An evaluation of post-production facial composite enhancement techniques

Josh P Davis, Stacie Simmons, Lucy Sulley

Department of Psychology, Social Work and Counselling, University of Greenwich, London, UK

Chris Solomon & Stuart Gibson

School of Physical Sciences, University of Kent, Canterbury, UK

Acknowledgments

The authors would like to thank Josi Cakebread, Amy Skelton, Sarah Thorniley, Helen Little, Sarah Poland, Natalie Baker, Sheena Belfon, and Lala Jammeh for their assistance in collecting data for this project.

The research was partly funded by two internal University of Greenwich grants to the first author.

Part of this research was presented at the American Psychology-Law Society Annual Meeting, 4th International Congress of Psychology and Law, Miami, Florida, USA, March 2011

Chris Solomon and Stuart Gibson are faculty members of the University of Kent and directors of VisionMetric Ltd. VisionMetric Ltd markets the EFITV and E-FIT facial composite systems. Solomon and Gibson’s contribution to this work was to facilitate the proportion of the data collection that took place at the University of Kent, development of the software and provision of software support. Data collection, analysis and interpretation were performed by Davis, Simmons and Sulley.

(Manuscript ID JFP-08-2014-0021)

An evaluation of post-production facial composite enhancement techniques

Abstract

Purpose: This article describes four experiments evaluating post-production enhancement techniques with facial composites mainly created using the EFIT-V holistic system.

Design/methodology/approach: Each experiment was conducted in two stages. In Stage 1, constructors created between one and four individual composites of unfamiliar targets. These were merged to create morphs, and in Experiment 3, composites were also vertically stretched. In Stage 2, participants familiar with the targets named or provided target-similarity ratings to the images.

Findings: In Experiments 1-3, correct naming rates were significantly higher to between-witness 4-morphs, within-witness 4-morphs and vertically stretched composites than to individual composites. In Experiment 4, there was a positive relationship between composite-target similarity ratings and between-witness morph-size (2-, 4-, 8-, 16-morphs).

Practical implications: The likelihood of a facial composite being recognised can be improved by morphing and vertical stretch.

Originality/value: A greater understanding of the theoretical underpinnings and applied advantage of post-production facial composite techniques should ensure greater acceptance by the criminal justice system, leading to better detection outcomes.

Keywords – Memory Eyewitness Facial Composite EFIT-V E-FIT Morphing Vertical Stretch

Paper type - Empirical paper

Introduction

In a criminal investigation, an eyewitness may from memory, construct a facial composite of the suspect. The aim is that someone familiar with that suspect may recognise the image. Traditional mechanical (e.g., Identikit, Photofit) and computerised (e.g., E-FIT, FACES, PRO-fit) *feature-based* systems required witnesses to describe the offender's face, and to select and assemble individual facial features. This is challenging, as humans mainly process faces holistically and not feature-by-feature (e.g., Tanaka and Farah, 1993). Consequently, feature-based composites are often poor target likenesses (for reviews see Davies and Valentine, 2007; Frowd, 2015; see also Fodarella et al., 2015, for a description of the composite system construction techniques described in this paper). *Holistic* systems have been designed to correspond with these processes (e.g., EFIT-V, EvoFIT, ID). Witnesses select, from a series of arrays, computer-generated realistic *whole* face images, until ideally the final composite matches their memory of the suspect. Accessing recognition, rather than recall, holistic composites are often better target likenesses than feature-based composites (Davis et al., 2010; Frowd et al., 2007b), and their use by the police has enhanced suspect identification rates (Frowd et al., 2010; Solomon et al., 2012).

Regardless of system, composite quality depends on factors including the witness' memory, the event (e.g. how long the offender was viewed) and the offender - higher composite naming rates are obtained with distinctive faces (e.g., Frowd et al., 2005). Post-production techniques can however enhance recognition. These include displaying multiple composites (Frowd et al., 2006), dynamic animated caricatures (Frowd et al., 2012), morphing composites constructed by the same (*within-witness morph*: e.g., Davis et al., 2010; Frowd et al., 2006; Valentine et al., 2010); or different witnesses (*between-witness morph*: e.g., Bruce et al., 2002; Hasel and Wells, 2007; Valentine et al., 2010), and *vertical stretch* (Frowd et al., 2013).

