodies were presented in front-view.
266
267 84 Most pertinently, the sensitivity of discrimination was predicted by the magnitude of the torso
268
269 85 width change detectable by the observer. As BMI increases, the degree of change in torso width
270
271 86 as a proportion of the total torso width, is greater in profile or at 45° than in front-view. This is
272
273 87 true for both CGI bodies and digital photographs of real bodies (Cornelissen et al., 2018). As
274
275 88 a result, judgements in profile or at 45° tend to be more accurate than those made in front-view.
276
277 89 This difference in performance and its correlation with the saliency of the visual cues to change
278
279 90 in torso width change, suggests that this is the cue that is being used to judge body size.

282 283 91 **1.2. Eye-Movement Studies**

284
285
286 92 Alternatively, there are also visual cues that are internal to the body outline that index
287
288 93 overall body mass, and several studies suggest that in practice these are the cues being used.
289
290 94 The evidence for this hypothesis is primarily based on eye-movement studies. For example,

296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354

95 women with anorexia nervosa fixate more on these body landmarks when making body size
96 judgements than control observers and are significantly better than the control observers at
97 judging the body size of low weight bodies (Cornelissen et al., 2015; George et al., 2012). This
98 suggests that the use of these cues may form the basis of a successful strategy in judging lower
99 BMI bodies. In addition, as mentioned above, increasing body fat changes the pattern of texture
100 gradients and shading cues across the surface of the body within the body outline (Cornelissen
101 et al., 2016b).

102 Several studies have suggested that stomach size, indexed through its depth, is a strong
103 cue to BMI (e.g., Rilling et al., 2009; Smith et al., 2007; Tovée et al., 1999). Eye-movement
104 data suggest control participants who are accurate at estimating their own BMI fixate primarily
105 on the stomach. Critically, these fixations fall *within* the body outline (Cornelissen et al., 2009b,
106 2016b; George et al., 2012). This is true whether observers are judging bodies seen in front-
107 view or viewed at a 45° angle. If they were simply viewing the degree to which the stomach
108 protrudes then their fixations should shift between central fixations on the torso in front-view
109 to fixations on the edge body outline in the 45° viewing angle. However, the fixations remain
110 centrally located (Cornelissen et al., 2009b, 2016b; George et al., 2012). This is surprising, as
111 if participants are asked to judge torso shape, they made eye-movements across the body and
112 sequentially fixated on either side of the torso edge (Cornelissen et al., 2009b). This suggests
113 that when viewing bodies at a 45° angle, the optimal fixation strategy for estimating stomach
114 depth would be to make fixations on both edges of the body corresponding to its outline.
115 However, under these viewing conditions, observers whose fixations are not concentrated
116 centrally within the body outline, and those who look more at the edge of the body are less
117 accurate in their body mass judgements (Cornelissen et al., 2016b). Eye movement data like
118 these therefore suggest that the principal cues being used to judge body mass are located within
119 the body outline.

1.3. Dissociation Between Fixation Patterns and the Allocation of Attention

A potential key flaw with these eye-movement studies is the assumption that visual attention is always aligned directly with the line of sight. A number of studies have suggested that this may not necessarily be the case (e.g., Datta & DeYoe, 2009; Ehinger & Rosenholtz, 2016; Gegenfurtner, 2016). For example, in judgements of a basketball scenario, a contingent-gaze paradigm suggests that the position of the player with the ball is used as an “anchor point” for an observer’s fixation while the relative position of the other players was estimated using the peripheral visual field (Ryu et al., 2013). Thus, a particular fixation point may just be a suitable point in the visual field from which to sample visual information using the retinal periphery, and not the complete focus of an observer’s attention. This therefore raises an alternative account of the eye-tracking studies of body size estimation. It is possible that instead of extracting information *within* the body outline, the eye-movement pattern is actually an efficient foraging strategy which allows a wider attentional window to extract *edge-based* cues from the torso while using a central looking strategy.

It is well known that resolution acuity (i.e., the smallest separation between two points that allows them to be perceived as separate) drops off dramatically from the central fovea towards the parafovea and beyond (Anderson, Mullen, & Hess, 1991; Carrasco, 2011; Pelli & Tillman, 2008). This necessarily means that the apparent sharpness of the torso edges when sampled by a strategy of viewing the centre of the body, would be reduced; put simply, the torso edges would appear blurry. However, it is important to remember that the visual system’s ability to resolve edge alignment, edge sharpness or smoothness, and curvature, i.e., exactly the kinds of low-level features that are likely to be needed to estimate the separation and shape of the torso edges, operate within the hyperacuity range (Carrasco, 2011). The phenomenon of hyperacuity is based not on the cone density of the retina, but on a cortical calculation which extrapolates from the limited sampling array to estimate a more detailed percept (Motter, 1998;

414
415
416 145 Gegenfurtner, 2016). This means that these spatial attributes can potentially be resolved to an
417
418 146 accuracy often an order of magnitude finer than that of resolution acuity, even in the presence
419
420 147 of a blurred stimulus. Therefore, there is no reason in principle why a foraging strategy that
421
422 148 appears to blur the edges of the object being judged will impair the visual system's ability to
423
424 149 discriminate the locations and shapes of those edges in calculating body size.

427 150 **1.4. The Bubbles Masking Technique**

430 151 A potential way of disambiguating these two possibilities, i.e., edge versus central
431
432 152 image information and gauging the location of the attentional window during the perceptual
433
434 153 judgement of body size, is the bubbles masking technique (Gosselin & Schyns, 2001). This
435
436 154 technique is a psychophysics paradigm that has been used to determine which visual cues are
437
438 155 being used in a categorisation task; i.e., which areas are diagnostic for a given judgement. For
439
440 156 example, the technique has been used to reveal which facial features drive the distinction
441
442 157 between neutral versus happy faces and male versus female faces. In the bubbles masking task,
443
444 158 parts of the stimuli are revealed by randomly allocated Gaussian "windows." These are circular
445
446 159 holes with blurred edges that perforate a uniform gray surface that overlies the stimulus (see
447
448 160 Figure 1 for an illustration). On each trial, observers make a categorical judgement based on
449
450 161 this partial information, e.g., "this is a male face" or, as in the current study, "that body is larger
451
452 162 than mine." Over multiple trials, all areas in the stimulus image are sampled and from this
453
454 163 unbiased sampling strategy, it is possible to calculate how effective each Gaussian window was
455
456 164 at independently determining the behavioural performance (Humphreys et al., 2006). Thus, it
457
458 165 should be possible to localise the areas of a body stimulus that are actually used when
459
460 166 participants make self-estimates of body size.

465 167 But the bubbles masking technique has its own potential flaw. It is possible that the
466
467 168 imposition of the bubble masks fundamentally changes the looking strategy (Gosselin &

473
474
475 169 Schyns, 2004; Murray & Gold, 2004). So, we address this problem by using an adapted bubbles
476
477 170 strategy which emphasises the distinction between central versus edge featural information
478
479
480 171 (Experiments 1 and 2). In addition, we also measure eye-movements to test whether the
481
482 172 underlying search strategy, reflected in eye fixation patterns, has changed from the up and
483
484 173 down the middle of the body fixation strategy reported by previous studies of self-estimation
485
486 174 of body size (Experiment 3).

487 488 489 175 **1.5. The Current Study**

490
491 176 Here we ask what visual cues do participants use when judging their own body size?
492
493 177 The literature reviewed above suggests that there are two potential sets of cues that participants
494
495
496 178 could be using to make these judgements: (1) information about the separation of the torso
497
498 179 edges and (2) information about body shape contained within the body outline. If the former
499
500 180 case is true, we should expect to find a dissociation between where participants look on the
501
502 181 stimulus bodies and the location of the regions on the bodies that are diagnostic for body size.
503
504 182 Specifically, we predict that the eye fixations should lie along the vertical midline of the body
505
506 183 stimuli, and the diagnostic regions should lie along the left and right torso edges. If, however,
507
508 184 the latter case is true, both the diagnostic regions and the eye fixations should be spatially
509
510 185 coincident, and both should be aligned with the vertical midline of the stimulus body.

511
512
513 186 In three experiments, we combine a modified bubbles masking technique together with
514
515 187 eye movement recording to distinguish between these two possibilities. All the studies were
516
517 188 completed by two sets of observers. In a pre-test screening process, we identified observers
518
519 189 who were accurate at estimating their own body size, and observers who were inaccurate. By
520
521
522 190 using both accurate and inaccurate observers we were able to compare the features important
523
524 191 for an accurate judgement with the regions which lead to a misestimation. As discussed above,
525
526 192 overestimation of body size in women with anorexia nervosa may arise from either one or both

532
533
534 193 of two factors; attitudinal or perceptual distortion. By testing nonclinical samples who
535
536 194 overestimate body size compared to those who are accurate at estimating body size and who
537
538 195 have the same psychological concerns, we can focus purely on perceptual factors as the basis
539
540 196 of the overestimation. Ultimately, we intend to extend this research to compare diagnostic
541
542 197 regions for self-estimates of body size in people with eating disorders with those from accurate
543
544 198 and overestimating individuals without eating disorders. However, these experiments make
545
546 199 heavy demands on participants. Therefore, as a first step in the introduction of the bubbles
547
548 200 paradigm into this research area, we felt it appropriate to recruit participants who had no history
549
550 201 of eating disorders.

553 554 202 **1.6. Overall Experimental Strategy**

555
556
557 203 In all three experiments, we recruited women with no history of eating disorders. For
558
559 204 each experiment, we used a standard yes-no body size estimation task (described below) to
560
561 205 identify a group of 12 women who estimated their body size accurately and a second group of
562
563 206 12 women who overestimated their body size. In addition, all participants were administered a
564
565 207 standard battery of psychometric tasks to estimate their psychological attitudes regarding their
566
567 208 body shape, weight, eating, and self-esteem, as well as report their symptoms of depression.
568
569 209 This allowed us to ensure that, in each of the three experiments, the groups of accurate body
570
571 210 size estimators and overestimators were comparable in terms of their psychological profiles,
572
573 211 chronological age, and BMI. We then used the bubbles masking technique with large
574
575 212 (Experiment 1) and small (Experiment 2) bubbles to identify the diagnostic regions that
576
577 213 allowed participants to judge their own body size against the stimulus presented. On each trial
578
579 214 of these tasks, participants had to decide whether the body in the masked image was smaller or
580
581 215 larger than they believe themselves to be. Because the two groups of participants differed only
582
583 216 in their accuracy at estimating their own body size (from the yes-no task), and did not differ in
584
585 217 any other way, we could use a spatial analysis to compare the diagnostic regions for self-

591
592
593 218 estimates of body size between them. Finally, in Experiment 3, we ran an eye movement
594
595 219 recording study to test whether the presence of the bubble masks caused a fundamental change
596
597 220 in looking strategy in Experiments 1 and 2. Specifically, we needed to know whether
598
599 221 participants had changed from an up-and-down the middle of the body viewing strategy, which
600
601 222 we would expect to see in the absence of bubbles, to an alternative strategy in which they
602
603 223 deliberately looked separately at the left and right torso edges.
604
605
606

607 224 **2. Experiment 1**

608 609 225 **2.1. Method**

610
611
612 226 The experimental procedures and methods for participant recruitment for this study
613
614 227 were approved by the local ethics committee at Northumbria University.
615
616

617 228 **2.1.1. Participants.**

618
619 229 Pilot testing showed that the maxima and minima in the group differences in correctly
620
621 230 responding in diagnostic areas that were biologically meaningful (e.g., edge of torso, central
622
623 231 abdomen, and gap between thighs) could be detected using a sample size of between 4 and 11
624
625 232 participants per group (alpha = 0.05 and power = 80%). To offset attrition in participant
626
627 233 numbers and/or unexpected sources of variability, we therefore recruited 12 participants per
628
629 234 group.
630
631
632

633 235 To be eligible to take part in this study, participants had to be female (as assigned at
634
635 236 birth), aged 18-35, with no history of eating disorders, and they had to have normal or
636
637 237 corrected-to-normal vision. We recruited 41 females into Experiment 1 from staff and students
638
639 238 at Northumbria University who carried out the initial psychometric and psychophysical tests.
640
641 239 We defined body size overestimators as those whose point of subjective equality (PSE) from
642
643 240 the yes-no-task (see below) was at least 2 BMI units above their measured BMI. Accurate
644
645 241 body-size estimators recorded a PSE within +/-1 BMI unit of their measured BMI. According
646
647
648
649

650
651
652 242 to these criteria, we identified 12 accurate body size estimators and 12 overestimators from the
653
654 243 initial sample of 41 consenting women and invited these individuals to complete the full study.
655
656 244 The characteristics of these 24 participants are reported in Table 1.
657
658

659 245 **2.1.2. Measures.**

660 246 **2.1.2.1. Psychometric and anthropometric measures.**

661
662
663
664 247 To measure the attitudinal component of body image, participants completed a number
665
666 248 of self-report questionnaires that measure body satisfaction and attitudes towards body shape,
667
668 249 weight and eating.
669
670

671 250 *2.1.2.1.1. Body Shape Questionnaire.* The 16-item Body Shape questionnaire (BSQ-
672
673 251 16) (Evans & Dolan, 1993) was used to assess participants' attitudes towards their body shape.
674
675 252 Items are rated along a 6-point Likert-type scale ranging from *never* (scored as 1) to *always*
676
677 253 (scored as 6). Items were summed to create a total score. A sample item is, "Have you been so
678
679 254 worried about your shape that you have been feeling you ought to diet."
680
681

