

1 **USING COMPOSITE IMAGES TO ASSESS ACCURACY IN PERSONALITY**
2 **ATTRIBUTION TO FACES**

3

4

Abstract

5 Several studies have demonstrated some accuracy in personality attribution
6 using only visual appearance. Using composite images of those scoring high
7 and low on a particular trait, the current study shows that judges perform
8 better than chance in guessing others' personality, particularly for the traits
9 conscientiousness and extraversion. This study also shows that
10 attractiveness, masculinity, and age may all provide cues to accurately assess
11 personality and that accuracy is affected by the sex of both of those judging
12 and being judged. Individuals do perform better than chance at guessing
13 another's personality from only facial information, providing some support for
14 the popular belief that it is possible to accurately assess personality from
15 faces. However, this accuracy is somewhat limited.

16

17 **Key words:** social perception, facial appearance, accuracy, zero-
18 acquaintance, stereotype

19

Introduction

20 Judging attractiveness of human faces takes only a moment, and we also
21 classify faces for broad and tangible qualities like age and sex. Alongside
22 these attributions we also examine more subtle social signals predicting the
23 behaviour and personality of others, such as deciding whether we think
24 someone is an extravert or an introvert, based on their appearance. Facial
25 characteristics influence attributions of various personality characteristics and,
26 because of their prominent and (in most cases) permanent display), can play
27 an important role in social perception.

28 Many individuals believe the face provides important guides to
29 character (Hassin & Trope 2000, Liggett 1974) and there are also studies
30 showing that observers can make reliable and somewhat accurate
31 judgements of others' personality traits on the basis of very little information.
32 Several studies have examined accuracy of personality attributions and many
33 utilise the five factor model of personality (or the Big 5) proposed by Norman
34 (1963). The factors are extraversion, agreeableness, conscientiousness,
35 neuroticism, and intellect-openness. Passini and Norman (1966) examined
36 small groups of undergraduates who were placed in groups without verbal
37 interaction for 15 minutes and asked to rate each other using scales
38 corresponding to the 'Big 5' personality factors. They found that correlations
39 between self and others' ratings were significantly greater than chance for
40 extraversion, conscientiousness and openness.

41 Replicating this study, Albright, Kenny, & Malloy (1988) also found that
42 when judges were asked to rate strangers they met in person without
43 interaction on personality factors, there was a high degree of agreement

44 between different judges on the personality characteristics attributed. The
45 judgements were also significantly correlated with the targets' own self-ratings
46 for extraversion and conscientiousness. Watson (Watson 1989) also found
47 evidence for accuracy when judging extraversion and conscientiousness. This
48 paradigm was referred to as "zero acquaintance" and there are now many
49 studies which reinforce the original findings (see Kenny, Albright, Malloy, &
50 Kashy 1994 for review). The phenomena of consensus and accuracy in
51 personality attributions from faces have also been identified in cross-cultural
52 studies. They can be found using photographs of still faces (Albright et al.
53 1988), video footage (Kenny, Horner, Kashy, & Chu 1992), and also using
54 acquaintances' judgements of targets' personality in comparison with the
55 unfamiliar judges' estimations (Borkenau & Liebler 1993). Amongst these
56 studies there have sometimes been indications of sex differences in accuracy.
57 For example, Ambady, Hallahan, and Rosenthal (1995) report that women are
58 more accurate judges of strangers' personality than men.

59 Accuracy in rating has also been documented for traits not related to
60 the Big 5. Berry and Brownlow (1989) found that unfamiliar judge's ratings of
61 male babyfacedness (possession of infant like facial traits) were positively
62 correlated with the face owner's self-reported approachability and warmth, but
63 negatively related to self-reported aggression. For female faces, babyishness
64 was associated with low self-reported levels of physical power and
65 assertiveness. Bond, Berry, and Omar (1994) have demonstrated that
66 individuals with faces rated as having low honesty are more likely to volunteer
67 for experiments that involve them deceiving others than people whose faces
68 are judged to look more honest. There is also evidence that intelligence can

69 be inferred from facial information (Zebrowitz, Hall, Murphy, & Rhodes 2002)
70 and that personality can also be manifested in the environments that people
71 construct around themselves, in that judges can accurately infer some
72 personality traits from brief viewing of targets' bedrooms and offices (Gosling,
73 Ko, Mannarelli, & Morris 2002).

74 The consistency in attributions must be due to certain visible
75 characteristics in the perceived. Three likely candidates which have received
76 much attention in stereotype research are masculinity, attractiveness, and
77 age. Males and females differ in facial form, and certain behavioural traits
78 such as dominance-submissiveness are thought to be associated with one
79 sex more than the other (it is essentially immaterial to the issue of consistency
80 of attributions whether such stereotypes are actually accurate, although of
81 course, this would be relevant to attribution accuracy). By extrapolation,
82 observers may perceive the differences in the masculinity of faces within
83 members of the same sex as relating to the dominance of the owner of that
84 face (Perrett et al. 1998). As well as potential sex stereotypes, other general
85 stereotypes also exist. For example, there exists a pervasive "what is beautiful
86 is good" stereotype (Dion, Berscheid, & Walster 1972), in which varied
87 positive personality attributions are projected on to those possessing attractive
88 faces (e.g., Feingold 1992). There also exists a "baby-face" stereotype (Berry
89 & McArthur 1986) whereby individuals whose faces most resemble infants are
90 seen as warmer, less likely to exhibit antisocial behaviour, more submissive,
91 more naive, and more irresponsible than those with more mature faces
92 (Zebrowitz & Montepare 1992). This may reflect attribution based on similarity
93 to a particular group, and since immaturity is associated with childhood,

94 childlike faces are perceived as immature (Berry & McArthur 1985). While
95 baby-facedness may not be the same as perceived age, infant-like faces do
96 appear younger than more mature looking faces. Given their prominent role in
97 social perception, any of these traits may provide cues to accurate personality
98 attributions. Accuracy could potentially be mediated by self-fulfilling
99 prophecies (Snyder, Tanke, & Berscheid 1977), the expressive habits of
100 individuals (Malatesta, Fiore, & Messina 1987), active manipulation, such as
101 use of grooming aids (Cash 1990), or putative links to biological mechanisms,
102 such as those between face shape, personality and hormone levels (Enlow
103 1982, Mazur & Booth 1998).

