Appendix F: Pilot Studies

Contents

Pilot Study 1

Pilot Study 1 aimed to locate twenty-four popular celebrities that could be used as targets for the subsequent experimental phases. Twenty-four provided a good average estimate and also provided a divisible number for the setup of the conditions of the experiments (male/female, six conditions). By establishing the twenty-four most popular celebrities, there was a higher likelihood of participants being familiar to them in the experiments.

Design: An online study approach was used to reach a sufficiently sized audience. Participants were asked to name their top ten celebrity names: ten was deemed a suitable number to request before participants would become bored or find it too difficult to think of more names. Ten was also chosen as a suitable number to generate enough power from the experiment using a moderate number of participants. The top twenty-four celebrity names from the results of this pilot study became the targets for the subsequent celebrity versions of the experiments.

Procedure: Participants were required to read an online information sheet and fill out a questionnaire and consent form (see Appendix G) prior to completing the pilot study. Participants were asked for their age, nationality, gender identity and ethnic group.

During the test, participants were asked to name their top ten celebrities. These could be any celebrities, including actors, singers, presenters etc. Data were collected via two online survey platforms: Bristol Online Surveys (LJMU Participants) and SurveyMonkey (UoD Participants) (see Appendix E for the testing layout). The experiment requested

participants' type in the names of their top ten celebrities into ten blank text boxes that were provided on the same page.

Participants: Twenty-eight participants were recruited, six males and twenty-two females (Age range 22–52 years of age, mean age 28.39 yrs.). Participants were of different nationalities and ethnicities and recruited via mailing lists (see Appendix G) throughout the University of Dundee and Liverpool John Moores University. All participants confirmed that they were aged 18 yrs. or older and indicated they had normal or corrected to normal vision.

Results: A total of 180 different celebrity names were mentioned across 280 responses, with forty-four names receiving two or more hits (names mentioned by participants). These forty-four names were subsequently assessed for the criteria for target selection;

- age – Young to middle aged adults were chosen as compositing faces with strong age differences is problematic.
- ethnicity – White European faces were chosen due to the potential influence of the own-race bias.
- face modifications –Facial tattoos and piercings (excluding ears) may provide non-facial feature related cues for recognition, so individuals with these adornments were excluded.
- facial hair – Facial hair may provide a non-facial feature cue and make compositing problematic. Individuals with heavy facial hair were excluded.

The assessment reduced the sample to thirty-six names. Further searches of the internet for suitable images for each of the targets was used to find those twenty-four names that had the most suitable images that could be used in the experiment (see **Error! Reference source not found.** for final results). Images needed to be of a frontal view with a neutral expression and mouth closed, good resolution and without extreme shadows.

Table 1: Pilot study 1 - Hits for top 24 celebrities

		HITS
celebrity target	Brad Pitt	10
	Angelina Jolie	9
	Johnny Depp	8
	Leonardo DiCaprio	6
	Jennifer Aniston	5
	Cameron Diaz	3
	Hugh Jackman	3
	Jennifer Lawrence	3
	Katy Perry	3
	Robert Downey Jr	3
	Sandra Bullock	3
	Anne Hathaway	2
	Benedict Cumberbatch	2
	Courtney Cox	2
	Daniel Craig	2
	Daniel Radcliffe	2
	Ellen DeGeneres	2
	Gerard Butler	2
	Hugh Grant	2
	Justin Timberlake	2
	Keira Knightley	2
	Scarlett Johansson	2
	Taylor Swift	2
	Tom Hardy	2

Final celebrity images were sourced and chosen for each of the twenty-four targets from the results of Pilot study 1. These images were sourced online using internet search

engines. Images needed to be of a good resolution and in focus. Editorial images were mostly excluded as these tended to appear unnatural and with specific lighting. Mostly "paparazzi" photographs were chosen as most were in natural outdoor lighting (a few editorial images were used where good paparazzi images were not available). All images contained flash which reduced any lighting inconsistencies and shadows.

Pilot Study 2

Pilot Study 2 investigated whether the compositing technique used to generate the unique composite face images was convincing enough for the images produced to pass as faces of real people.