682 255 *2.1.2.1.2. The Eating Disorders Examination Questionnaire.* The Eating Disorders
683
684 256 Examination Questionnaire (EDE-Q) is a 28-item self-report version of the Eating Disorder
685
686 257 Examination (EDE) interview (Fairburn & Beglin, 1994). It contains four subscales: the
687
688 258 Restraint subscale investigates the restrictive nature of eating, the Eating Concern subscale
689
690 259 measures the preoccupation with food and social eating, the Shape Concern subscale measures
691
692 260 dissatisfaction with body shape, and the Weight Concern subscale measures dissatisfaction
693
694 261 with body weight. Participants report how many days out of the past four weeks they have
695
696 262 experienced an item (e.g., "Have you been deliberately trying to limit the amount of food you
697
698 263 eat to influence your shape or weight [whether or not you have succeeded]") on a 7-point
699
700 264 Likert-type scale ranging from *No days* (scored as 0) to *Every day* (scored as 6). A global score
701
702 265 of overall disordered eating behaviour and subscale scores were calculated by averaging the
703
704
705
706
707
708

709
710
711 266 appropriate items, and frequency data on key behavioural features of eating disorders is
712
713 267 provided.
714
715

716 268 *2.1.2.1.3. Beck Depression Inventory.* The Beck Depression Inventory (BDI) was used
717
718 269 to measure levels of depressive symptomatology (Beck, Ward, Mendelson, Mock, & Erbaugh,
719
720 270 1961). It is a behavioural checklist that contains 21 items. Each item is rated on a 4-point scale,
721
722 271 ranging from 0 (no symptom of depression) to 3 (severe expression of depressive symptom).
723
724 272 Items are summed.
725
726

727 273 *2.1.2.1.4. Body mass index.* BMI was calculated from their weight and height measured
728
729 274 with a set of calibrated clinical SECA scales and a stadiometer, respectively.
730
731

732 275 ***2.1.2.2. Psychophysical measurements.***
733
734

735 276 *2.1.2.2.1. Yes-no task.* In this study, we apply classical psychophysical methods (cf.
736
737 277 Gardner, 1996) to measure two components of the participants' judgements of their own body
738
739 278 size: (a) the point of subjective equality (PSE) and (b) the difference limen (DL). The PSE is
740
741 279 the participant's subjective estimate of their body size. The DL is an estimate of how sensitive
742
743 280 a participant is to changes in body size and equates to the smallest difference in body size that
744
745 281 she can detect. To obtain these measurements, we use the method of constant stimuli in a yes-
746
747 282 no forced choice paradigm. This allows a psychometric function to be estimated. Here, the
748
749 283 psychometric function is a plot of the percentage of 'this image is larger than me responses' as
750
751 284 a function of the BMI of the stimuli presented, and the curve tends to have a sigmoidal shape.
752
753 285 The PSE is defined from the psychometric function as the BMI at which participants would
754
755 286 respond 'larger than me' 50% of the time. The DL is the difference in the BMI of the stimuli
756
757 287 falling between the 25% and 75% 'larger than me' response points (see Gescheider, 1997).
758
759 288 This range captures the steepness of the psychometric curve. Participants who are very
760
761
762
763
764
765
766
767

768
769
770 289 sensitive to small differences in body size will have a steeper psychometric function with a
771
772 290 correspondingly small DL.
773
774

775 291 In the yes-no task, participants were presented with a randomized sequence of images
776
777 292 of a standard CGI female model, standing in three-quarter view (for details of stimulus image
778
779 293 generation, see Cornelissen, 2016). Across the image set, BMI varied continuously from 12.5
780
781 294 to 44.5. On each trial of the task, one image was presented, and participants were required to
782
783 295 decide whether the body depicted was larger or smaller than they believed themselves to be.
784
785 296 Stimuli were presented on a 19" flat panel LCD screen (1280w × 1024h pixel native resolution,
786
787 297 32-bit colour depth) for as long as it took participants to make a decision. At the standard
788
789 298 viewing distance of ~60cm, the image frame containing the female body subtended ~26°
790
791 299 vertically and ~8° degrees horizontally. Each participant first judged seven images covering
792
793 300 the whole BMI range (from 12.5 to 44.5 in equal BMI steps) presented in two separate blocks.
794
795 301 Each stimulus image appeared 10 times in each block, and the order of presentation was
796
797 302 randomized. Based on the responses from each block, the participants' point of subjective
798
799 303 equality or PSE (the BMI they believe themselves to be) was calculated automatically by fitting
800
801 304 a cumulative normal distribution. These two values were then averaged to give an initial
802
803 305 estimate of the participant's PSE. Based on this initial estimate, the program presented a further
804
805 306 set of 21 images (spread over a range of 5 BMI units centred on the participant's initial PSE,
806
807 307 at a spacing of 0.25 units per image) for the participants to judge. Each image was presented
808
809 308 ten times in randomized order. This final set of judgements allowed us to plot the full
810
811 309 psychometric function (i.e. the percentage of 'larger than me' responses on the y-axis as a
812
813 310 function of stimulus BMI on the x-axis) and use probit analysis off-line to calculate a definitive
814
815 311 estimate of PSE as well as the difference limen or DL (that is how sensitive participants are to
816
817 312 changes in BMI). Participants were classified as accurate at body size estimation if their PSE
818
819
820
821
822
823
824
825
826

827
828
829 313 was within +/-1 BMI unit of their measured BMI and overestimators if their PSE was > 2 BMI
830
831 314 units above their measured BMI.
832
833

834 315 *2.1.2.2.2. Bubbles masking task.* We built a bubbles masking task that was inspired by,
835
836 316 but different from, the Bubbles paradigm developed by Gosselin and Schyns (2001). In these
837
838 317 authors' task, like ours, on every trial, participants are given a partial view of a stimulus through
839
840 318 a set of Gaussian windows (i.e., circular holes with blurred edges, see Figure 1). The holes are
841
842 319 punched, as it were, through a gray overlay that covers the stimulus image. In Gosselin and
843
844 320 Schyns (2001), the centre of any one Gaussian bubble can be located at any pixel location in
845
846 321 the stimulus image. However, in the current study, we were asking whether information from
847
848 322 the edges of the body outline, or the midline of the body, primarily drives decisions about self-
849
850 323 estimates of body size. For this reason, we wanted to constrain the location of the mask bubbles
851
852 324 into three columns. Bubbles in the left column of the stimulus overlay the right body edge and
853
854 325 allowed participants to see this edge only. (Here we use the anatomical convention where left
855
856 326 refers to the left side of the person in the stimulus image, from their point of view.) Bubbles in
857
858 327 the middle column overlay the midline of the woman in the stimulus, thereby restricting
859
860 328 participants' view to the midline of the body only. Bubbles in the right column of the stimulus
861
862 329 overlay the left body edge, and restricted participants' view to that region only (see Figure 1a).
863
864 330 This approach meant that we could carry out a spatial analysis of percentage correct responses
865
866 331 at each fixed bubble location, and explicitly test for differences in body size classification
867
868 332 between bubbles in the midline versus the two edge columns.
869
870
871
872

873 333 Bubbles were created dynamically as the program ran the task. On each trial, a stimulus
874
875 334 image was covered by an opaque grey overlay (RGB: 64, 64, 64 on a 0-256 range), punctured
876
877 335 by transparent bubbles whose centres were defined by the centres of an invisible, rectangular
878
879 336 grid of squares $3(w) \times 9(h)$, corresponding to the three columns (left edge, midline, and right
880
881 337 edge). Each square of the grid measured 100×100 pixels. In Experiment 1, the transparency
882
883
884
885

886
887
888 338 of the bubbles followed a 2D Gaussian distribution with a standard deviation of 0.56 degrees.
889
890 339 On each trial of the task, a subset of the bubble locations was chosen at random from this $3 \times$
891
892 340 9 array to be transparent, and participants had to decide, and respond by button press, whether
893
894 341 the underlying image (drawn from the same stimulus set as the yes-no task) was larger or
895
896 342 smaller than the participant believed themselves to be. Half of the images presented were
897
898 343 larger, and half of them smaller, and the order of image presentation was randomized across
899
900 344 trials. The particular pair of images presented to each participant were chosen based on their
901
902 345 difference limen (DL) in the yes-no task. The smaller image corresponded to the 25% response
903
904 346 rate in the yes-no task and the larger image the 75% response rate. Like Gosselin and Schyns
905
906 347 (2001), we sought to maintain participants' performance in the bubble mask task at ~75%
907
908 348 correct across the 2000 trials of the task. To do this, we calculated the correct response rate
909
910 349 after every 20 trials, and reduced the bubble count by 1, kept it the same or increased it by 1
911
912 350 depending on whether the participant's responses were below, at or above criterion (within +/-
913
914 351 15 %).

915
916
917
918
919 352 **2.1.3. Procedure.** To maximize participant's vigilance and minimize their fatigue,
920
921 353 they typically completed the experiment over the course of three sessions on three consecutive
922
923 354 days. On the first day, in a quiet, private testing room, participants gave written consent to take
924
925 355 part having read the study information sheet. Next, over the course of ~ 40 minutes, their height
926
927 356 and weight were measured, they were asked to complete the psychometric questionnaires, and
928
929 357 finally complete the yes-no psychophysical task. Participants who were eligible to complete
930
931 358 the full study (i.e., they fit the criteria either for accurate or overestimation of body size) carried
932
933 359 out the bubble masking task over the course of the next two sessions, each of which lasted
934
935 360 about 60 minutes. Trials were presented back to back, each new trial triggered by the
936
937 361 participant's button response. A pause was included after every 140 trials, giving the
938
939
940
941
942
943
944

945
946
947 362 opportunity for a break. Once all tasks were completed, participants were verbally debriefed
948
949 363 and given the opportunity to ask questions about the study.
950
951

952 364 **2.2. Results**

953
954
955 365 **2.2.1. Univariate statistics.** The right-hand columns in Table 1 show the output of
956
957 366 pairwise comparisons of the two group means, adjusted for multiple comparisons, using the
958
959 367 bootstrap resampling method with 10,000 bootstrap samples in PROC MULTTEST (SAS v9.4,
960
961 368 SAS Institute, North Carolina, USA). The effect sizes (Cohen's *d*) for these comparisons,
962
963 369 together with their 95% CI, are also included (Kadel & Kip, 2012). Despite some of the
964
965 370 Cohen's *d* values representing medium-to-large effect sizes, almost all of them include 95%
966
967 371 confidence intervals that include zero. This is likely attributable to the relatively small number
968
969 372 of participants. The only confidence intervals that do not include zero, correspond to very large
970
971 373 effect sizes, are these also associated with statistically significant pairwise comparisons. Table
972
973 374 1 confirms that accurate estimators were within ~0.25 BMI units of their actual BMI, on
974
975 375 average, as compared to overestimators who overestimated by ~4 BMI units. With respect to
976
977 376 the World Health Organization's BMI classification scheme (World Health Organization,
978
979 377 2003), the numbers of participants who fell into the underweight, normal, overweight, and
980
981 378 obese categories for the accurate and overestimating groups, respectively, were: 0, 11, 1, 0, and
982
983 379 1, 9, 2, 0. The mean BSQ scores shown in Table 1 are consistent with mild concern with body
984
985 380 shape (Evans & Dolan, 1993). The mean BDI scores for the accurate and overestimating groups
986
987 381 are both consistent with the mild range. The EDEQ subscales in both groups were within 1SD
988
989 382 of the normative means for women within this age group (Mond et al., 2006). Cronbach's
990
991 383 alphas for the BDI, BSQ, and EDEQ in the two groups (combined) were .92, .95, and .94,
992
993 384 respectively.
994
995
996
997
998
999
1000
1001
1002
1003

1004
1005
1006 385 **2.2.2. Where are the diagnostic regions for the accurate and overestimating**
1007
1008 386 **groups?** In Experiment 1, on each trial, the stimulus to be judged was visible through bubbles
1009
1010 387 picked at random from an array of $3(w) \times 9(h)$ bubble locations. By the end of the task, the
1011
1012 388 number of times that any particular bubble location had been used, as well as the percentage of
1013
1014 389 those presentations that were associated with a correct response were recorded for each
1015
1016 390 participant. Therefore, a percentage correct could be calculated for every bubble location,
1017
1018 391 separately for each participant.
1019
1020
1021

1022 392 The adaptive procedure ensured that participants' responses tracked close to the
1023
1024 393 criterion we set for the masking task, namely that 75% of the choices they made across 2000
1025
1026 394 trials should be correct, and Table 1 confirms this. To achieve this criterion performance, both
1027
1028 395 groups required on average a bubble count of ~ 5 (see Table 1). As Gosselin and Schyns (2001)
1029
1030 396 argue, if all regions in our stimuli were equally informative about participants' perceptions of
1031
1032 397 their own body size, then the percentage of correct responses at each location in our mask array
1033
1034 398 should match the same criterion: i.e., the response rate for every bubble location should also
1035
1036 399 be 75% correct. However, if there is a subset of areas in the stimuli that are particularly
1037
1038 400 informative about the body size participants' believed they have, then we should expect the
1039
1040 401 response rates in bubbles overlying these regions to be significantly higher than 75%. Such
1041
1042 402 areas should correspond to regions that are diagnostic of participants' body size beliefs,
1043
1044 403 according to the terminology of Gosselin and Schyns (2001). However, for this to be true, and
1045
1046 404 for average performance across the set of trials to be 75% correct, we should also expect the
1047
1048 405 response rates in bubble locations that overlie non-informative regions in the stimuli to be
1049
1050 406 lower than 75% correct. Note that the non-informative regions do not necessarily need to be
1051
1052 407 significantly lower than 75%. They might reach perhaps only $\sim 72\%$ for example, but
1053
1054 408 nevertheless be widely distributed enough across the sample space so that the average across
1055
1056 409 the whole space is 75%.
1057
1058
1059
1060
1061
1062