104 In this study we created composite images of individuals who had rated
105 themselves as high or low on each of the five-factor traits. We had the
106 resulting images rated for the same traits so that we could assess accuracy
107 and determine whether there were consistent facial cues to accurate
108 personality attribution. Galton (1878) devised the basic technique of
109 combining individual images to produce composites. Galton was also
110 interested in how behaviour may be reflected in faces and he produced,
111 amongst other images, a composite image of criminals. Composite creation
112 techniques have been developed in recent years, yielding ever more realistic
113 looking composites (Benson & Perrett 1993, Tiddeman, Burt, & Perrett 2001).
114 Characteristics common to the individual faces combined in composites are
115 maintained and highlighted, while idiosyncratic variations that are not common
116 to the set are 'averaged out'. Therefore, if individuals high or low on a
117 particular trait have similar facial appearance, the facial characteristics they

118 have in common should be maintained in composites, while characteristics
119 they do not share will disappear.

120

121

Methods

A). Making average faces based on self-rated personality

Participants

124 68 males (aged 18-24, $M = 20.1$, $SD = 1.4$) and 123 females (aged 17-32, $M =$
125 20.7 , $SD = 2.5$) participated in this part of the study.

126

Materials

128 A 40-item questionnaire was administered that was developed from trait pairs
129 presented in McCrae and Costa (1987). McCrae and Costa (1985) present an
130 80-item questionnaire, and McCrae and Costa (1987) present the five-factor
131 loadings with varimax rotation for 738 raters judging one of their peers for
132 these 80 adjective pairs. To reduce this questionnaire to the most valid 40
133 items, the 8 highest loading questions from each factor were taken. The
134 questionnaire was presented via a computer with participants using a mouse
135 to click the point on a 7-point scale between the 40 adjective pairs. The forty
136 pairs can be seen in the Appendix.

137

Photography

139 Each participant was photographed to provide a full-face colour image.
140 Photographs were taken with a digital camera (resolution set at 1200x1000
141 pixels) under standardised diffuse lighting conditions and against a constant
142 background. Participants were asked to pose with a neutral facial expression

143 and were asked to pull their hair back from their face. Participants were also
144 asked to any remove spectacles and males were clean shaven in
145 appearance.

146

147 **Factor analysis of personality questionnaire**

148 Factor analysis extracting 5 factors and using varimax rotation was carried out
149 separately for males and females. The five factors accounted for 48.9% (1=
150 12.3, 2= 10.0, 3= 9.4, 4= 8.8, 5= 8.5) of the variance of the original scores in
151 females and 52.9% (1= 15.7, 2= 10.7, 3= 10.5, 4= 9.3, 5= 6.6) of the variance
152 of scores in males. For females, factor 1 was labelled extraversion, factor 2
153 conscientiousness, factor 3 neuroticism, factor 4, agreeableness, and factor 5
154 openness to experience. For males, the factors were labelled similarly apart
155 from factor 4 which was labelled openness to experience and factor 5 which
156 was labelled agreeableness. The factor loadings for the adjective pairs can be
157 seen in the Appendix. As can be seen in the Appendix, the 40-item
158 questionnaire appears to capture the big 5 factors for both males and females.
159

160 **Making the composite faces**

161 From the factor analysed personality scores the 15 highest and lowest scorers
162 on the five-factors for males and females were selected to make up the
163 composites. Fifteen faces was deemed sufficient to capture the average
164 configuration of high and low scoring individuals, as the perception of
165 individuality or distinctiveness in composite images changes little after the
166 merging of 6 images (Little & Hancock 2002). The average mean difference
167 between the highest rated 15 and lowest rated 15 was 2.68 for men and 3.24

168 for women. For both males and females, the personality scores of the
169 individuals selected as high for each trait were significantly higher than those
170 selected for the low group for the relevant trait (all $p < .001$), while no
171 difference was found between the high and low groups for any of the
172 personality traits for which the individuals were not selected (all $p > .31$). For
173 example, the high extravert group had significantly higher scores for
174 extraversion than the low extraversion group but did not significantly differ on
175 any other personality trait

176 For each set of 15 face images a single composite face was produced
177 for a total of 20 composites: 2 (high, low) X 5 (personality traits) X 2 (male,
178 female). The composite faces were created using specially designed software.
179 Key locations (174 points) were manually marked around the main features
180 (e.g., points outline, eyes, nose, and mouth) and the outline of each face (e.g.,
181 jaw line, hair line). The average location of each point in the 15 faces for each
182 composite was then calculated. The features of the individual faces were then
183 morphed to the relevant average shape before superimposing the images to
184 produce a photographic quality result. For more information on this technique
185 see Tiddeman, Burt, and Perrett (2001). The male and female composite
186 images can be seen in Figure 1.

187

188 **Figure 1 around here**

189

190 **B). Rating the composite faces for perceived personality**

191

192 **Participants**

193 Forty participants (15 male, 25 female, aged 19-35, mean = 22.9) rated the
194 composite faces for perceived personality. Thirty-three of these individuals
195 rated the faces for perceived attractiveness, masculinity and age.

196

197 **Ratings**

198 Participants were asked to rate the 20 composite faces for: agreeableness,
199 conscientiousness, extraversion, neuroticism, openness to experience,
200 attractiveness, masculinity and age. Ratings were on a 7-point scale (1=very
201 low, 7=very high) except for age judgements, for which participants were
202 asked to guess at the actual age of the face.

203

204 Faces were presented to participants on computer screen individually and in a
205 random order. Rating the face from 1-7 brought up the next face. Participants
206 rated the faces on a single dimension at a time (e.g., if asked to rate
207 agreeableness, all faces were rated for agreeableness followed by the next
208 rating block) and the order in which the traits were rated was randomised
209 between participants. There was no time limit for the ratings. Due to the length
210 of the rating task, participants were given the option of not rating the physical
211 traits (attractiveness, masculinity, age).