The effects from morphing are substantial. Valentine et al. (2010) found that correct naming rates to between-witness EFIT-V 4-morphs (44%; constructed by four different participants) were higher than within-witness 4-morphs (32%), and individual composites (24%). The authors explain the *morph-superiority effect* as a consequence of distinctive (e.g., Valentine, 1991) and external features of unfamiliar faces tending to be better remembered (e.g., Ellis, Shepherd, and Davies, 1979), and reproduced in composites (Frowd et al., 2007a). Individual composite errors are unlikely to be replicated by two or more different witnesses; or even the same witness. Morphing strengthens the accurate representations – being more plentiful, while errors are diluted. Internal features are most error prone and morphing improves these (Valentine et al., 2010) - fortuitous, as familiar face recognition (e.g., Ellis et al., 1979) and composite recognition is primarily internal feature driven (Frowd et al., 2007a).

Morphing can also increase image similarity to non-targets (Hasel and Wells, 2007). These findings can be explained by *face space* models of face recognition (e.g., Valentine, 1991), in which internal representations of familiar distinctive faces are proposed to be located in isolated peripheral areas of a multidimensional internal space, whereas typical faces cluster more densely at the centre and due to this increased density are more easily misidentified. Hasel and Wells (2007) suggest that morphing *averages* the appearance of composites, inducing a *morph-prototype* effect so that in terms of the face space model, they become less distinctive than the individual composites from which they are constructed. As such, they are more likely to possess the properties of a *prototype*, or typical face. The combined effects of the morph-superiority effect and the morph-prototype effect imply both an increased likelihood of identifying the actual offender, and of identifying an innocent suspect.

Another innovation, vertical stretching, in which the perceived height of the composite is doubled, while keeping the width constant – has, like morphing, been found to enhance composite naming rates (from 30% to 42%; Frowd et al., 2013). The authors suggest that like morphing, vertical stretch induces an averaging effect. They base this on Hole et al.'s (2002) proposal that the viewer automatically internally restructures the appearance of a vertically stretched image to match the normal expected veridical perspective. Nevertheless, Hasel and Wells (2007) based their morph-prototype theory on feature-based (FACES) composite-target similarity ratings, provided to targets and foils depicted in an array by *unfamiliar* raters. However, familiar and unfamiliar faces are processed using different mechanisms (see Bruce and Young, 1986), and recognising holistic composites – possibly of better quality - from a pool of potentially familiar candidates may draw on different cognitive processes.

The research base suggests that vertical stretch, between-witness morphing, and within-witness morphing should increase the likelihood of a correct suspect identification in a real police investigation. However, it is as important for the police to be aware of the potential negative consequences of application. No published research has previously reported as to whether, even when the viewer is familiar with the person depicted, these techniques also increase incorrect naming with EFIT-Vs, as would be predicted by the averaging hypothesis (Hasel & Wells, 2007). This was the primary aim of the current research. Experiment 1 therefore compared correct and incorrect naming rates to between-witness 4-morphs and individual composites constructed using EFIT-V (holistic) and E-FIT (feature-based) to compare between-system outcomes. Experiment 2 compared naming rates to within-witness

4-morphs and EFIT-V individual composites. Experiment 3 examined naming rates to EFIT-V individual composites, within-witness morphs and vertically stretched composites, as well as to composites manipulated using both techniques. Finally, Experiment 4 was designed to further investigate the theoretical underpinnings of the morphing effect. Between-witness morphs were created from up to 16 E-FITs and EFIT-Vs, and images were rated for target-composite similarity, and other 'personality' characteristics. In all experiments, composite *constructors* were unfamiliar, *participant* '*namers*,' or '*raters*,' were familiar with targets.

Experiment 1

Previous research has found that compared to individual composites, higher naming rates are associated with between-witness and within-witness EFIT-V 4-morphs (Valentine et al., 2010), and higher target-composite similarity ratings to within-witness E-FIT 2- and 3-morphs (Davis et al., 2010). However, incorrect naming rates were not reported, and no published research has examined between-witness morphing of E-FITs. Experiment 1 therefore compared correct and incorrect naming rates to E-FIT and EFIT-V individual composites and between-witness 4-morphs derived from these systems.