1063
1064
1065 410 To test these predictions, we ran three generalized linear mixed models (GLMMs) of
1066
1067 411 the normalized percentage responses across different bubble locations, using PROC MIXED
1068
1069 412 in SAS v9.4 (SAS Institute, North Carolina, USA). To normalize the data, we calculated the
1070
1071 413 mean percentage correct across all 3(w) × 9(h) bubble locations for each participant, and then
1072
1073 414 subtracted these global means from the percentage correct for each individual bubble location,
1074
1075 415 separately for each participant. For spatially sampled data, we cannot assume that the
1076
1077 416 percentage correct responses at each bubble location are statistically independent of each other.
1078
1079 417 Specifically, we must assume that percentage correct will covary across bubble locations, and
1080
1081 418 that the magnitude of this spatial covariation is inversely proportional to the bubbles' proximity
1082
1083 419 to each other. Therefore, in all three models we took account of the repeated measures within
1084
1085 420 subjects – i.e., each subject was presented 27 mask locations in all (defined by row and column
1086
1087 421 co-ordinates). In addition, we controlled for spatial covariance by incorporating the spatial
1088
1089 422 variability into the statistical models by specifying a Gaussian spatial correlation model for the
1090
1091 423 model residuals (Littell et al., 2006). The general form of the model we fitted was:

$$1092$$
$$1093$$
$$1094$$
$$1095$$
$$1096$$
$$1097$$
$$1098$$
$$1099$$
$$1100$$
$$1101$$
$$1102$$
$$1103$$
$$1104$$
$$1105$$
$$1106$$
$$1107$$
$$1108$$
$$1109$$
$$1110$$
$$1111$$
$$1112$$
$$1113$$
$$1114$$
$$1115$$
$$1116$$
$$1117$$
$$1118$$
$$1119$$
$$1120$$
$$1121$$
$$E[Y|u] = X\beta + Zu + e$$

1098 425 Where $E[Y|u]$ is the conditional probability of the outcome given the random model
1099
1100 426 effects, $X\beta$ are the fixed effects, Zu are the random effects, and e the error term. Spatial
1101
1102 427 correlation was reflected in R , the covariance matrix of the model errors. The fixed effects in
1103
1104 428 all models comprised two class variables: ROW (i.e., the index for each row of the grid of
1105
1106 429 bubbles which could take values 1 to 9 inclusive) and COLUMN (i.e., the index for each
1107
1108 430 column of the grid of bubbles which could take values 1 to 3 inclusive). This means that the
1109
1110 431 location of each bubble in the 3 × 9 mask array was uniquely addressed, like an x,y coordinate,
1111
1112 432 by the combination of the two fixed effect variables, ROW and COLUMN. Where relevant,
1113
1114 433 we also included GROUP (i.e., accurate body size estimators versus overestimators) as a fixed
1115
1116
1117
1118
1119
1120
1121

1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180

434 effect when we wanted to compare performance between accurate body size estimators versus
435 overestimators. The most important outcomes from the statistical modelling were to identify:

436 MODEL 1: Where were the areas diagnostic of body size (i.e., > 75% correct) for
437 accurate estimators?

438 MODEL 2: Where were the areas diagnostic of body size (i.e., > 75% correct) for
439 overestimators?

440 MODEL 3: Where were the significant differences in diagnostic areas for body size
441 comparing accurate estimators with overestimators?

442 To do this, for each model, we computed the predicted population margins from the
443 GLMMs and compared them using tests for simple effects by partitioning the interaction
444 effects, controlling for multiple comparisons. In other words, for MODELS 1 and 2, we used
445 the fitted GLMMs to predict the percentage of correct responses in each bubble location and
446 asked whether that percentage was significantly greater than 75%. These predictions are
447 corrected for the repeated measures design, the spatial covariance in the data and the fact that
448 we carried out multiple comparisons. For MODEL 3 we used the fitted GLMM to predict the
449 difference in the percentage of correct responses comparing accurate body size estimators and
450 overestimators, and asked whether each of these differences was significantly different from
451 zero. An additional constraint for MODEL 3 was that a bubble location was only deemed to
452 show a statistically significant difference between accurate and overestimators if that location
453 had a response rate significantly greater than 75% ($p < .01$) from either MODEL 1 or MODEL
454 2, as well as showing a significant group difference. For completeness, we report the fixed
455 effects in each model below, and then show the key outcomes, i.e., the predicted percentages
456 of correct responses in each bubble location, in Figure 2.

1181
1182
1183 457 The Type III tests of fixed effects for MODEL 1 were: ROW $F(4, 44) = 1.04, p = .40$;
1184
1185 458 COLUMN $F(10, 110) = 25.02, p < .001$; ROW \times COLUMN $F(40, 440) = 5.19, p < .001$.
1186
1187

1188 459 The Type III tests of fixed effects for MODEL 2 were: ROW $F(4, 44) = 0.27, p = .90$;
1189
1190 460 COLUMN $F(10, 110) = 12.98, p < .001$; ROW \times COLUMN $F(40, 440) = 7.37, p < .001$.
1191
1192

1193 461 The Type III tests of fixed effects for MODEL 3 were: GROUP $F(1, 22) = 0.00, p =$
1194
1195 462 $.99$; ROW $F(4, 88) = 0.23, p = .92$, COLUMN $F(10, 220) = 36.91, p < .001$; GROUP \times ROW
1196
1197 463 $F(4, 88) = 1.16, p = .33$; COLUMN \times GROUP $F(10, 220) = 2.39, p = 0.01$; ROW \times COLUMN
1198
1199 464 $F(40, 880) = 10.28, p < .001$; GROUP \times ROW \times COLUMN $F(40, 880) = 2.05, p < .001$.
1200
1201

1202 465 In principle, a significant fixed effect of ROW means that, averaged across columns,
1203
1204 466 there would be a significant linear increase/decrease in percentage correct responses as a
1205
1206 467 function of ROW – i.e., a tilt to the 2D regression plane. Similarly, a significant fixed effect of
1207
1208 468 COLUMN would mean that, averaged across rows, there would be a significant linear
1209
1210 469 increase/decrease in percentage correct responses as a function of COLUMN. A significant
1211
1212 470 interaction between ROW \times COLUMN would mean that the degree of tilt in the 2D regression
1213
1214 471 plane with respect to ROW, say, changes as a function of COLUMN. As the foregoing
1215
1216 472 description of the fixed effects in the GLMMs makes clear, it is encouraging that we see
1217
1218 473 statistically significant interactions between ROW and COLUMN in all three models. This
1219
1220 474 strongly suggests that there are indeed statistically significant diagnostic regions of interest.
1221
1222 475 However, analysis of the fixed effects alone cannot reveal the specific locations of the
1223
1224 476 diagnostic bubbles. For this, we need post-hoc comparisons, to which we now turn.
1225
1226
1227

1228 477 The first two columns in Figure 2a show the outcomes of the analyses of simple effects
1229
1230 478 from MODEL 1 and MODEL 2, for accurate body size estimators and overestimators,
1231
1232 479 respectively. Circles correspond to mask locations where correct response rates were
1233
1234 480 significantly higher than criterion (i.e., 75%), based on the GLMMs, and which can therefore
1235
1236
1237
1238
1239

1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298

481 be considered diagnostic regions. The red/orange/yellow coloured overlay represents the
482 averaged and smoothed raw data above criterion, referred to henceforth as a heat map.

483 For the accurate estimators, the circles *a* (80.4%, 95%CI 79.0 – 81.8%) and *c* (82.0%,
484 95%CI 80.7 – 83.4%) correspond to the peak LSmean response rates for the left and right
485 columns of mask bubbles respectively, and circle *b* (78.4%, 95%CI 77.1 – 79.8%) is the closest
486 mask bubble adjacent to both *a* and *c*. Circle *d* (78.4%, 95%CI 77.1 – 79.8%) corresponds to
487 the peak LSmean response rate for the central column of mask bubbles. Therefore, while it is
488 true that the central abdomen provides information that is diagnostic about body size for
489 accurate estimators, the left and right torso edges appear to provide more information, and this
490 difference is statistically significant for the left torso edge (i.e., the 95% confidence interval for
491 *c* does not overlap with those for *b* or *d*).

492 For the overestimators, circles *e* (82.0%, 95%CI 80.8 – 83.3%) and *g* (80.2%, 95%CI
493 78.9 – 81.4%) correspond to the peak LSmean response rate for the left and right sides of the
494 torso, and circle *f* (77.1%, 95%CI 75.9 – 78.4%) is the closest mask bubble adjacent to both *e*
495 and *g*. Circle *h* (77.9%, 95%CI 76.7 – 79.2%) corresponds to the peak LSmean response rate
496 for the central column of mask bubbles. Therefore, unlike the accurate estimators, the midline
497 is providing diagnostic information about the face. As with the accurate estimators, the midline
498 is also providing diagnostic information about the abdomen. However, the upper right torso
499 and the left hip are providing more, and this difference is statistically significant for the upper
500 right torso (i.e. the 95% confidence interval for circle *e* does not overlap with those for *f* or *h*).

501 The right most column in Figure 2a shows where diagnostic information about body
502 size differs significantly between accurate and overestimators. Specifically, accurate estimators
503 made significantly more use of information from the upper thigh gap and the left edge of the

1299
1300
1301 504 abdomen (red/yellow colours), whereas overestimators made significantly more use of
1302
1303 505 information from the right upper torso/arm and the face (blue/cyan colours).
1304
1305

1306 506 **2.3. Discussion**

1307
1308

1309 507 The results of Experiment 1 suggest that while both groups utilised information from
1310
1311 508 the middle of the stimulus body as well as its edges, the edges provided the most diagnostic
1312
1313 509 information (i.e., were more influential in driving participants' decisions in the categorisation
1314
1315 510 task). Additionally, the two groups differed significantly in the edge cues used. While the
1316
1317 511 accurate estimators made most use of the left flank and thigh gap, the overestimators used the
1318
1319 512 face and right arm/chest area more. Interestingly, eye-tracking studies suggest that women with
1320
1321 513 anorexia nervosa, who also overestimate body size, also fixate more on the face than
1322
1323 514 nonclinical controls who accurately estimate body size (Cornelissen et al., 2016b). Accurate
1324
1325 515 estimators also showed a distribution of diagnostic areas that are more evenly spread onto both
1326
1327 516 sides of the body, whereas the diagnostic areas of overestimators showed a bias onto one side
1328
1329 517 of the torso (see Figure 2a).
1330
1331
1332

1333 518 Even though the evidence from Experiment 1 suggests that body edges provide
1334
1335 519 diagnostic information for body size judgements, some mid-body features were still used, i.e.,
1336
1337 520 the face and thigh gap. Therefore, in order to provide a more detailed picture of the edge cues
1338
1339 521 used, we decreased the size of the bubbles from 100×100 pixels to 40×40 pixels in
1340
1341 522 Experiment 2. With this strategy, by providing more bubbles that are smaller in size, a more
1342
1343 523 detailed picture of the diagnostic information may be gathered.
1344
1345

1346 524 **3. Experiment 2**

1347
1348

1349 525 **3.1. Method**

1350
1351

1352 526 **3.1.1. Participants.** The selection criteria and methods of participant recruitment were
1353
1354 527 the same as for Experiment 1. Accordingly, we identified 12 accurate body size estimators and
1355
1356
1357

1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416

528 12 overestimators from an initial sample of 41 consenting women, to take part in the complete
529 study. These participants' characteristics are reported in Table 2.

530 **3.1.2. Measures.** The psychometric and psychophysical tasks were identical to
531 Experiment 1. The only difference in the bubble mask task was that we used a finer scale
532 rectangular grid of 9(w) × 21(h) squares (each of which measured 40 × 40 pixels), to locate the
533 bubble centres. The transparency of these smaller bubbles followed a 2D Gaussian distribution
534 with a standard deviation of 0.29 degrees, and the bubble count was increased or decreased by
535 2.

536 **3.2. Results**

537 **3.2.1. Univariate statistics.** Table 2 confirms that accurate estimators were within
538 ~0.25 BMI units of their actual BMI, on average, as compared to overestimators who
539 overestimated by ~4 BMI units. With respect to the World Health Organization's BMI
540 classification scheme (WHO, 2003), the numbers of participants who were classified into the
541 underweight, normal, overweight, and obese categories for the accurate and overestimating
542 groups, respectively, were: 0, 10, 1, 1, and 2, 8, 2, 0. Cronbach's alphas for the BDI, BSQ, and
543 EDEQ in the two groups (combined) were .92, .96, and .97, respectively. The mean BSQ scores
544 shown in Table 2 are consistent with mild concern with body shape (Evans & Dolan, 1993).
545 The mean BDI scores for the accurate and overestimating groups are consistent with the
546 minimal and mild ranges respectively. The EDEQ subscales in both groups were within 1SD
547 of the normative means for women within this age group (Mond et al., 2006). Table 2 shows
548 that the adaptive procedure maintained participant performance very close to 75% correct in
549 both groups, and that they required ~18-19 bubbles on average to achieve this performance.