212

213 **Results**

214 **Calculating difference scores**

215 Difference scores were calculated for each type of rating of low and high
216 personality trait face pairs (high-low). For example, if a participant judging
217 extraversion rates 7 for the high extravert face and 5 for the low extravert face

218 this would give a difference score of +2. This single score thus represented
219 whether judges were accurate, as indicated by a positive score, or not and
220 allowed comparison between accuracy and trait.

221

222 **Reliability of raw ratings**

223 Co-efficient α was calculated for each trait across the 20 rated faces (40
224 participants for judgements of agreeableness, conscientiousness,
225 extraversion, neuroticism, and openness, and 33 participants for judgements
226 of attractiveness, masculinity, and age). This revealed moderate to high
227 agreement for all traits (age = .84, masculinity = .98, agreeableness = .76,
228 conscientiousness = .79, extraversion = .83, neurotic = .64, openness = .56)
229 bar attractiveness which was relatively low (.35, but splitting by sex of face
230 revealed greater agreement within each sex, female = .54 and male = .55).

231

232 **Accuracy by trait**

233 Using 1-sample t-tests against chance (0), for female faces, a significant
234 difference was found for rated agreeableness for the agreeableness faces (t_{40}
235 = 2.1, $p = .039$), rated conscientiousness for the conscientiousness faces (t_{40}
236 = 2.6, $p = .014$), rated extraversion for the extravert faces ($t_{40} = 2.4$, $p = .026$),
237 and neuroticism for neuroticism faces ($t_{40} = 2.2$, $p = .033$). Only the extravert
238 faces rated for extraversion ($t_{40} = 2.4$, $p = .022$) were significantly different
239 from chance for the males. For females, the score for openness to experience
240 faces did not significantly differ from chance ($t_{40} = 2.0$, $p = .057$) but was very
241 close to the 0.05 criterion for significance. For males, scores for face pairs
242 rated for the relevant personality trait did not significantly differ for

243 agreeableness ($t_{40} = -0.7, p = .52$), neuroticism ($t_{40} = 0.7, p = .50$), or
244 openness to experience ($t_{40} = 0.6, p = .58$), and there was marginal trend for a
245 relationship with conscientiousness ($t_{40} = 1.7, p = .098$). All significant
246 differences were in line with accurate ratings (as in fact were the non-
247 significant differences, in all but one case). Mean difference scores can be
248 seen in Table 1.

249

250 **Table 1 around here**

251

252 **Attractiveness, masculinity, and age**

253 Difference scores were also calculated for composite pairs rated for
254 attractiveness, masculinity, and age. Again using 1-sample t-tests against
255 chance, for attractiveness ratings, significant differences for the female
256 agreeableness pair ($t_{32} = 2.8, p = .008$) and the male neuroticism pair ($t_{32} =$
257 $2.5, p = .018$) were found. The high agreeable face was rated higher than the
258 low agreeableness face for females and the high neuroticism face was rated
259 higher than the low neuroticism faces for males. For females,
260 conscientiousness ($t_{32} = 0.9, p = .37$), extraversion ($t_{32} = -0.4, p = .68$),
261 neuroticism ($t_{32} = 0.6, p = .55$), and openness to experience ($t_{32} = 0.7, p = .47$)
262 face pairs did not generate a difference score that was different from chance.
263 For males, conscientiousness ($t_{32} = -1.8, p = .083$) and openness to
264 experience ($t_{32} = -1.8, p = .083$) faces pairs had difference scores that were
265 marginally significant while agreeableness ($t_{32} = 0.6, p = .55$) and extraversion
266 ($t_{32} = -0.2, p = .88$), face pairs did not generate a difference score that was
267 different from chance. Mean difference scores can be seen in Table 2.

268

269 **Table 2 around here**

270

271 For ratings of masculinity, significant differences for the male agreeableness
272 pair ($t_{32} = -2.1, p = .044$), the male extraversion pair ($t_{32} = -2.1, p = .040$), and
273 the female neuroticism pair ($t_{32} = -2.1, p = .044$) were found. The high
274 agreeableness face was rated lower than the low agreeableness face and the
275 high extraversion face was rated higher than the low extraversion face for
276 males. The high neuroticism face was rated lower than the low neuroticism
277 faces for females. There was a marginally significant effect for a similar
278 pattern for agreeableness in females ($t_{32} = -1.9, p = .062$), with low
279 agreeableness faces appearing more masculine. For females,
280 conscientiousness ($t_{32} = -1.6, p = .13$), extraversion ($t_{32} = 0.4, p = .66$), and
281 openness to experience ($t_{32} = 0.0, p = 1.0$) face pairs did not generate a
282 difference score that was different from chance. For males, conscientiousness
283 ($t_{32} = -0.2, p = .87$), neuroticism ($t_{32} = -0.1, p = .90$), and openness to
284 experience ($t_{32} = 0.5, p = .63$) face pairs did not generate a difference score
285 that was different from chance. Mean difference scores can be seen in Table
286 2.

287 For ratings of age the differences for the male extraversion pair ($t_{32} =$
288 $2.2, p = .032$), the male neuroticism pair ($t_{32} = -3.0, p = .005$), the male
289 openness to experience pair ($t_{32} = 2.7, p = .010$), and the female
290 conscientiousness pair ($t_{32} = -3.5, p < .001$) was found with the high
291 extraversion, neuroticism, and openness to experience faces appearing older
292 than low traits faces for males and the high conscientiousness face being

293 rated as younger than the low conscientiousness for females. For females, the
294 difference score for extraversion had a trend towards significance ($t_{32} = 2.2$, p
295 $= .094$) while the agreeableness ($t_{32} = -0.7$, $p = .49$), neuroticism ($t_{32} = 0.1$, $p =$
296 $.92$), and openness to experience ($t_{32} = -0.5$, $p = .66$) face pairs did not
297 generate a difference score that was different from chance. For males,
298 agreeableness ($t_{32} = 0.2$, $p = .85$) and conscientiousness ($t_{32} = -1.4$, $p = .17$)
299 face pairs did not generate a difference score that was different from chance.
300 Mean difference scores can be seen in Table 2.