Our predictions were consistent with the morph-superiority effect. Correct naming rates were expected to be higher to 4-morphs than to individual composites. No predictions were made in relation to incorrect naming rates, as unlike in Hasel and Well's (2007) study, participants were familiar with the targets. Finally, holistic composites tend to be named more accurately than feature-based composites and similar effects might have been predicted here. However, Experiment 1 was conducted at training courses for police operators. The primary aim for the trainee operators assisting the constructors was system familiarisation, and not the production of high-quality composites. Indeed, Davis et al. (2010) found no differences in E-FITs and EFIT-Vs constructed in similar circumstances, and therefore correct naming rates were expected to be roughly equal.

Method

Design

In Stage 1, constructors created facial composites of between two and four target-actors. In Stage 2, participants named the composites in an independent-measures design with two factors: - image type (between-witness 4-morph, individual composite) and system (E-FIT, EFIT-V). The dependant variables were correct and incorrect naming rates.

Participants

Adult participant-constructors (n = 10) and operators (n = 10) were attendees at police operator training courses. Adult participant-namers (n = 226) were University of Greenwich students and watched *Eastenders* (BBC1) or *Hollyoaks* (Channel 4) regularly. Thirty-five participant-namers were excluded for failure to name real photographs of the *Eastenders* actors.

Materials and Procedure

Stage 1: Constructors viewed one of four 2-min videos white male target-actors from *Eastenders* knowing they were to subsequently create a composite. Approximately 15 minutes later, and following the administration of the Cognitive Interview by the operator (CI; see Memon, Meissner, & Fraser, 2010 for a review of the CI's positive effects on the quality of eyewitness testimony), to extract a verbal description of the target, they created an E-FIT or an EFIT-V (see Fodarella et al., 2015 for a description of the systems' procedures). Using the same composite system, they subsequently followed the same procedure to construct up to three additional composites of different targets with different operators (quantity was time-limited by the five hour sessions). In total, 17 E-FITs and 19 EFIT-Vs were created. For the purposes of Experiment 1, four E-FITS and four EFIT-Vs of each actor were randomly selected to be employed as individual composites in Stage 2.

Stage 2: Using the method described by Valentine *et al.* (2010), Morph Studio software (The Learning Company, Cambridge, MA) merged the four individual composites with each contributing 25% to between-witness morphs. All 40 images were transferred into ten counterbalanced and randomly ordered testing sets. Each set contained one image of each of the targets and a further four *Hollyoaks* EFIT-Vs created previously (Valentine et al., 2010) and included as distracters. Distracter data were not analysed.

Participant-namers were randomly assigned to one of the ten testing sets, and asked to provide names or individuating semantic information to the images. Responses were recorded as correct, incorrect, or no response. Participants subsequently attempted to name photographs of the targets to ensure they were familiar with those depicted.

Results

Naming data analyses were conducted by items. Namers provided names to 49.2% of the 4-morphs (64.6% of these were correct) and 38.1% of the individual composites (53.1% correct). Naming rates are displayed in Table 1. A 4 (actor: A-D) x 2 (image type: morph, individual composite) x 2 (system: E-FIT, EFIT-V) x 2 (response: correct name vs. other response: – no response or incorrect name) hierarchical backward elimination loglinear analysis revealed significant interactions between actor, system and correct responses, *and*, between image type and correct responses, $\chi^2(14) = 10.08$, $p = .758$.

Table 1 about here

A chi-squared analysis examining the interaction between system and response for each actor revealed a significant effect for Actor B only, $\chi^2_B(1, 235) = 8.29$, $p = .004$, $\Phi = .187$. E-FIT correct naming rates (34.6%) were higher than EFIT-Vs (18.2%). A similar test on the combined actor data was not significant, $\chi^2(1, 917) < 1$, $\Phi = .021$, correct responses to E-FITs (26.9%) and EFIT-Vs (25.1%) did not significantly differ. The interaction between image type and response revealed higher correct naming rates to 4-morphs (31.8%) than individual composites (20.3%), $\chi^2(1, 917) = 15.86$, $p < .001$, $\Phi = .131$.

A 4 (actor: A-D) x 2 (image type) x 2 (system) x 2 (response: incorrect name vs. other response: - no response or correct name) loglinear analysis revealed a model with a significant two-way interaction between actor and response only, $\chi^2(24) = 21.14$, $p = .630$; incorrect naming rates varied by actor.