550 **3.2.2. Where are the diagnostic regions for the accurate and overestimating**
551 **groups?** The rationale for the analysis procedures in Experiment 2 was identical to those for

1417
1418
1419 552 Experiment 1. Therefore, the treatment of data was the same, and we fitted the same 3 GLMMs
1420
1421 553 as in Experiment 1. The only difference was in the resolution of the bubble mask, which
1422
1423 554 comprised $9(w) \times 21(h)$ bubble locations.
1424
1425

1426 555 The Type III tests of fixed effects for MODEL 1 were: ROW $F(10, 110) = 10.44, p <$
1427
1428 556 $.001$; COLUMN $F(22, 242) = 5.88, p < .001$; ROW \times COLUMN $F(220, 2420) = 2.39, p <$
1429
1430 557 $.001$.
1431
1432

1433 558 The Type III tests of fixed effects for MODEL 2 were: ROW $F(10, 110) = 15.51, p <$
1434
1435 559 $.001$; COLUMN $F(22, 242) = 8.21, p < .001$; ROW \times COLUMN $F(220, 2420) = 3.13, p <$
1436
1437 560 $.001$.
1438
1439

1440 561 The Type III tests of fixed effects for MODEL 3 were: GROUP $F(1, 22) = 0.00, p =$
1441
1442 562 $.99$; ROW $F(10, 220) = 24.7, p < .001$; GROUP \times ROW $F(10, 220) = 1.21, p = .28$; COLUMN
1443
1444 563 $F(22, 484) = 12.56, p < .001$; COLUMN \times GROUP $F(22, 484) = 1.51, p = .06$; ROW \times
1445
1446 564 COLUMN $F(220, 4840) = 4.13, p < .001$; GROUP \times ROW \times COLUMN $F(220, 4840) = 1.38,$
1447
1448 565 $p < .001$.
1449
1450

1451 566 As before, the first two columns in Figure 2b show the outcomes from MODEL 1 and
1452
1453 567 MODEL 2, for accurate body size estimators and overestimators respectively. Circles
1454
1455 568 correspond to mask locations where correct response rates were significantly higher than
1456
1457 569 criterion (i.e., 75%), based on the GLMMs. The heat maps represent the smoothed, averaged
1458
1459 570 raw data above criterion. For the accurate estimators, the bubble locations corresponding to
1460
1461 571 significant diagnostic information about body size are clustered continuously along the edge of
1462
1463 572 the right lower chest and abdomen, the edge of the left waist and upper hip, and the thigh gap
1464
1465 573 (again using anatomical conventions for left and right). The overestimators show a very similar
1466
1467 574 pattern along the right edge of the upper body and a more extensive cluster along the left body
1468
1469 575 edge extending to the chest. However, it appears that the overestimators do not make use of the
1470
1471
1472
1473
1474
1475

1476
1477
1478 576 thigh gap. The right-hand column in Figure 2b shows where diagnostic information about body
1479
1480 577 size differs significantly between accurate and overestimators. Specifically, accurate estimators
1481
1482 578 made significantly more use of information from the upper thigh gap and a small region just to
1483
1484 579 the right of midline in the upper abdomen (red/yellow colours). In comparison, the
1485
1486 580 overestimators made more use of information on the right abdominal edge, as well as the left
1487
1488 581 upper quadrant of the abdomen (blue/cyan colours).

1491 582 **3.3. Discussion**

1494 583 The results of Experiment 2 suggest that for both groups the edges of the body stimuli
1495
1496 584 were instrumental in driving self-estimates of body size. Again, the two groups differed to
1497
1498 585 some extent in cues used, with accurate estimators using the information about the thigh gap,
1499
1500 586 and a region in the upper abdomen, while the overestimators used more cues from the right
1501
1502 587 edge of the abdomen and an upper area of the abdomen. These results provide a more detailed
1503
1504 588 picture of the diagnostic areas driving self-estimates of body size.

1507 589 However, as described in the Introduction, it is possible that the presence of the bubbles
1508
1509 590 and the partial view of the stimulus that this provides, changes the observer's looking strategy.
1510
1511 591 Therefore, we have measured the eye-movements of our participants to identify if the up-down
1512
1513 592 looking pattern reported by prior studies of size estimation changes when a bubble mask task
1514
1515 593 is used (Cornelissen et al., 2016b).

1519 594 **4. Experiment 3**

1521 595 **4.1. Rationale**

1524 596 We wanted to know where participants were fixating when they carried out the bubble
1525
1526 597 masking task with large and small bubbles. Therefore, in a third sample of participants, we
1527
1528 598 recorded the movements of the right eye during 200 trials of each version of the bubble mask
1529
1530 599 task. In addition, we also wanted to identify any differences in gaze patterns between the bubble

1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593

600 mask task as carried out in Experiments 1 and 2, compared to using the same size bubbles and
601 the same task – i.e., judging whether the presented image was larger or smaller than the
602 participant believed themselves to be, but now with all of the bubbles always set to transparent.
603 These latter conditions, 200 trials with large bubbles all open and 200 trials with small bubbles
604 all open, were the closest we could get to normal viewing using the bubbles task, and still
605 permitting maximum visibility of all parts of the stimuli simultaneously, on every trial. Given
606 that the view of the body per trial during the actual bubbles mask task is so restricted, we fully
607 expected that there should be greater dispersion of fixations across space, when the data were
608 binned over the course of 200 trials. Nevertheless, the critical question was whether participants
609 adopted a different viewing strategy compared to what is usually seen when participants view
610 non-masked bodies: i.e., looking up and down the midline of the body (see e.g., Cornelissen et
611 al., 2016b). Specifically, given the evidence from Experiments 1 and 2 that the body edges
612 provide diagnostic information for self-estimates of body size, we needed to know whether
613 fixation patterns during the bubble masking task also split into two distinct distributions, with
614 their peaks similarly aligned with the left and right body edges, instead of the midline.

4.2. Method

4.2.1. **Participants.** The selection criteria and methods of participant recruitment were the same as for Experiments 1 and 2. Accordingly, we identified 12 accurate body size estimators and 12 overestimators from an initial sample of 36 consenting women, to take part in the complete study. The characteristics of these 24 participants are reported in Table 3.

4.2.2. **Measures.** The psychometric and psychophysical tasks were identical to Experiments 1 and 2.

4.2.3. **Eye movement recordings.** Movements of the right eye were recorded with an Eyelink 1000 eye-tracker at a sample rate of 1000Hz. Stimuli were presented on a flat 19” CRT monitor while participants sat at a table with their heads restrained by a combined chin and

1594
1595
1596 625 forehead rest. At the standard viewing distance of ~60cm, the image frame containing the
1597
1598 626 female body subtended ~26° vertically and ~8° degrees horizontally. At the start of each block
1599
1600 627 of 200 trials, participants' eye movements were calibrated using a 9-point calibration screen.
1602
1603 628 Once the calibration procedure was validated, the experimental task began. We randomized the
1604
1605 629 order of the four versions of the masking task: large bubbles, large bubbles open, small bubbles,
1606
1607 630 and small bubbles open. While we did record participants' button responses in the task, there
1608
1609 631 were not enough trials to warrant a spatial analysis of these behavioural data (i.e., 1/10th of the
1610
1611 632 number of trials in Experiments 1 and 2). Nevertheless, the average accuracy of responding
1612
1613 633 over the 200 trials for large bubbles, large bubbles open, small bubbles, and small bubbles open
1614
1615 634 was: 69%, 88%, 67%, and 87%, respectively, for accurate estimators. The equivalent
1616
1617 635 performance for overestimators was: 69%, 98%, 69%, and 96%, respectively. Tests of location
1618
1619 636 showed that all these values are significantly better than guessing (i.e., 50% accuracy), even
1620
1621 637 though participants' performance had not stabilized at the ~75% criterion, which would be
1622
1623 638 expected had they carried out all 2000 trials of the main tasks.

1624
1625
1626
1627 639 The Eyelink 1000 system uses a saccade-picker approach to identify saccades by
1628
1629 640 applying an exclusive OR rule to three thresholds: velocity (30 degrees/sec), acceleration (8000
1630
1631 641 degrees/sec²), and distance moved between samples (0.1 degrees). It then treats the rest of the
1632
1633 642 (non-blink) data as fixations, assuming that the 'not in a saccade' condition is maintained for
1634
1635 643 at least 50ms. The stated accuracy of the system is down to a resolution of 0.15°, though 0.25°
1636
1637 644 to 0.5° is typical.

1640 645 **4.3. Results**

1641
1642
1643 646 **4.3.1. Univariate statistics.** Table 3 confirms that accurate estimators were within
1644
1645 647 ~0.25 BMI units of their actual BMI, on average, as compared to overestimators who
1646
1647 648 overestimated by ~4 BMI units. With respect to the World Health Organization's weight

1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711

649 classification scheme (WHO, 2003), the numbers of participants who fell into the underweight,
650 normal, overweight, and obese categories for the accurate and overestimating groups,
651 respectively, were: 0, 11, 0, 1, and 1, 9, 1, 1. Cronbach's alphas for the BDI, BSQ, and EDE-
652 Q in the two groups were .90, .93, and .94, respectively. The mean BSQ scores shown in Table
653 3 are both consistent with mild concern with body shape (Evans & Dolan, 1993). The mean
654 BDI scores for the accurate and overestimating groups are consistent with the minimal and
655 mild ranges, respectively. The EDE-Q subscales in both groups all fall within 1SD of the
656 normative means for women within this age group (Mond et al., 2006).

4.3.2. Where were participants fixating? The main question we wanted to address

658 was whether participants were fixating primarily within the midline of the stimuli or along the
659 body edges, during each of the four conditions: i.e., masking task with: large bubbles; large
660 bubbles open; small bubbles; and small bubbles open. Therefore, our analyses focus on within
661 task comparisons rather than between task comparisons. After blinks and saccades were
662 removed from the eye movement time series, the only additional data filtering we applied was
663 to remove the first 300msec post stimulus onset, as otherwise this would include the initial
664 fixation which was determined by the fixation cross and not by the observer. In order to
665 examine the spatial distributions of fixations, we constructed a sampling grid of square cells
666 (20 × 20 pixels each) and applied it to the fixation data that were recorded within the central
667 600(w) × 1020(h) pixels of the stimulus array. This cell size (20 × 20 pixels) represents a
668 compromise between capturing as many fixation samples per cell as possible to optimize
669 statistical power (which ideally requires large cells) versus retaining good anatomical
670 resolution (which ideally requires small cells) (cf. George et al., 2012). Having binned the
671 fixation data in this way, we calculated the percentage of the total fixation samples in each bin,
672 separately for each task and participant. These fixation density data were then converted to z-
673 scores which are presented as heat maps in Figure 3.

1712
1713
1714 674 Figure 3 shows clearly that, irrespective of whether they viewed stimuli through small
1715
1716 675 or large bubble masks, or whether they were accurate body size estimators or overestimators,
1717
1718 676 participants always showed a spatially more distributed gaze pattern during the bubble masking
1719
1720 677 task as compared to viewing the stimuli when all bubbles were open. The critical question for
1721
1722 678 the current study, however, is whether the gaze patterns for the bubbles task remain centred on
1723
1724 679 the midline, or whether they break apart into two distributions: one centred on the left torso
1725
1726 680 edge and the other on the right. Inspection of the black contours in Figure 3, which represent
1727
1728 681 the three standard deviation limits in each heat map, would suggest that participants' fixations
1729
1730 682 remained densest in the midline irrespective of task type or group assignment. To quantify this,
1731
1732 683 we split each fixation density map into three columns of equal width, corresponding to the large
1733
1734 684 bubble diameters at 100 pixels. We then calculated the total percentage of the fixation samples
1735
1736 685 in each column, separately for each participant and for each task, and used PROC MIXED in
1737
1738 686 SAS v9.4 (SAS Institute, North Carolina, USA) to test for differences between the average
1739
1740 687 fixation density in each column. Table 4 shows the outcome including the post-hoc
1741
1742 688 comparisons, controlled for multiple comparisons, between the left and middle columns and
1743
1744 689 the right and middle columns of fixations. There is no case in Table 4 where both left and right
1745
1746 690 columns of fixation data are significantly larger than the middle column. Therefore, we found
1747
1748 691 no compelling evidence that participants' fixation patterns divided into separate distributions
1749
1750 692 coincident with the edge regions diagnostic of body size. However, for accurate observers
1751
1752 693 during the masking task, there was evidence that their gaze patterns shifted to the left,
1753
1754 694 particularly in the chest region.