301

302 **Relationships across traits and faces**

303 A repeated measures ANOVA ($5 \times 5 \times 2 \times 2$) was conducted with 'trait', 'face', and
304 'sex of face' as within-participant factors and 'sex of rater' as a between-
305 participant factor. This revealed a significant within participant effect of 'trait'
306 ($F_{4,152} = 3.2$, $p = .014$) and significant interactions between 'face' x 'trait'
307 ($F_{16,608} = 3.2$, $p < .001$), 'face' x 'trait' x 'sex of rater' ($F_{16,608} = 2.8$, $p < .001$),
308 and 'face' x 'trait' x 'sex of face' ($F_{16,608} = 2.3$, $p = .003$). No other significant
309 effects or interactions were found (all $p > .14$).

310 The significant effect of 'trait' x 'face' can be seen in Figure 2. There
311 are many relationships in the data showing cross-talk between face and trait
312 and the discussions below focus on differences related to accuracy in trait
313 ratings. Figure 2 shows that for four of the five traits pairs differ most on their
314 own rated trait. The predicted means taking into account ratings on different
315 traits show that conscientiousness, extraversion, neuroticism, and openness
316 rating differences have the most positive difference score for their relevant
317 traits at levels comparable to or greater than the original raw scores. The

318 agreeableness face pair differed most on rated conscientiousness suggesting
319 that there was little accuracy in judging agreeableness but those individuals
320 who are agreeable have faces that appear more conscientiousness. Other
321 interactions are not followed up here.

322

323 **Figure 2 around here**

324

325

Discussion

326 This study shows that, when judging composite facial images, individuals are
327 able to infer the personality of others somewhat accurately based only on
328 facial information. This may mean that individuals are indeed correct in
329 thinking their judgements of others' personality based only on facial
330 information are accurate (Hassin and Trope, 2000; Ligget, 1974). Such
331 judgements are far from perfect, particularly considering that we only
332 examined extremes of personality scores – accuracy is likely to be lower when
333 individuals are more similar in personality.

334 Analysis of individual traits revealed that some traits were judged with
335 more accuracy than others. In previous studies accuracy was most
336 consistently seen for judgements of extraversion and conscientiousness
337 (Albright et al. 1988, Passini & Norman 1966, Watson 1989) which also
338 appears to be reflected in the current study. In the original ratings, across both
339 males and females, both extraversion and conscientiousness face pairs were
340 rated accurately (though $p = .096$ for male conscientiousness pair). For the
341 female faces there were also indications of significantly accurate judgements

342 for agreeableness and neuroticism. This potential sex difference is discussed
343 in more detail below.

344 The repeated measures analysis revealed that overall there was a trait
345 by face interaction. Ratings of conscientiousness, extraversion, neuroticism,
346 and openness faces were most different in regard to their relevant trait. The
347 agreeableness face pair rated for agreeableness produced a difference score
348 near 0. This analysis confirms much of what was seen in the analysis of
349 original difference scores. Ignoring sex of face, judges were more accurate
350 than chance at estimating others' personality traits with the largest differences
351 being for conscientiousness and extraversion. Two interactions suggested
352 that accuracy was influenced by both the sex of the face judged and the sex
353 of the rater, though the small sample size of raters means we draw no strong
354 conclusion from the latter interaction.

355 Looking at the raw scores, accuracy is higher when judging female
356 than male faces. This may partially reflect the way in which the composite
357 images were made. The average mean difference between the highest rated
358 15 and lowest rated 15 was greater in the women than the men (male = 2.68,
359 female = 3.24). This is likely due to the size of the pool from which the
360 participants were drawn – nearly double the number of females participated in
361 the first part of the study and so a greater potential for variation in personality.
362 The difference means that the male composites were less extreme in their
363 actual personality than the females and so we might expect it to be harder to
364 accurately judge their personality. Of course that male and female faces are
365 judged differently may also reflect that female faces contain more cues to their
366 actual personality than do male faces.

367 As well as general accuracy it appears that some traits are more
368 accurately judged depending on whether male or female faces are being
369 rated. There appeared to be higher accuracy for conscientiousness in male
370 faces and neuroticism and openness for female faces. Again this may reflect
371 differences in sample size between male and females or that the validity of
372 cues differs depending on the judged face either in the face themselves or the
373 attention judges pay to particular traits in male and female faces.

374 The use of composite faces in this study shows that there exist
375 consistent cues to the personality of an individual which are available from
376 their face. As reviewed in the introduction, three likely candidates which have
377 received much attention in stereotype research are attractiveness,
378 masculinity, and age. For female faces, the high agreeableness composite
379 was more attractive than the low agreeableness composite and for male faces
380 the high neuroticism composite was more attractive than the low neuroticism
381 composite.

382 Attractiveness could then be a cue to accurate personality attribution,
383 although in fact judges were not accurate in assessing agreeableness. Of
384 course, if men base their partner choice on attractiveness judgments they are
385 also expressing a preference for partners who are actually agreeable. This
386 suggests that even without conscious information about personality,
387 individuals may make other potentially important judgements based on a link
388 between facial appearance and personality. The attractiveness of the high
389 neurotic male composite is somewhat surprising but may be explained by the
390 fact that this face was also seen as over a year younger than the low neurotic
391 composite. It is possible that a preference for youth (Buss 1989) explains the

392 attractiveness of the high neurotic male composite. The one year difference in
393 age is not large and the exact interplay between neuroticism, youth and
394 preferences remains to be examined.

395 Masculinity judgments also differed between high and low face pairs. In
396 female faces masculinity related negatively to agreeableness ($p = .062$) and
397 significantly positively to neuroticism, and in male faces masculinity was
398 significantly negatively related to agreeableness and positively to
399 extraversion. Such findings are consistent with perceptions of computer
400 graphic manipulations of sexual dimorphism which show that masculinity is
401 negatively related to personality traits associated with agreeableness (e.g.,
402 warmth, cooperativeness, Perrett et al., 1998). Though perceived masculinity
403 was related to real personality it appears that this trait was not used to
404 accurately assess personality as out of all the traits agreeableness showed
405 lowest overall accuracy.