Design: A yes/no paradigm (selected/not selected) was used to inadvertently obtain 'suspicious' or 'non-convincing' responses by requesting participants to decide if a face was 'not trustworthy' as a proxy. Participants were asked to indicate the composite faces that they felt were *not* trustworthy. It is thought that a low trustworthiness rating would suggest that there is something in the face or image that is not quite "right" and participants would perceive an implausible face as unbelievable and in turn assess it as not trustworthy. Therefore, a unique composite face image with a high number of 'not trustworthy' responses would suggest that the image was not convincing as a real face. By using 'not trustworthy' decisions as a proxy to photographic realism the aim was to be able to obtain observations of any problems with the unique composite images without specifically asking participants to attend to the formal qualities of the image and instead look at the face. Those faces considered the least 'trustworthy' were not used to match with a celebrity/lecturer target face counterpart. Only those faces that scored the lowest (low number of hits) for 'not trustworthy' assessment were used for matching with the target faces.

Materials: One hundred unique composites were generated as described and presented as stimuli.

Procedure: An online study approach was used to reach a larger audience. Using the online data collection platform, Qualtrics, participants were required to read an online information sheet and fill out a questionnaire and consent form (See Appendix G). All participants were shown the 100 composites in arrays of five faces at a time. The array consisted of five randomly chosen faces presented in a row, of the same size. The order of faces in the row and order presentation of arrays was also randomised across participants. Participants were asked to click on those faces that they felt did not look 'trustworthy'. Participants could select more than one face or none at all from the array (see **Error! Reference source not found.**).
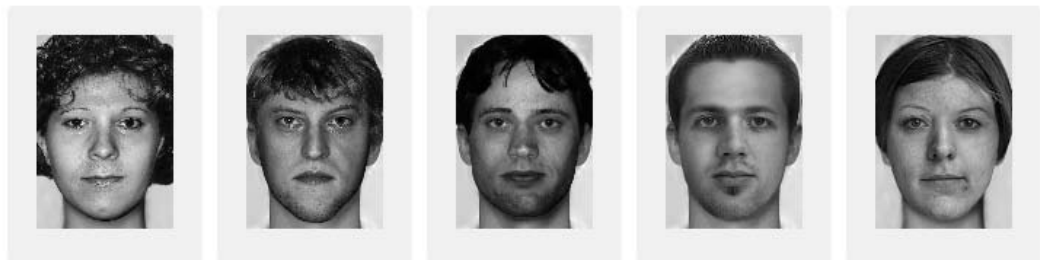

Figure 1: Pilot study 2 stimuli arrays

Arrays of five unique composites were shown one at a time, including instructions shown to participants

Once one set of selections had been completed, the program would continue onto the next array of five faces, after a blank screen interval of 1500ms. A total of 20 arrays were presented (20 x 5 = 100 unique composites).

Participants: Thirteen participants (male=3, female=10, *M*=27.92yrs, Range 21-45yrs) completed the online study using the data collection platform, Qualtrics, and were recruited via mailing lists (see Appendix G) at Liverpool John Moores University. Thirteen participants were recruited as ten or more would provide sufficient participant numbers

to indicate differences between items. All participants indicated they had normal or corrected to normal vision.

Results: Unique composites (n=100 [50 male, 50 female]) that received 'not trustworthy' hits when participants selected the face from the array were assigned a value of "1" (hit) (see Table 2).

Table 2: Pilot 2 - Histogram of 'Not trustworthy' hits for composite images

	'not trustworthy' hits										
	0	1	2	3	4	5	6	7	8	9	10
No. of composite Images	9	21	17	19	8	8	8	4	3	2	1

Results showed that composite images received between zero and ten hits (not trustworthy responses). Only composite images that received two hits or less were considered convincing enough to be used to match with target counterpart faces in order to generate the subsequent stimuli for the experiments. Forty-seven (20 male and 27 female) unique composite faces received two hits or less out of the 100 unique composite faces tested and formed the pool for selection during the target matching process (see **Error! Reference source not found.**).