1755 695 **4.3.3. Direct comparison between eye fixations and psychophysical performance.**

1756 696 Clearly, direct comparisons between Experiments 1 and 3 and between Experiments 2 and 3
1757
1758 697 were not feasible because the outcome measures, tasks, and participant groups were all
1759
1760 698 different. Moreover, the spatial sampling of data in the three experiments was not directly
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770

1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829

699 comparable. Nevertheless, we attempted to make approximate comparisons as follows. First,
700 we resampled the eye movement data for each participant to match that for the bubble masking
701 tasks. To do this, we used 20×20 pixel sample bins placed at the centres of the small and,
702 separately, the large bubble masks. This procedure spatially co-registered the eye-movement
703 data precisely with the large and small bubble mask psychophysical data. Then, we converted
704 both the behavioural psychophysical data and the eye-movement data to z -scores, and re-ran
705 the GLMMs, separately for the psychophysics and eye-movement data. This allowed us to
706 compute marginal means (i.e., LSmeans in SAS) with their accompanying 95% confidence
707 intervals for the data at each sample point, and these are plotted in Figure 4. In each case, the
708 solid black lines represent the eye-movement data, and the solid white lines the psychophysical
709 data. All error bars represent 95% confidence intervals in units of z -scores. The locations of
710 the horizontal slices through the combined datasets are indicated by letter groups: *A, B, & C*
711 and *D, E, & F*, for the large and small bubble mask datasets, respectively. Finally, there is a
712 small horizontal offset in the x-axes for the eye-movement and psychophysical data, so that
713 error bars do not overlap. Figure 4 confirms that eye fixations remained densest in the mid-line
714 of the body, while the regions diagnostic of body size were concentrated on the edges.

5. General Discussion

716 In Experiment 1, the results of the modified bubbles technique (using the larger
717 bubbles) suggest that the key areas of the image for accurate self-assessment of body size are
718 on the edge of the torso at waist height on either side of the body. Both the left and right edges
719 of the torso are of equal importance in making the judgement. Overestimating observers favour
720 the right side of the image relative to the left side, as illustrated by the comparison of accurate
721 and overestimators in Figure 2a. In Experiment 2, the results of the bubbles technique (using
722 the smaller bubbles) suggests that the key areas are located along the outline of the torso on
723 either side of the body and at the thigh gap. Once again, both sides of the body have equal

1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888

724 importance in accurate judgements, but there is a bias towards one side of the body in
725 overestimators as illustrated by the comparison of accurate and overestimators in Figure 2b. It
726 seems that an equal division of visual attention to both side of the torso outline may be key to
727 accurate judgements.

728 A potential concern is that the use of the bubble masks significantly changes the looking
729 strategy used to assess the stimuli (Gosselin & Schyns, 2004; Murray & Gold, 2004). However
730 in face experiments, the diagnostic areas of the face identified by the bubbles techniques for a
731 particular task are consistent with those identified using other methods, such as comparing the
732 performance with isolated parts of the face (e.g., Bassili, 1979; Calder, Young, Keane, & Dean,
733 2000), using reverse correlation (Jack, Caldara, & Schyns, 2012; Yu et al., 2012), and eye-
734 tracking (Blais et al., 2017). In Experiment 3, the addition of eye-tracking to the bubbles
735 paradigm shows the visual fixations are clearly in the centre of the torso (Figures 3 & 4). This
736 pattern of fixations is very similar to that reported by previous studies which have not used a
737 masking paradigm, but have instead allowed a free, unoccluded view of the body stimuli during
738 self-estimates of body size (Cornelissen et al., 2016b; George et al., 2012). This suggests that
739 the use of the bubbles technique is not qualitatively altering the fixation pattern that our
740 observers are using to estimate the size of their own body (Gosselin & Schyns, 2004). However,
741 although the fixations fall within the centre of the stimulus torso, the key regions of the torso
742 for accurate judgements are clearly on its edge (Figure 4). In short, the eye-movement results
743 suggest a clear dissociation between fixation location and the location of the regions of the
744 body stimuli that are diagnostic for self-estimates of body size.

745 At first, this dissociation might seem counterintuitive. The physical constraints of the
746 retina mean that detailed spatial information can only be sampled from a small central area of
747 around 2°, corresponding to the fovea (Levi, Klein, & Aitsebaomo, 1985). As a result,
748 information in detail and colour can only be collected in small snapshots corresponding to an

1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947

749 observer's individual fixations (Miller & Bockisch, 1997). Thus, the failure to fixate the key
750 regions of the body (as identified by the bubbles paradigm) so that the corresponding part of
751 the image formed on the retina falls on the fovea is unexpected. Such a strategy should allow
752 detailed analysis of the shape of these regions. Moreover, in a previous study in which
753 participants were explicitly asked to judge torso shape (indexed by the waist-to-hip ratio), eye-
754 tracking shows that fixations are initially made on one edge of the torso and then the
755 participants' gaze moves across the torso to fixate the other edge (Cornelissen et al., 2009b).
756 They do not make a simple central fixation as is seen here.

757 It is possible that the fixation on the centre of the torso may be serving as a convenient
758 way to locate an image of the torso's left and right edges on the parafoveal region (the region
759 of the retina surrounding the fovea). The parafoveal region supports a less detailed, lower
760 resolution sampling than the fovea, but which is still sufficient to support the detection of the
761 edges of the torso. This perception may be enhanced by the phenomenon of hyperacuity. In
762 this perceptual process, the cortex extrapolates detail from the limited sampling of the
763 parafoveal cone array and so is capable of finer discrimination than the retinal structure would
764 suggest (Gegenfurtner, 2016; Motter & Belky, 1998; Ryu et al., 2013). So even though the
765 centre of the torso is being fixated, information about the relative position of both torso edges
766 can be derived from the periphery of the visual field and an estimate of the body width can be
767 made. After all, just because the observer is fixating in the centre of the torso, that does not
768 mean that her visual attention is focussed at the same position. Numerous studies have
769 suggested that it is possible to direct attention at different parts of the visual field while at the
770 same time fixating a separate part of the image (Evans et al., 2011; Motter & Belky, 1998),
771 although it is unclear whether this allocation of attention across different parts of the visual
772 field is achieved simultaneously or in rapid succession (Evans et al., 2011; Hutterman et al.,
773 2013).

1948
1949
1950 774 Thus, if one accepts that the width of the torso is a good cue to overall body mass, then
1951
1952 775 the most efficient way of sampling the visual information that will allow you to make that
1953
1954 776 judgement may not be to fixate on one edge of the torso and then move the eyes to fixate on
1955
1956 777 the other edge of the torso. Instead, it may be quicker and simpler to foveate within the centre
1957
1958 778 of the torso while directing your attention to the parafoveal regions of the retina corresponding
1959
1960 779 to the edges of the torso. The previously reported difference in the pattern of eye-movements
1961
1962 780 when estimating body size as opposed to judging body shape may be because although the
1963
1964 781 parafovea can support enough spatial resolution to judge the relative position of the left and
1965
1966 782 right torso edges (and so judge width), it may lack sufficient resolution to detect the subtler
1967
1968 783 changes in the outline necessary to judge differences in torso shape (Cornelissen et al., 2009b).
1969
1970
1971

1972 784 This dissociation between the fixation pattern and the visual cues used in self-estimates
1973
1974 785 of body size illustrates the danger of making assumptions based on eye-tracking data. Just
1975
1976 786 because someone appears to look at a certain part of the body, it does not mean they are
1977
1978 787 necessarily directing their visual attention to the same place. The assumption that these two
1979
1980 788 visual activities are the same can lead to a misinterpretation of the data and mean that wrong
1981
1982 789 conclusions are drawn on which body features are key to self-estimates of body size. In future
1983
1984 790 research, it is important that eye-movement studies are paired with other techniques to localise
1985
1986 791 which body features are used in a judgement, to either corroborate or clarify the results of the
1987
1988 792 eye-tracking and avoid the wrong conclusions being made.
1989
1990
1991

1992 793 **5.1. Clinical Implications**

1993
1994

1995 794 Given the dissociation between eye fixation and diagnostic regions we have found in
1996
1997 795 this study of nonclinical women, it is clearly important to make the same measurements in
1998
1999 796 women who have eating disorders. Based on an extensive review of the literature on visual
2000
2001 797 processing in anorexia nervosa, Madsen, Bohon, and Feusner (2013) conclude that women with
2002
2003
2004
2005
2006

2007
2008
2009 798 anorexia nervosa struggle to process global features and tend to over-value local detail.
2010
2011 799 Therefore, one possible outcome of applying the bubbles technique to a body size self-
2012
2013 800 estimation task in anorexia nervosa might be to reveal a very non-specific, or diffuse pattern
2014
2015 801 of diagnostic regions. On each trial, it is possible that participants might lock onto one or a very
2016
2017 802 few bubbles to process only those local details. However, the particular bubble locations that
2018
2019 803 they choose to focus on may be quite different from one trial to the next. When averaged over
2020
2021 804 multiple trials, this could lead to widely dispersed and diffuse diagnostic regions. An alternative
2022
2023 805 possibility might be that, in the face of such over-attention, women with anorexia nervosa may
2024
2025 806 cling to a single well focused diagnostic region, say along just one body edge. If either of these
2026
2027 807 outcomes were true, such findings might suggest new intervention strategies to retrain how
2028
2029 808 sufferers attend to images of their body, thereby helping to prevent body size overestimation.
2030
2031 809 We know that such an outcome could be useful, because recent perceptual training studies have
2032
2033 810 shown clinically meaningful reductions in psychological concerns about body size, shape, and
2034
2035 811 eating that last for up to a month post-intervention (Gledhill et al., 2016; Szostak, 2018).
2036
2037
2038
2039

2040 812 **5.2. Conclusion**

2041
2042 813 In conclusion, the results of these studies using the modified bubbles technique suggest
2043
2044 814 that the key visual cue used when making self-estimates of body size is the width of the torso,
2045
2046 815 as judged from the relative position of the edges of the torso on either side of the body. Previous
2047
2048 816 studies have found that the width of the torso increases with increasing BMI and so this would
2049
2050 817 be a reliable cue to BMI status (e.g., Cornelissen et al. 2009a; Tovée & Cornelissen, 2001;
2051
2052 818 Tovée et al., 1999). In the small bubbles condition, there is an additional important area of the
2053
2054 819 image located at the position corresponding to the gap between the upper thighs. The diameter
2055
2056 820 of the thighs is correlated with overall BMI (Ryan & Niklas, 1999) and so the “thigh gap” is a
2057
2058 821 potential cue to overall adiposity, particularly for lower BMI bodies. The addition of eye-
2059
2060 822 tracking to the paradigm suggests that observers use an efficient fixation strategy when
2061
2062
2063
2064
2065

2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124

823 sampling the cues to body size, fixating centrally within the torso outline to estimate its width
824 and thereby the BMI of the body.

825

826 **Declarations of interest:** There are no conflicts of interest

827

828 **Funding:** The first author was funded by a PhD studentship from Northumbria University.

2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183

References

- 829
830 Anderson, S. J., Mullen, K. T., & Hess, R. F. (1991). Human peripheral spatial resolution for
831 achromatic and chromatic stimuli: Limits imposed by optical and retinal factors.
832 *Journal of Physiology*, 442, 47-64. doi:10.1113/jphysiol.1991.sp018781.
- 833 Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative
834 importance of upper and lower areas of the face. *Journal of Personality and Social
835 Psychology*, 37, 2049–2058. doi:10.1037/0022-3514.37.11.2049
- 836 Beck, A. T., Ward., C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for
837 measuring depression. *Archives of General Psychiatry*, 4, 561-571.
838 doi:10.1001/archpsyc.1961.01710120031004
- 839 Berkman, N. D., Lohr, K. N., & Bulik, C. M. (2007). Outcomes of eating disorders: A
840 systematic review of the literature. *International Journal of Eating Disorders*, 40,
841 293–309. doi:10.1002/eat.20369
- 842 Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window
843 to basic emotions. *Neuropsychologia*, 50, 2830–2838.
844 doi:10.1016/j.neuropsychologia.2012.08.010
- 845 Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial
846 expression perception. *Journal of Experimental Psychology: Human Perception and
847 Performance*, 26, 527–551. doi:10.1037/0096-1523.26.2.527
- 848 Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51, 1484–1525.
849 doi:10.1016/j.visres.2011.04.012

2184
2185
2186 850 Castro, J., Gila, A., Puig, J., Rodriguez, S., & Toro, J. (2004). Predictors of rehospitalization
2187
2188 851 after total weight recovery in adolescents with anorexia nervosa. *International*
2189
2190 852 *Journal of Eating Disorders*, 36, 22–30. doi:10.1002/eat.20009
2191
2192
2193
2194 853 Channon, S., & De Silva, W. (1985). Psychological correlates of weight gain in patients with
2195
2196 854 anorexia nervosa. *Journal of Psychiatric Research*, 19, 267-271. doi:10.1016/B978-0-
2197
2198 855 08-032704-4.50031-0.
2199
2200
2201 856 Cornelissen, K. (2016). *What does it mean to have distorted body image in anorexia*
2202
2203 857 *nervosa?* Doctoral thesis, Northumbria University.
2204
2205 858 <http://nrl.northumbria.ac.uk/30330/>
2206
2207
2208 859 Cornelissen, K. K., Bester, A., Cairns, P., Tovée, M. J., & Cornelissen, P.L. (2015). The
2209
2210 860 influence of personal BMI on body size estimations and sensitivity to body size
2211
2212 861 change in anorexia spectrum disorders. *Body Image*, 13, 75-85.
2213
2214 862 doi:10.1016/j.bodyim.2015.01.001
2215
2216
2217
2218 863 Cornelissen, K. K., Cornelissen, P. L., Hancock, P. J. B., & Tovée, M. J. (2016b). Fixation
2219
2220 864 patterns, not clinical diagnosis, predict body size over-estimation in eating disordered
2221
2222 865 women and healthy controls. *International Journal of Eating Disorders*, 49, 507–518.
2223
2224 866 doi:10.1002/eat.22505
2225
2226
2227 867 Cornelissen, K. K., Gledhill, L. J., Cornelissen, P. L., & Tovée, M. J. (2016a). Visual biases
2228
2229 868 in judging body weight. *British Journal of Health Psychology*, 21, 555-569.
2230
2231 869 doi:10.1111/bjhp.12185
2232
2233
2234 870 Cornelissen, P. L., Cornelissen, K. K., Groves, V., McCarty, K., & Tovée, M. J. (2018).
2235
2236 871 View-dependent accuracy in body mass judgements of female bodies. *Body Image*,
2237
2238 872 24, 116-123. doi:10.1016/j.bodyim.2017.12.007
2239
2240
2241
2242