406 Age judgements were related to conscientiousness in female faces and
407 extraversion, neuroticism and openness in male faces. The high
408 conscientiousness composite was rated as younger than the low
409 conscientiousness composite for female faces and the high extravert, high
410 openness, and low neuroticism composite faces were rated as older than their
411 counterparts. Again this shows that perceptual age is a potential cue to
412 accurate personality attribution.

413 Previous studies demonstrate accuracy in perceived personality using
414 more information than that shown here. Even in still facial photographs more
415 information is available to judges, such as clothing and hair style. Here
416 judges were accurate based only on facial information. Accuracy of

417 personality based on facial information may come about via self-fulfilling
418 prophecies (Snyder et al. 1977), whereby facial appearance affects social
419 perception leading individuals to behave in the way they are perceived to.
420 However, the causal direction could operate in the opposite direction, with
421 personality and behaviour affecting facial appearance. People with a tense
422 irritable temperament may tense certain facial muscles in a way that yields
423 different jaw development from that shown in people who are more relaxed
424 (Kreiborg, Jensen, Moller, & Bjork 1978). Personality may also be seen in
425 expressive habits. There is evidence that the personality dispositions of
426 elderly people are reflected in their faces, with those of a hostile disposition
427 tending to look angry even when posing in a neutral expression (Malatesta et
428 al. 1987). Accuracy can also be mediated through the environment, for
429 example, Cash (1990) reports that those who are highly sociable may choose
430 grooming aids that have a beneficial effect on their appearance. Another
431 potential source of accuracy comes from a biological link between personality
432 and facial appearance. For example, testosterone is proposed to be
433 responsible for masculine male facial traits (Enlow 1982) and is also linked to
434 male dominance behaviours (Mazur & Booth 1998), potentially providing a
435 biological link between the two. The reasons why personality is accurately
436 perceived remain to be studied and assessing the cues that people use may
437 be an important source of data in addressing this question.

438 The idea of judging an individual's personality from their appearance
439 may be seen as inherently undesirable (e.g., the common phrase "don't judge
440 a book by its cover") but this in no way implies that it is not important to
441 attempt to understand this area. In fact, the evidence that people appear to

442 make personality judgements based on only minimal information despite
443 society's discouragement implies that this is an area of fundamental
444 importance in social perception.

445

References

- 446 Albright, L., Kenny, D. A. & Malloy, T. E. (1988) Consensus in personality
447 judgments at zero acquaintance. *Journal of Personality and Social Psychology*
448 55(3): 387-395.
- 449
- 450 Ambady, N., Hallahan, M. & Rosenthal, R. (1995) On judging and being judged
451 accurately in zero-acquaintance situations. *Journal of Personality and Social*
452 *Psychology* 69(3): 518-529.
- 453
- 454 Benson, P. J. & Perrett, D. I. (1993) Extracting prototypical facial images from
455 exemplars. *Perception* 22: 257-262.
- 456
- 457 Berry, D. S. & Brownlow, S. (1989) Were the physiognomists right? Personality
458 correlates of facial babyishness. *Personality and Social Psychology Bulletin*
459 15: 266-279.
- 460
- 461 Berry, D. S. & McArthur, L. Z. (1985) Some components and consequences of a
462 babyface. *Journal of Personality and Social Psychology* 48(2): 312-323.
- 463
- 464 Berry, S. B. & McArthur, L. Z. (1986) Perceiving character in faces: The impact of
465 age-related craniofacial changes on social perception. *Psychological Bulletin*
466 100: 3-18.
- 467
- 468 Bond, C. F., Berry, D. S. & Omar, A. (1994) The kernel of truth in judgements of
469 deceptiveness. *Basic and Applied Social Psychology* 15: 523-534.

470

471 Borkenau, P. & Liebler, A. (1993) Convergence of stranger ratings of personality and
472 intelligence with self-ratings, partner ratings, and measured intelligence.

473 *Journal of Personality and Social Psychology* 65(3): 546-553.

474

475 Buss, D. M. (1989) Sex differences in human mate preferences: Evolutionary
476 hypotheses tested in 37 cultures. *Behavioural and Brain Sciences* 12: 1-49.

477

478 Cash, T. F. (1990) The psychology of physical appearance: Aesthetics, attributes, and
479 images. In: *Body Images: Development, Deviance, and Change*, ed. eds. T. F.
480 Cash & T. Pruzinsky, Guilford.

481

482 Dion, K., Berscheid, E. & Walster, E. (1972) What is beautiful is good. *Journal of*
483 *Personality and Social Psychology* 24: 285-290.

484

485 Enlow, D. M. (1982) *Handbook of Facial Growth*, Saunders.

486

487 Feingold, A. (1992) Good-looking people are not what we think. *Psychological*
488 *Bulletin* 111: 304-341.

489

490 Galton, F. J. (1878) Composite portraits. *Nature* 18: 97-100.

491

492 Gosling, S. D., Ko, S. J., Mannarelli, T. & Morris, M. E. (2002) A room with a cue:
493 Personality judgments based on offices and bedrooms. *Journal of Personality*
494 *and Social Psychology* 82(3): 379-398.

495

496 Hassin, R. & Trope, Y. (2000) Facing faces: Studies on the cognitive aspects of
497 physiognomy. *Journal of Personality and Social Psychology* 78(5): 837-852.

498

499 Kenny, D. A., Albright, L., Malloy, T. E. & Kashy, D. A. (1994) Consensus in
500 interpersonal perception: acquaintance and the big five. *Psychological Bulletin*
501 116: 245-258.

502

503 Kenny, D. A., Horner, C., Kashy, D. A. & Chu, L. (1992) Consensus at zero
504 acquaintance: replication, behavioral cues and stability. *Journal of Personality*
505 *and Social Psychology* 62: 88-97.

506

507 Kreiborg, S., Jensen, B. L., Moller, E. & Bjork, A. (1978) Craniofacial Growth in a
508 Case of Congenital Muscular-Dystrophy - Roentgencephalometric and
509 Electromyographic Investigation. *American Journal of Orthodontics and*
510 *Dentofacial Orthopedics* 74(2): 207-215.

511

512 Liggett, J. (1974) *The Human Face*, Constable.

513

514 Little, A. C. & Hancock, P. J. B. (2002) The role of masculinity and distinctiveness in
515 judgments of human male facial attractiveness. *British Journal of Psychology*
516 93: 451-464.

517

- 518 Malatesta, C. Z., Fiore, M. J. & Messina, J. J. (1987) Affect, Personality, and Facial
519 Expressive Characteristics of Older-People. *Psychology and Aging* 2(1): 64-
520 69.
- 521
- 522 Mazur, A. & Booth, A. (1998) Testosterone and dominance in men. *Behavioural and*
523 *Brain Sciences* 21: 353-371.
- 524
- 525 McCrae, R. R. & Costa, P. T. (1985) Updating Norman Adequate Taxonomy -
526 Intelligence and Personality Dimensions in Natural-Language and in
527 Questionnaires. *Journal of Personality and Social Psychology* 49(3): 710-721.
- 528
- 529 McCrae, R. R. & Costa, P. T. (1987) Validation of the Five Factor model of
530 personality across instruments and observers. *Journal of Personality and*
531 *Social Psychology* 52(1): 81-90.
- 532
- 533 Norman, W. T. (1963) Toward an adequate taxonomy of personality attributes:
534 Replicated factor structure in peer nominated personality ratings. *Journal of*
535 *Abnormal and Social Psychology* 66: 574-583.
- 536
- 537 Passini, F. T. & Norman, W. T. (1966) A universal conception of personallity
538 structure? *Journal of Personlity and Social Psychology* 4: 44-49.
- 539
- 540 Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D. R., Yoshikawa, S., Burt, D.
541 M., Henzi, S. P., Castles, D. L. & Akamatsu, S. (1998) Effects of sexual
542 dimorphism on facial attractiveness. *Nature* 394: 884-887.

543

544 Snyder, M., Tanke, E. D. & Berscheid, E. (1977) Social perception and interpersonal

545 behaviour: On the self-fulfilling nature of social stereotypes. *Journal of*546 *Personality and Social Psychology* 35: 656-666.

547

548 Tiddeman, B. P., Burt, D. M. & Perrett, D. I. (2001) Prototyping and transforming

549 facial texture for perception research. *IEEE Computer Graphics and*550 *Applications* 21: 42-50.

551

552 Watson, D. (1989) Strangers' ratings of the five robust personality factors: Evidence

553 of a surprising convergence with self-report. *Journal of Personality and Social*554 *Psychology* 57: 120-128.

555

556 Zebrowitz, L. A., Hall, J. A., Murphy, N. A. & Rhodes, G. (2002) Looking smart and

557 looking good: Facial cues to intelligence and their origins. *Personality and*558 *Social Psychology Bulletin* 28(2): 238-249.

559

560 Zebrowitz, L. A. & Montepare, J. M. (1992) Impressions of babyfaced males and

561 females across the life span. *Developmental Psychology* 28: 1143-1152.

Table 1: Mean difference scores for personality face pairs rated for their relevant trait for male and female faces

	Difference score on relevant trait	
	male	female
Agreeableness	-.13	.43*
Conscientiousness	.43	.55*
Extraversion	.53*	.53*
Neuroticism	.15	.48*
Openness to experience	.10	.45

***Significant at $p < 0.05$**

Table 2: Mean difference scores for personality face pairs (rating for high minus rating low) for each personality factor and for male and female faces. Positive scores indicate that the high scoring composite face is seen as more attractive, masculine or old and negative scores indicate that the low scoring composite is seen as more attractive, masculine or old.

	Attractiveness		Masculinity		Age	
	male	female	male	female	male	female
Agreeableness	.15	.64*	-.36*	-.39	.06	-.24
Conscientiousness	.36	-.21	-.03	-.33	-.45	-1.12**
Extraversion	.03	-.09	.39*	.09	.67*	.61
Neuroticism	.48*	.15	-.03	-.36*	-1.06*	.03
Openness to experience	-.36	.15	.09	.00	.97*	-.18