2243
2244
2245 873 Cornelissen, P. L., Hancock, P. J. B., Kiviniemi, V., George, H. R., & Tovée, M. J. (2009b).
2246
2247 874 Patterns of eye movements when male and female observers judge female
2248
2249
2250 875 attractiveness, body fat and waist-to-hip ratio. *Evolution and Human Behavior*, 30,
2251
2252 876 417–428. doi:10.1016/j.evolhumbehav.2009.04.003.
2253
2254
2255 877 Cornelissen, P. L., Johns, A., & Tovée, M. J. (2013). Body size over-estimation in women
2256
2257 878 with anorexia nervosa is not qualitatively different from female controls. *Body Image*,
2258
2259 879 10, 103-111. doi:10.1016/j.bodyim.2012.09.003.
2260
2261
2262 880 Cornelissen, P. L., Tovée, M. J., & Bateson, M. (2009a). Patterns of subcutaneous fat
2263
2264 881 deposition and the relationship between body mass index and waist-to-hip ratio:
2265
2266 882 Implications for models of physical attractiveness. *Journal of Theoretical Biology*,
2267
2268 883 256, 343-350. doi:10.1016/j.jtbi.2008.09.041
2269
2270
2271 884 Datta, R., & DeYoe, E. A. (2009). I know where you are secretly attending! The topography
2272
2273 885 of human visual attention revealed with fMRI. *Vision Research*, 49, 1037–1044.
2274
2275
2276 886 doi:10.1016/j.visres.2009.01.014
2277
2278
2279 887 Ehinger, K. A., & Rosenholts, R. (2016). A general account of peripheral encoding also
2280
2281 888 predicts scene perception performance. *Journal of Vision*, 16, 1-19.
2282
2283 889 doi:10.1167/16.2.13
2284
2285
2286 890 Evans, C., & Dolan, B. (1993). Body Shape Questionnaire: Derivation of shortened “alternate
2287
2288 891 forms.” *International Journal of Eating Disorders*, 13, 315-321. doi:10.1002/1098-
2289
2290 892 108x(199304)13:3<315::aid-eat2260130310>3.0.co;2-3
2291
2292
2293 893 Evans, K. K., Horowitz, T. S., Howe, P., Pedersini, R., Reijnen, E., & Pinto, Y., et al. (2011).
2294
2295 894 Visual attention. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 503–514.
2296
2297 895 doi:10.1002/wcs.127
2298
2299
2300
2301

2302
2303
2304 896 Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorder psychopathology:
2305
2306 897 Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*,
2307
2308 898 363-370. doi:10.1002/1098-108X(199412)16:4<363::AID-
2309
2310 EAT2260160405>3.0.CO;2-#
2311 899
2312
2313
2314 900 Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating
2315
2316 901 disorders: A “transdiagnostic” theory and treatment. *Behaviour Research and*
2317
2318 902 *Therapy*, *41*, 509–528. doi:10.1016/s0005-7967(02)00088-8
2319
2320
2321 903 Gardner, R. M., & Bokenkamp, E. D. (1996). The role of sensory and nonsensory factors in
2322
2323 904 body size estimations of eating disorder subjects. *Journal of Clinical Psychology*, *52*,
2324
2325 905 3-15. doi:10.1002/(sici)1097-4679(199601)52:1<3::aid-jclp1>3.0.co;2-x
2326
2327
2328 906 Gegenfurtner, K. R. (2016). The interaction between vision and eye movements. *Perception*,
2329
2330 907 *45*, 1333–1357. doi:10.1177/0301006616657097
2331
2332
2333 908 George, H. R., Cornelissen, P. L., Hancock, P. J. B., Kiviniemi, V. V., & Tovée, M. J. (2011).
2334
2335 909 Differences in eye-movement patterns between anorexic and control observers when
2336
2337 910 judging body size and attractiveness. *British Journal of Psychology*, *102*, 340-354.
2338
2339 911 doi:10.1348/000712610X524291.
2340
2341
2342
2343 912 Gledhill, L. J., Cornelissen, K. K., Cornelissen, P. L., Penton-Voak, I. S., Munafò, M. R., &
2344
2345 913 Tovée, M. J. (2016). An interactive training programme to treat body image
2346
2347 914 disturbance. *British Journal of Health Psychology*, *22*, 60–76.
2348
2349 915 doi:10.1111/bjhp.12217
2350
2351
2352 916 Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information
2353
2354 917 in recognition tasks. *Vision Research*, *41*, 2261–2271. doi:10.1016/s0042-
2355
2356 918 6989(01)00097-9
2357
2358
2359
2360

2361
2362
2363 919 Gosselin, F., & Schyns, P. G. (2004). No troubles with bubbles: A reply to Murray and Gold.
2364
2365 920 *Vision Research*, 44, 471–477. doi:10.1016/j.visres.2003.10.007
2366
2367
2368
2369 921 Herzog, D. B., Dorer, D. J., Keel, P. K., Selwyn, S. E., Ekeblad, E. R., & Flores, A. T., . . .
2370
2371 922 Keller, M. B. (1999). Recovery and relapse in anorexia nervosa and bulimia nervosa:
2372
2373 923 A 7.5-year follow-up study. *Journal of the American Academy of Childhood and*
2374
2375 924 *Adolescent Psychiatry*, 38, 829-837. doi:10.1097/00004583-199907000-00012
2376
2377
2378 925 Hüttermann, S., Memmert, D., Simons, D. J., & Bock, O. (2013). Fixation strategy influences
2379
2380 926 the ability to focus attention on two spatially separate objects. *PLoS ONE*, 8, e65673.
2381
2382 927 doi:10.1371/journal.pone.0065673
2383
2384
2385
2386 928 Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural
2387
2388 929 diversity in expectations of facial expressions of emotion. *Journal of Experimental*
2389
2390 930 *Psychology: General*, 141, 19–25. doi:10.1037/a0023463
2391
2392
2393 931 Junne, F., Wild, B., Resmark, G., Giel, K. E., Teufel, M., & Martus, . . . Zipfel, S. (2019).
2394
2395 932 The importance of body image disturbances for the outcome of outpatient
2396
2397 933 psychotherapy in patients with anorexia nervosa: Results of the ANTOP-study.
2398
2399 934 *European Eating Disorders Review*, 27, 49-58. doi:10.1002/erv.2623
2400
2401
2402
2403 935 Kadel R. P., & Kip K. E. (2012, October). A SAS macro to compute effect size (Cohen’s *d*)
2404
2405 936 and its confidence interval from raw survey data. *Proceedings of the Annual*
2406
2407 937 *Southeastern SAS Users Group Conference*, Durham, NC.
2408
2409
2410 938 Keel, P. K., Dorer, D. J., Franko, D. L., Jackson, S. C., & Herzog, D. B. (2005).
2411
2412 939 Postremission predictors of relapse in women with eating disorders. *The American*
2413
2414 940 *Journal of Psychiatry*, 162, 2263-2268. doi:10.1176/appi.ajp.162.12.2263
2415
2416
2417
2418
2419

2420
2421
2422 941 Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Verner acuity, crowding and cortical
2423
2424 942 magnification. *Vision Research*, 25, 963-977. doi:10.1016/0042-6989(85)90207-X
2425
2426
2427 943 Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberber, O. (2006).
2428
2429 944 *SAS for mixed models* (2nd ed.). Cary, NC: SAS.
2430
2431
2432 945 Madsen, S. K., Bohon, C., & Feusner, J. D. (2013) Visual processing in anorexia nervosa and
2433
2434 946 body dysmorphic disorder: Similarities, differences, and future research directions.
2435
2436 947 *Journal of Psychiatric Research*, 47, 1483-1491.
2437
2438 948 doi:10.1016/j.jpsychires.2013.06.003
2439
2440
2441 949 Miller, J. M., & Bockisch, C. (1997). Where are the things we see? *Nature*, 386, 550-551.
2442
2443 950 doi:10.1038/386550a0
2444
2445
2446 951 Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). Eating Disorder Examination
2447
2448 952 Questionnaire (EDE-Q): Norms for young adult women. *Behaviour Research and*
2449
2450 953 *Therapy*, 44, 53–62. doi:10.1016/j.brat.2004.12.003
2451
2452
2453 954 Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual
2454
2455 955 search. *Vision Research*, 38, 1805–1815. doi:10.1016/s0042-6989(97)00349-0
2456
2457
2458 956 Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research*, 44, 461-470.
2459
2460 957 doi:10.1016/j.visres.2003.10.006
2461
2462
2463 958 Pelli, D., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature*
2464
2465 959 *Neuroscience*, 11, 1129-1135. doi:10.1038/nn.2187
2466
2467
2468 960 Pike, K. M. (1998). Long-term course of anorexia nervosa. *Clinical Psychology Review*, 18,
2469
2470 961 447–475. doi:10.1016/s0272-7358(98)00014-2
2471
2472
2473
2474
2475
2476
2477
2478

2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537

962 Probst, M., Vandereycken, W., Van Coppenolle, H., & Pieters, G. (1998). Body size
963 estimation in anorexia nervosa patients: The significance of overestimation. *Journal*
964 *of Psychosomatic Research, 44*, 451-456. doi:10.1016/s0022-3999(97)00270-5

965 Rilling, J. K., Kaufman T. L., Smith, E. O., Patel, R., & Worthman, C. M. (2009). Abdominal
966 depth and waist circumference as influential determinants of human female
967 attractiveness. *Evolution and Human Behavior, 30*, 21–31.
968 doi:10.1016/j.evolhumbehav.2008.08.007

969 Ryan, A. S., & Nicklas, B. J. (1999). Age-related changes in fat deposition in mid-thigh
970 muscle in women: Relationships with metabolic cardiovascular disease risk factors.
971 *International Journal of Obesity, 23*, 126-132. doi:10.1038/sj.ijo.0800777

972 Ryu, D., Abernethy, B., Mann, D. L., Poolton, J. M., & Gorman, A. D. (2013). The role of
973 central and peripheral vision in expert decision making. *Perception, 42*, 591-607.
974 doi:10.1068/p7487

975 Slade, P., & Russell, G. (1973). Awareness of body dimensions in anorexia nervosa: Cross-
976 sectional and longitudinal studies. *Psychological Medicine, 3*, 188-199.
977 doi:10.1017/s0033291700048510

978 Smith, K. L., Tovée, M. J., Hancock, P. J., Bateson, M., Cox, M. A. A., & Cornelissen, P. L.
979 (2007). An analysis of body shape attractiveness based on image statistics: Evidence
980 for a dissociation between expressions of preference and shape discrimination. *Visual*
981 *Cognition, 15*, 927-953. doi:10.1080/13506280601029515

982 Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and
983 decoding facial expressions. *Psychological Science, 16*, 184–189. doi:10.1111/j.0956-
984 7976.2005.00801.x

2538
2539
2540 985 Szostak, N. M. (2018). Negative body image and cognitive biases to body size. (Doctoral
2541
2542 986 dissertation). Retrieved from: <https://hydra.hull.ac.uk/resources/hull:16444>
2543
2544
2545 987 Tovée, M. J., Benson, P. J., Emery, J. L., Mason, S. M., & Cohen-Tovée, E. M. (2010).
2546
2547 988 Measurement of body size and shape perception in eating-disordered and control
2548
2549 989 observers using body-shape software. *British Journal of Psychology*, *94*, 501-516.
2550
2551 990 doi:10.1348/000712603322503060
2552
2553
2554
2555 991 Tovée, M. J., & Cornelissen, P. L. (1999). The mystery of female beauty. *Nature*, *399*, 215–
2556
2557 992 216. doi:10.1038/20345
2558
2559
2560 993 Tovée, M. J., & Cornelissen, P. L. (2001). Female and male perceptions of female physical
2561
2562 994 attractiveness in front-view and profile. *British Journal of Psychology*, *92*, 391–402.
2563
2564 995 doi:10.1348/000712601162257
2565
2566
2567 996 Tovée, M. J., Hancock, P. J. B., Mahmoodi, S., Singleton, B. R. R., & Cornelissen, P. L.
2568
2569 997 (2002). Human female attractiveness: Waveform analysis of body shape. *Proceedings*
2570
2571 998 *of the Royal Society of London B: Biological Sciences*, *269*, 2205-2213.
2572
2573 999 doi:10.1098/rspb.2002.2133
2574
2575
2576
2577 1000 World Health Organisation (2003). Shaping the future report. Retrieved from
2578
2579 1001 <http://www.who.int/whr/2003/en/>
2580
2581
2582 1002 Yu, H., Garrod, O. G., & Schyns, P. G. (2012). Perception-driven facial expression synthesis.
2583
2584 1003 *Computers & Graphics*, *36*, 152–162. doi:10.1016/j.cag.2011.12.002
2585
2586
2587 1004
2588
2589
2590 1005
2591
2592
2593 1006
2594
2595
2596