***Significant at $p < 0.05$ /**Significant at $p < 0.01$**

Figure 1: Composite facial images based on self-reported personality

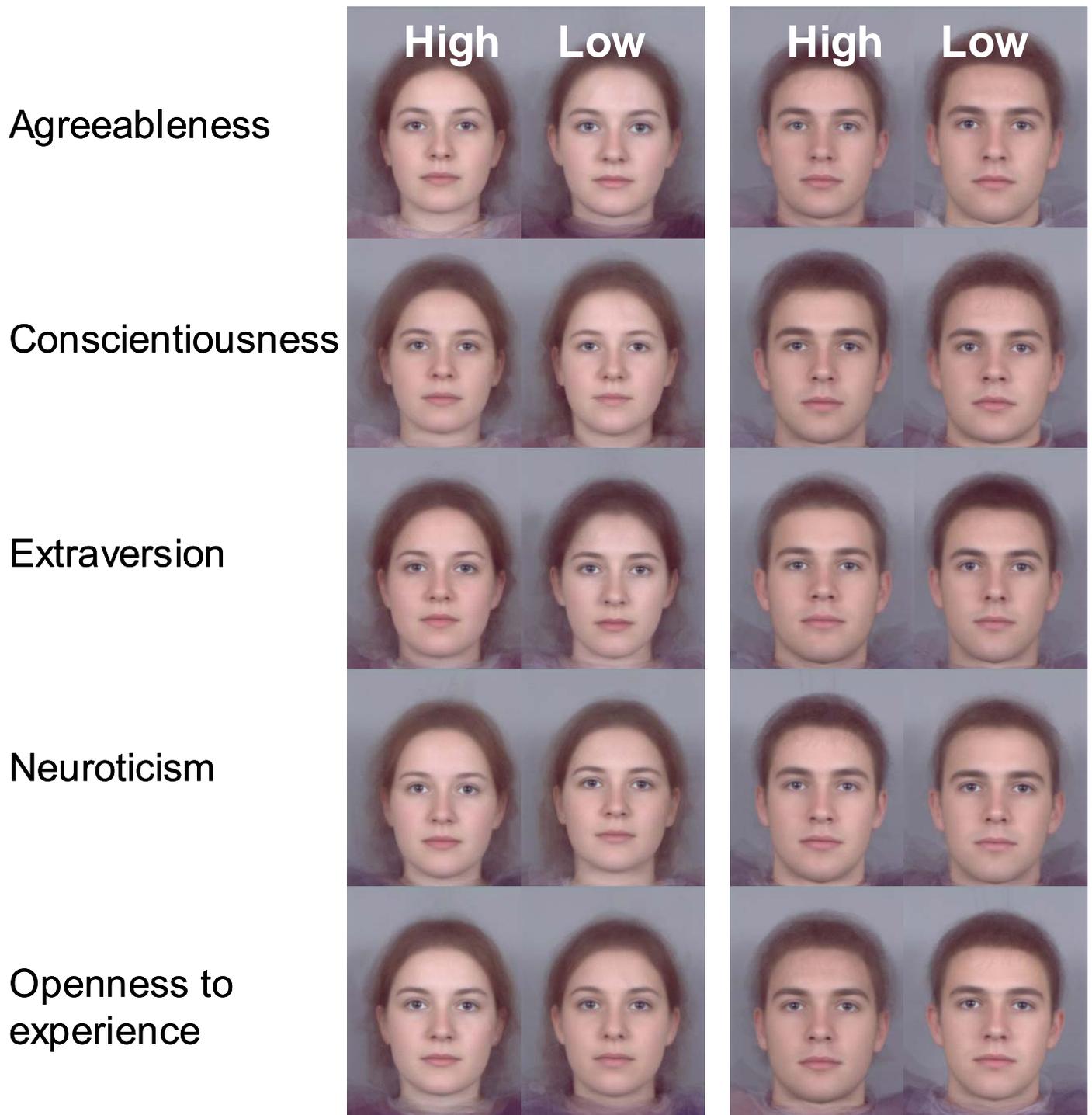
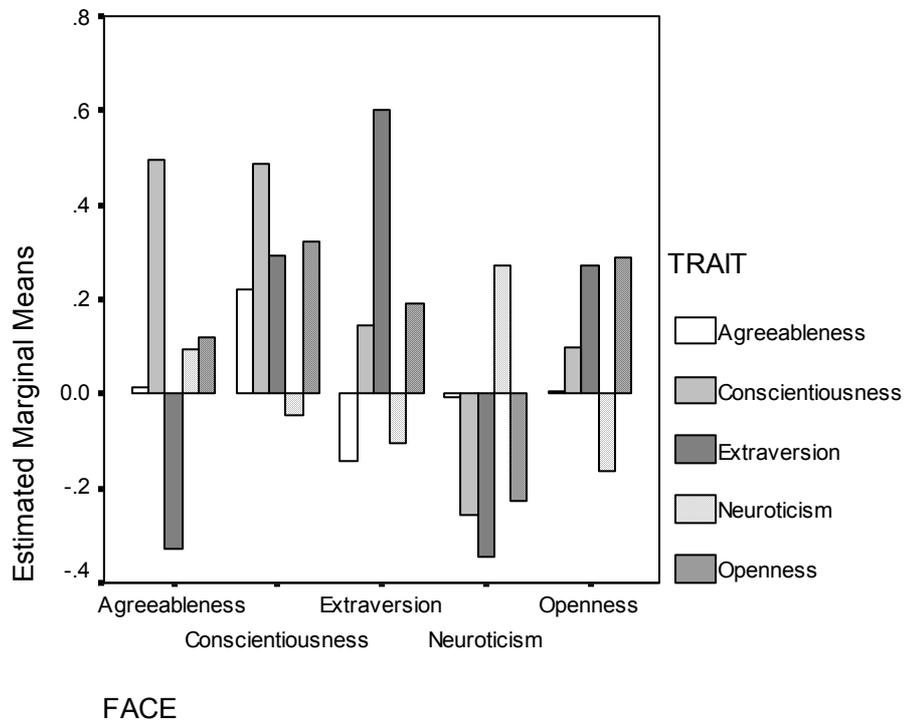


Figure 2: Estimated means of difference scores by trait and face from Face by Trait ANOVA.



Appendix: Factor loadings of 40-item questionnaires for males and females.

Table A: factor loadings of 40-item personality questionnaire for males (for trait, a = agreeableness, c = conscientiousness, e = extraversion, n = neuroticism, o = openness to experience). Loadings over 0.4 in larger type.

			factor 1	factor 2	factor 3	factor 4	factor 5
trait	low	high	Extraversion	Conscientiousness	Neuroticism	Open to experience	Agreeableness
a	Antagonistic	Acquiescent	0.02	0.37	-0.12	0.02	0.44
a	Callous	Sympathetic	0.59	0.06	0.27	0.04	0.36
a	Flexible	Stubborn	0.07	-0.06	-0.06	-0.13	-0.52
a	Forgiving	Vengeful	-0.44	0.07	0.06	0.28	-0.29
a	Lenient	Critical	-0.12	0.10	0.10	0.00	-0.75
a	Ruthless	Soft-hearted	0.59	0.02	0.31	-0.13	0.18
a	Selfish	Selfless	0.40	0.02	-0.12	-0.01	0.36
a	Trusting	Suspicious	-0.28	0.11	0.33	0.23	-0.35
c	Careless	Careful	0.04	0.65	0.19	0.04	-0.06
c	Conscientious	Negligent	-0.25	-0.73	-0.05	0.04	-0.02
c	Late	Punctual	-0.04	0.60	-0.23	-0.30	0.01
c	Lazy	Hardworking	-0.05	0.74	-0.18	-0.01	0.05
c	Persevering	Quitting	-0.33	-0.37	0.02	-0.33	0.05
c	Self-disciplined	Weak-willed	0.07	-0.79	0.27	-0.04	0.06
c	Undependable	Reliable	0.04	0.61	-0.22	-0.12	0.23
c	Well-organised	Disorganised	0.13	-0.53	0.25	-0.06	0.28
e	Friendly	Aloof	-0.76	-0.24	-0.07	0.13	-0.08
e	Inhibited	Spontaneous	0.45	-0.08	-0.26	0.60	0.09
e	Joiner	Loner	-0.76	-0.14	0.18	0.09	0.04
e	Quiet	Talkative	0.77	0.03	-0.04	0.20	-0.15
e	Reserved	Affectionate	0.67	-0.18	-0.11	0.11	0.12
e	Sober	Fun loving	0.48	-0.38	-0.32	0.30	0.12

e	Sociable	Retiring	-0.73	0.02	0.04	-0.16	0.29
e	Warm	Cold	-0.88	-0.15	0.01	-0.10	0.02
n	At ease	Nervous	-0.16	-0.10	0.58	-0.09	-0.36
n	Insecure	Secure	0.05	0.17	-0.78	0.12	-0.16
n	Relaxed	Highly Strung	-0.12	0.01	0.61	-0.04	-0.36
n	Self-conscious	Comfortable	-0.02	0.18	-0.73	0.07	0.08
n	Self-satisfied	Self-pitying	-0.04	-0.40	0.43	-0.11	-0.13
n	Unemotional	Emotional	0.69	-0.23	0.19	0.29	0.02
n	Vulnerable	Hardy	-0.35	0.16	-0.63	-0.03	-0.23
n	Worrying	Calm	-0.13	0.18	-0.71	-0.06	0.16
o	Broad interests	Narrow interests	-0.08	-0.05	0.07	-0.59	-0.03
o	Complex	Simple	0.17	-0.32	-0.32	-0.50	-0.09
o	Conforming	Independent	-0.16	-0.02	0.01	0.70	-0.02
o	Conservative	Liberal	-0.16	-0.07	-0.18	0.36	0.60
o	Conventional	Original	0.25	-0.23	-0.16	0.73	0.22
o	Daring	Unadventurous	-0.37	0.07	0.43	-0.31	0.07
o	Down to earth	Imaginative	-0.03	-0.15	-0.03	0.63	0.00
o	Uncreative	Creative	0.32	0.08	-0.07	0.61	0.13

Table B: factor loadings of 40-item personality questionnaire for females
(for trait, a = agreeableness, c = conscientiousness, e = extraversion, n =
neuroticism, o = openness to experience). Loadings over 0.4 in larger
type.

			factor 1	factor 2	factor 3	factor 4	factor 5
trait	low	high	Extraversion	Conscientiousness	Neuroticism	Agreeableness	Open to experience
a	Antagonistic	Acquiescent	-0.15	0.03	-0.06	0.49	-0.19
a	Callous	Sympathetic	0.24	0.11	0.05	0.71	0.09
a	Flexible	Stubborn	-0.18	0.05	0.34	-0.45	-0.06
a	Forgiving	Vengeful	-0.10	-0.12	0.22	-0.58	-0.02
a	Lenient	Critical	0.04	0.16	0.15	-0.54	-0.15
a	Ruthless	Soft-hearted	0.07	-0.04	0.18	0.75	-0.25
a	Selfish	Selfless	-0.03	0.09	-0.10	0.46	0.13
a	Trusting	Suspicious	-0.18	-0.02	0.37	-0.44	-0.09
c	Careless	Careful	0.03	0.57	0.05	0.28	-0.08
c	Conscientious	Negligent	0.09	-0.71	0.08	-0.13	0.02
c	Late	Punctual	-0.13	0.41	-0.17	-0.14	-0.18
c	Lazy	Hardworking	0.06	0.79	0.01	0.06	-0.02
c	Persevering	Quitting	-0.07	-0.54	0.06	-0.02	-0.42
c	Self-disciplined	Weak-willed	-0.07	-0.82	-0.10	0.14	-0.03
c	Undependable	Reliable	-0.11	0.44	-0.32	0.17	-0.03
c	Well-organised	Disorganised	0.02	-0.73	-0.01	0.05	0.09
e	Friendly	Aloof	-0.59	-0.08	0.28	-0.38	-0.14
e	Inhibited	Spontaneous	0.54	-0.22	-0.23	-0.10	0.47
e	Joiner	Loner	-0.67	0.08	0.24	-0.22	0.19
e	Quiet	Talkative	0.79	-0.13	-0.05	-0.15	0.04
e	Reserved	Affectionate	0.81	0.07	0.17	0.04	0.13
e	Sober	Fun loving	0.68	-0.22	-0.27	0.15	0.28
e	Sociable	Retiring	-0.75	-0.08	0.26	-0.10	-0.03
e	Warm	Cold	-0.47	-0.08	-0.02	-0.53	-0.14
n	At ease	Nervous	-0.49	0.02	0.54	-0.06	-0.02
n	Insecure	Secure	0.21	0.12	-0.63	0.19	-0.11

n	Relaxed	Highly Strung	-0.10	-0.06	0.68	-0.19	-0.12
n	Self-conscious	Comfortable	0.45	0.22	-0.40	0.00	0.10
n	Self-satisfied	Self-pitying	-0.29	-0.29	0.22	-0.20	-0.09
n	Unemotional	Emotional	0.46	0.06	0.51	0.21	0.19
n	Vulnerable	Hardy	0.12	0.24	-0.61	-0.10	0.28
n	Worrying	Calm	0.18	-0.13	-0.61	0.24	0.08
o	Broad interests	Narrow interests	-0.19	-0.11	0.40	-0.15	-0.42
o	Complex	Simple	-0.11	-0.28	-0.37	0.05	-0.43
o	Conforming	Independent	-0.01	0.03	-0.25	-0.01	0.68
o	Conservative	Liberal	0.23	-0.11	-0.04	0.04	0.51
o	Conventional	Original	-0.05	-0.07	-0.03	-0.02	0.77
o	Daring	Unadventurous	-0.36	0.02	0.35	0.00	-0.31
o	Down to earth	Imaginative	0.09	-0.37	0.09	0.09	0.53
o	Uncreative	Creative	0.05	0.01	0.00	0.09	0.64