2597
2598
2599 **1007 Figure Legends**
2600
2601

2602 **1008** *Figure 1.* Screenshots of the stimuli from two consecutive trials from: (a) Experiment 1 with
2603
2604 **1009** large bubbles, and (b) Experiment 2 with small bubbles. The first two columns show the stimuli
2605
2606 **1010** as presented to the participant. Columns three and four show the same images but with a red
2607
2608 **1011** outline to indicate the outline of the female model in the stimulus, beneath the gray overlay.
2609
2610 **1012** On every trial, participants are given a partial view of the female model through a set of so-
2611
2612 **1013** called Gaussian bubbles. These are circular holes with blurred edges that perforate the gray
2613
2614 **1014** overlay that covers the model in the stimulus image. Please note that much visual detail will be
2615
2616 **1015** lost in this illustration, compared to the original stimuli displayed on a PC monitor.
2617
2618
2619

2620 **1016**
2621
2622 **1017** *Figure 2.* Diagnostic images for (a) the big bubbles mask Experiment 1, top row, and (b) the
2623
2624 **1018** small bubbles mask Experiment 2, bottom row. For the Accurate and Overestimate figures (left
2625
2626 **1019** and middle columns), the white circles show the locations of bubbles where correct response
2627
2628 **1020** rates were significantly above the 75% criterion based on the GLMMs. The heat maps represent
2629
2630 **1021** the averaged and smoothed raw data that contributed to the GLMMs. For the Accurate –
2631
2632 **1022** Overestimate figure (right column), the white circles show where the differences between the
2633
2634 **1023** two groups of observers are significantly different from zero. The blue-cyan colours in the heat
2635
2636 **1024** map show where over-estimators made more correct responses than accurate estimators. The
2637
2638 **1025** red-yellow colours in the heat map show where accurate estimators made more correct
2639
2640 **1026** responses than over-estimators.
2641
2642
2643

2644 **1027**
2645
2646 **1028** Figure 3: Fixation density maps for accurate and overestimators across the four eye-tracking
2647
2648 **1029** conditions. Each image represents the same stimulus model with a semi-transparent coloured
2649
2650 **1030** overlay to indicate fixation density, reported in z -scores. The higher the z -score (from gray,
2651
2652
2653
2654
2655

2656
2657
2658 1031 through green and yellow to red), the more time participants spent looking at a particular region
2659
2660 1032 on the body. Black contours represent $3SDs$, within which most fixations lie.
2661
2662
2663 1033
2664
2665
2666 1034 Figure 4: Shows predicted marginal means together with their 95%CIs, for co-registered eye
2667
2668 1035 fixation and psychophysical data, across a set of horizontal slices. The locations on the model's
2669
2670 1036 body of the horizontal slices through the combined datasets are indicated by letter groups: *A*,
2671
2672 1037 *B*, & *C* and *D*, *E*, & *F*, for the large and small bubble mask datasets, respectively. For accurate
2673
2674 1038 estimators, solid black lines with black circles represent the eye-movement data and solid white
2675
2676 1039 lines with white circles the psychophysical data. For overestimators, solid black lines with
2677
2678 1040 black triangles represent the eye-movement data and solid white lines with white triangles the
2679
2680 1041 psychophysical data. There is a small horizontal offset in the x-axes between the eye-movement
2681
2682 1042 and psychophysical data, so that error bars do not overlap.
2683
2684
2685
2686 1043
2687
2688
2689 1044
2690
2691
2692 1045
2693
2694 1046
2695
2696
2697 1047
2698
2699
2700 1048
2701
2702
2703 1049
2704
2705 1050
2706
2707
2708 1051
2709
2710
2711 1052
2712
2713
2714

2715
 2716
 2717
 2718
 2719
 2720
 2721
 2722
 2723
 2724
 2725
 2726
 2727
 2728
 2729
 2730
 2731
 2732
 2733
 2734
 2735
 2736
 2737
 2738
 2739
 2740
 2741
 2742
 2743
 2744
 2745
 2746
 2747
 2748
 2749
 2750
 2751
 2752
 2753
 2754
 2755
 2756
 2757
 2758
 2759
 2760
 2761
 2762
 2763
 2764
 2765
 2766
 2767
 2768
 2769
 2770
 2771
 2772
 2773

Table 1. Experiment 1 with large bubble masks: Participant characteristics

	Accurate (<i>n</i> =12)		Overestimate (<i>n</i> =12)		Accurate vs. Overestimate		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>	<i>d</i>	95% <i>CI</i>
Participant characteristics							
Age (years)	23.67	5.65	22.25	4.37	.99	-0.28	(-1.12 – 0.56)
BMI (kg/m ²)	21.97	2.89	22.16	3.22	1.00	0.06	(-0.77 – 0.90)
Depression							
BDI score	15.17	9.84	17.75	11.28	.99	0.24	(-0.60 – 1.09)
Body shape and eating concerns							
BSQ-16 score	38.92	20.78	47.67	23.03	.93	0.40	(-0.45 – 1.25)
EDE-Q global score	1.32	1.02	2.23	1.57	.50	0.69	(-0.48 – 1.55)
EDE-Q res score	1.40	1.37	2.03	1.32	.86	0.47	(-0.38 – 1.32)
EDE-Q eat score	0.45	0.52	1.28	1.42	.35	0.78	(-0.09 – 1.65)
EDE-Q wc score	1.58	1.45	2.47	1.84	.76	0.53	(-0.32 – 1.39)
EDE-Q sc score	1.86	1.67	3.14	2.17	.55	0.66	(-0.20 – 1.52)
Psychophysical performance							
PSE (kg/m ²)	22.16	2.98	25.85	3.40	.07	1.15	(0.25 – 2.06)
DL (kg/m ²)	0.67	0.26	1.15	0.88	.43	0.73	(-0.14 – 1.59)
Overestimation (PSE - BMI)	0.19	0.78	3.69	1.31	< .001	3.25	(1.98 – 4.53)
Mean bubble count	5.00	1.08	5.12	1.21	.92	0.15	(-0.69 – 0.99)
Mean percentage trials correct	74.41	0.86	74.33	1.10	.97	-0.08	(-0.92 – 0.76)

Note. BDI = Beck Depression Inventory; BSQ-16 = Body Shape Questionnaire; EDE-Q = Eating Disorders Examination Questionnaire global score; EDE-Q subscales: res = restraint; eat = eating concerns; wc = weight concerns; sc = shape concerns.

2774
 2775
 2776
 2777
 2778
 2779
 2780
 2781
 2782
 2783
 2784
 2785
 2786
 2787
 2788
 2789
 2790
 2791
 2792
 2793
 2794
 2795
 2796
 2797
 2798
 2799
 2800
 2801
 2802
 2803
 2804
 2805
 2806
 2807
 2808
 2809
 2810
 2811
 2812
 2813
 2814
 2815
 2816
 2817
 2818
 2819
 2820
 2821
 2822
 2823
 2824
 2825
 2826
 2827
 2828
 2829
 2830
 2831
 2832

1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072

Table 2. Experiment 2 with small bubble masks: Participant characteristics

	Accurate (<i>n</i> = 12)		Overestimate (<i>n</i> = 12)		Accurate vs. Overestimate		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>	<i>d</i>	95% <i>CI</i>
Participant characteristics							
Age (years)	22.58	6.40	20.92	3.75	.98	-0.32	(-1.16 – 0.53)
BMI (kg/m ²)	22.55	4.80	21.99	2.94	0.99	-0.14	(-0.98 – 0.70)
Depression							
BDI score	9.50	8.60	16.83	11.34	.45	0.73	(-0.14 – 1.59)
Body shape and eating concerns							
BSQ-16 score	45.67	23.91	53.17	18.41	.97	0.35	(-0.49 – 1.20)
EDE-Q global score	1.74	1.66	2.65	1.29	.64	0.62	(-0.24 – 1.48)
EDE-Q res score	1.35	1.43	2.22	1.43	.65	0.61	(-0.25 – 1.46)
EDE-Q eat score	1.08	1.48	1.63	1.40	.95	0.38	(-0.46 – 1.23)
EDE-Q wc score	2.08	2.09	2.97	1.63	.85	0.47	(-0.38 – 1.32)
EDE-Q sc score	2.43	1.92	3.79	1.39	.33	0.81	(-0.06 – 1.68)
Psychophysical performance							
PSE (kg/m ²)	22.40	4.68	25.96	2.72	.20	0.93	(0.05 – 1.81)
DL (kg/m ²)	0.75	0.30	1.03	0.23	.11	1.04	(0.15 – 1.94)
Overestimation (PSE – BMI)	-0.15	0.57	3.97	1.35	<.001	3.98	(2.53 – 5.42)
Mean bubble count	18.42	5.51	19.62	3.86	.74	0.25	(-0.59 – 1.09)
Mean percentage trials correct	74.75	1.46	74.03	1.05	.61	-0.33	(-1.17 – 0.51)

Note. BDI = Beck Depression Inventory; BSQ-16 = Body Shape Questionnaire; EDE-Q = Eating Disorders Examination Questionnaire global score; EDE-Q subscales: res = restraint; eat = eating concerns; wc = weight concerns; sc = shape concerns.

1073 Table 3. Experiment 3: Participant characteristics

	Accurate (<i>n</i> = 12)		Overestimate (<i>n</i> = 12)		Accurate vs. Overestimate		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>	<i>d</i>	<i>95%CI</i>
Participant characteristics							
Age (years)	20.58	1.98	23.33	5.35	.55	0.68	(-0.18 – 1.54)
BMI (kg/m ²)	23.42	6.38	21.73	3.57	.98	-0.33	(-1.17 – 0.52)
Depression							
BDI score	12.50	8.12	15.17	8.12	.98	0.33	(-0.52 – 1.17)
Body shape and eating concerns							
BSQ-16 score	38.92	16.81	46.33	12.61	.83	0.50	(-0.35 – 1.35)
EDE-Q global score	1.38	0.97	2.15	1.24	.53	0.69	(-0.35 – 1.35)
EDE-Q res score	1.15	0.92	2.37	1.89	.33	0.82	(-0.05 – 1.69)
EDE-Q eat score	0.65	0.69	1.12	0.96	.75	0.56	(-0.29 – 1.41)
EDE-Q wc score	1.55	1.48	2.32	1.11	.71	0.58	(-0.27 – 1.44)
EDE-Q sc score	2.16	1.55	2.79	1.35	.91	0.44	(-0.41 – 1.28)
Psychophysical performance							
PSE (kg/m ²)	23.45	6.15	25.35	3.27	.95	0.39	(-0.46 – 1.23)
DL (kg/m ²)	0.78	0.40	0.92	0.65	.99	0.25	(-0.59 – 1.09)
Overestimation (PSE – BMI)	0.03	0.64	3.62	2.03	< .001	2.39	(1.29 – 3.49)

Note. BDI = Beck Depression Inventory; BSQ-16 = Body Shape Questionnaire; EDE-Q = Eating Disorders Examination Questionnaire global score; EDE-Q subscales: res = restraint; eat = eating concerns; wc = weight concerns; sc = shape concerns.

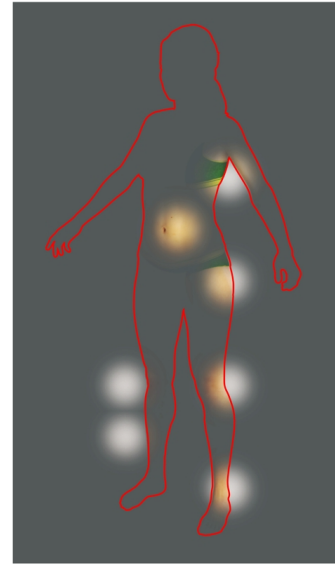
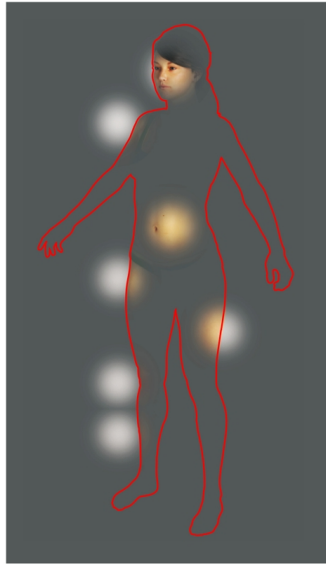
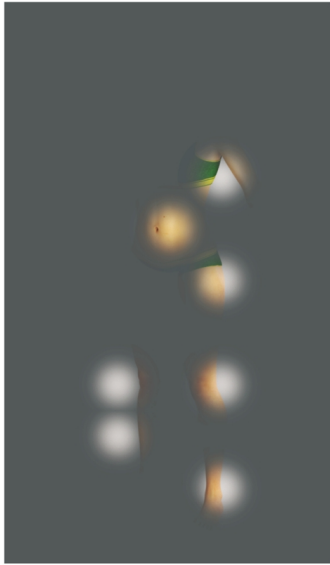
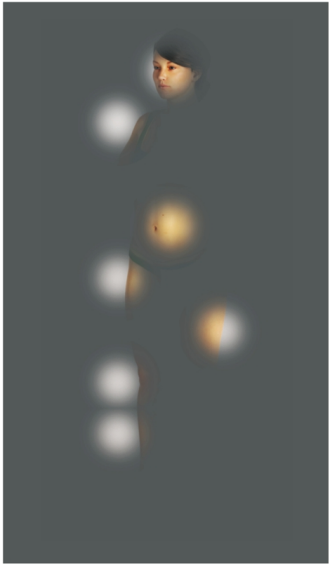
2892
 2893
 2894
 2895
 2896
 2897
 2898
 2899
 2900
 2901
 2902
 2903
 2904
 2905
 2906
 2907
 2908
 2909
 2910
 2911
 2912
 2913
 2914
 2915
 2916
 2917
 2918
 2919
 2920
 2921
 2922
 2923
 2924
 2925
 2926
 2927
 2928
 2929
 2930
 2931
 2932
 2933
 2934
 2935
 2936
 2937
 2938
 2939
 2940
 2941
 2942
 2943
 2944
 2945
 2946
 2947
 2948
 2949
 2950

Table 4. Comparison of fixation density in each of the three columns.

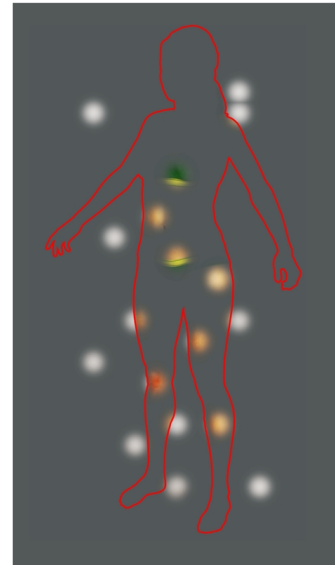
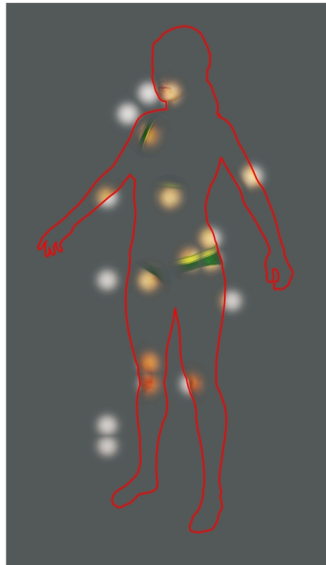
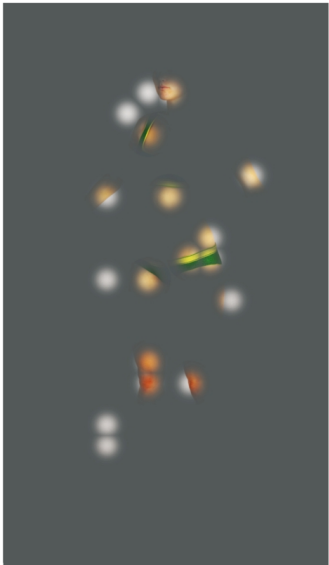
Group	Bubble Size	Task	Left column (%)		Middle column (%)		Right column (%)		Left vs. Middle	Right vs. Middle
			<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>p</i>	<i>p</i>
Accurate	Big	Mask	32.41	4.66	56.98	3.74	10.60	2.26	< .001	< .001
		No mask	20.56	5.99	75.21	6.18	5.63	3.20	< .001	< .001
	Small	Mask	41.17	5.19	50.75	3.58	8.82	2.39	.08	< .001
		No mask	17.89	4.74	77.45	4.66	5.08	2.14	< .001	< .001
Over-estimate	Big	Mask	35.09	5.26	52.80	3.44	12.11	3.48	< .001	.003
		No mask	25.53	7.57	70.98	6.99	4.19	2.53	< .001	< .001
	Small	Mask	39.64	4.86	50.82	3.19	9.55	3.29	.05	< .001
		No mask	31.34	9.76	63.18	8.71	6.58	2.92	< .001	< .001

1084

a)



b)



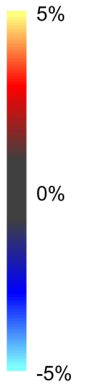
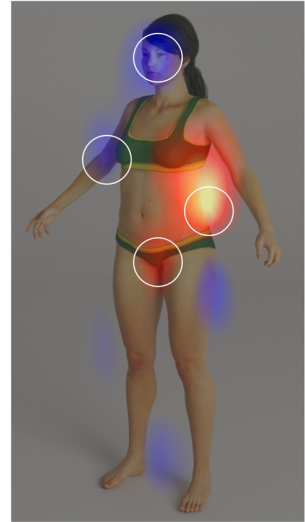
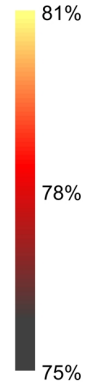
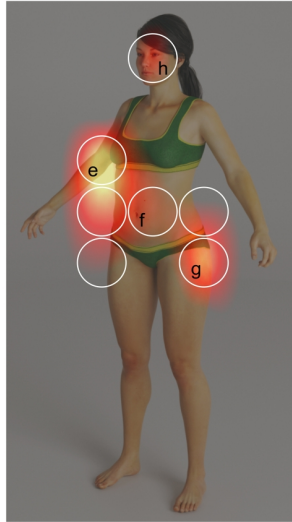
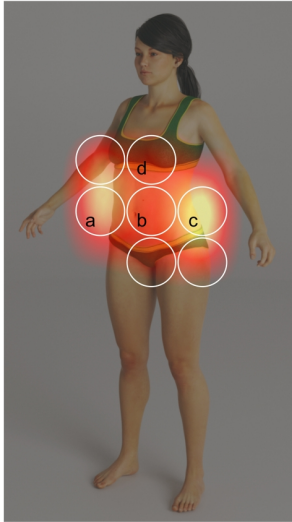
a)

Accurate

Over-estimate

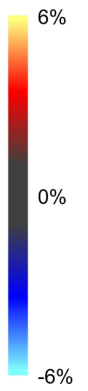
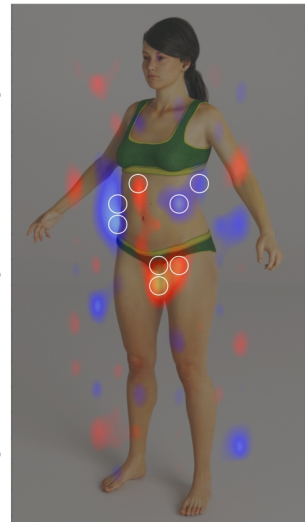
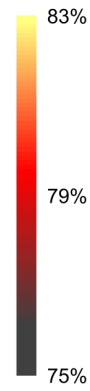
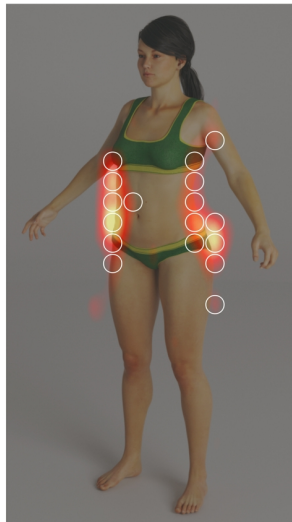
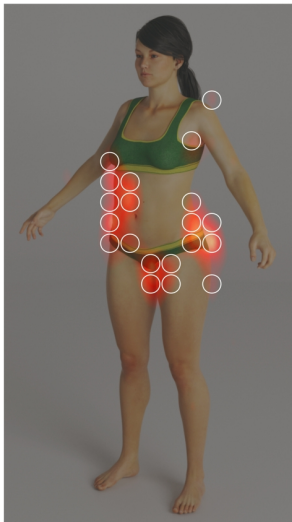
Accurate - Over-estimate

Big Bubbles Mask



b)

Small Bubbles Mask



Accurate Estimators

Over-estimators

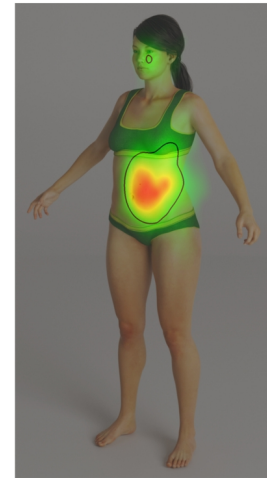
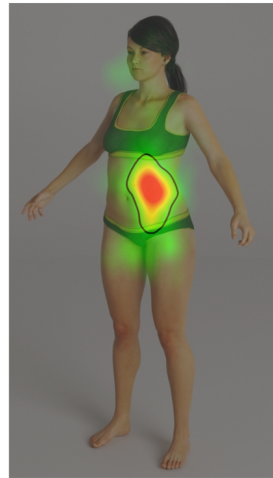
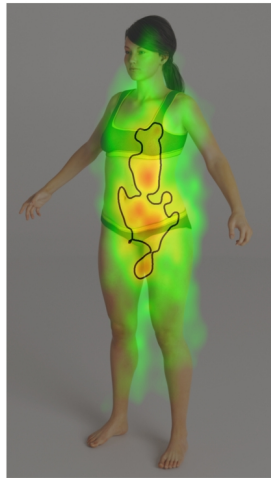
Masking Task

All Mask Bubbles Open

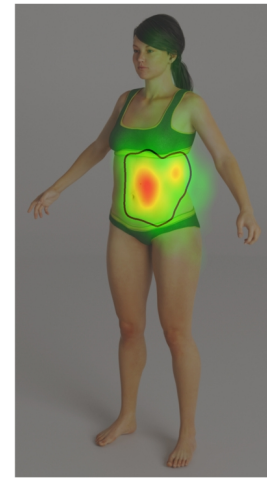
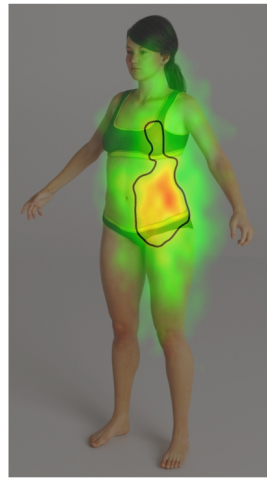
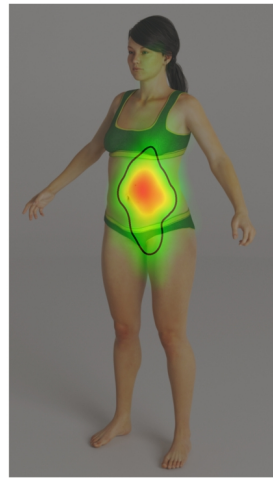
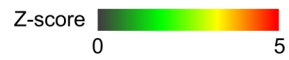
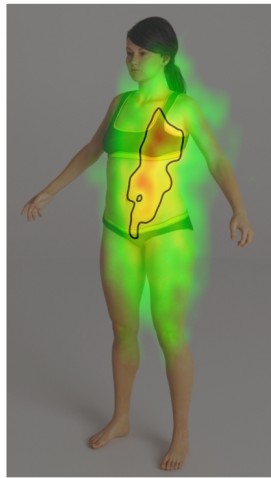
Masking Task

All Mask Bubbles Open

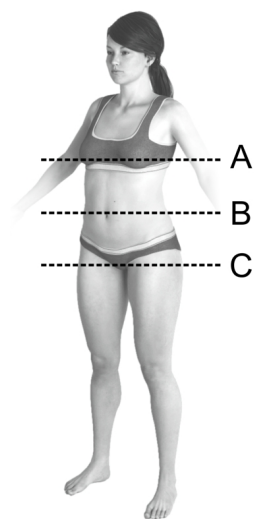
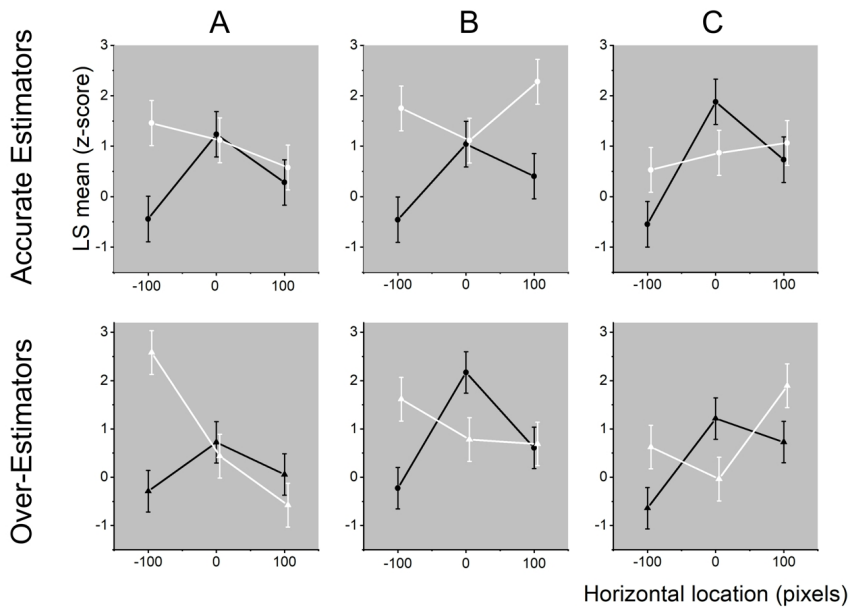
Big
Bubbles



Small
Bubbles



Big Bubbles Mask



Small Bubbles Mask