Discussion

Consistent with the morph-superiority effect, correct naming rates were higher to between-witness 4-morphs than individual E-FIT and EFIT-V composites. There was no evidence of a morph-prototype effect, as overall incorrect naming rates were similar. The E-FITs (4-morphs and individual composites) of one actor (Actor B) were named more often than EFIT-Vs. This actor-specific effect may be a consequence of some faces being better recreated using a feature-based, rather than a holistic system, or that one or more of the constructors creating E-FITs of that actor might have possessed exceptionally good face recognition ability. However, as there were no overall differences in correct or incorrect naming rates between E-FITs or EFIT-Vs, it is possible that this effect was an anomaly. Nevertheless, Experiment 1 was limited by the use of only four targets, and greater numbers were included in Experiments 2 and 3. The applied and theoretical implications are returned to in the General Discussion.

Experiment 2

In some police investigations there may only be a single witness. In anecdotal conversations with the first author of this paper, experienced police operators have reported that many witnesses create more than one composite – if they feel they could produce a better image on the next attempt. Previous research has found a within-witness morph advantage for EFIT-Vs (Valentine et al., 2010), but no such research has investigated incorrect naming rates. Therefore, for Experiment 2, within-witness EFIT-V morphs were created to determine whether the morph-superiority effects for between-witness morphs found in Experiment 1 would be replicated.

In an investigation, the police would have to decide whether to publicise a morph or an individual composite. For this, they may request composite quality feedback from witnesses. Unfortunately, constructors rarely provide accurate assessments (e.g., Hasel & Wells, 2007). Therefore, in Experiment 2, within-witness 4-morph naming rates were compared with all individual composites produced by their constructors, and separately to the constructors' self-rated best composites. The hypotheses were consistent with the results of Experiment 1, in that correct naming rates were expected to be higher to 4-morphs than to individual composites, and no effects were expected in relation to incorrect naming rates.

Design

In an independent measures design with one variable (image type: individual composite, within-witness 4-morph), participants viewed eight counterbalanced and randomly ordered images. The dependent variables were correct and incorrect naming rates. However, as some of the participant-namers did not watch both TV soaps, conditionalised naming rates (CNR[1]) were calculated.

Participants

[1] CNRs for each participant are calculated by dividing the number of composites correctly or incorrectly named by the number of real photos correctly named.

Constructors (n = 8) unfamiliar with *Eastenders* or *Hollyoaks* attended EFIT-V operator (n = 8) training courses. Participant-namers (n = 250) were University of Greenwich students or visitors to the *Guerilla Science* tent at the *Secret Garden Party* (2009) music festival, Huntingdonshire. All regularly viewed at least one of the soaps.

Materials and procedure

Replicating the same procedure as Experiment 1, constructors viewed 2-min video clips of one white male target from *Eastenders*, or *Hollyoaks*. Thirty minutes later, with a different operator each time, they successively created four EFIT-Vs in a five hour time period. In total, four individual composites of each target were created, although in advance, constructors were only aware they were to create a single composite. Using Morph Studio, within-witness 4-morphs were weighted based on constructor target-composite rankings (the 'best' contributed 40%; 2nd 30%; 3rd 20%; 4th 10%). In total 40 images were created. Stage 2 replicated Experiment 1 except five counterbalanced and randomly ordered testing sets each contained one of each of the eight targets.

Results

Overall CNR rates to 4-morphs were 74.8% (72.6% correct), to the combined individual composites they were 63.9% (56.0% correct). The CNR rates to 4-morphs, combined individual composites (*comb*), and self-rated best individual composites (*best*) are presented in Table 2. An 8 (actor: A-H) x 2 (image type: morph, individual composite) x 2 (response: correct CNR, other response: incorrect CNR or no name provided) hierarchical backward elimination loglinear analyses revealed significant three-way highest order interactions, $\chi^2_{comb}(7) = 36.07$, $p < .001$; $\chi^2_{best}(7) = 39.66$, $p < .001$. Chi-squared analyses revealed that significantly more correct CNRs were provided to morphs than to the combined (A, C, D, E), and best composites (A, C, E, H) for four of the actors (see Table 2). Similar tests on the overall combined actor data were also significant, $\chi^2_{comb}(1, 1883) = 43.56$, $p < .001$, $\Phi = .152$, $\chi^2_{best}(1,754) = 43.55$, $p < .001$, $\Phi = .240$, with higher correct naming rates to morphs (54.3%) than combined (35.8%), and best-rated individual composites (30.6%).

Similar 8 (actor) x 2 (image type) x 2 (response: incorrect CNR, other response) analyses on the incorrect CNR data revealed significant three-way highest order interactions, $\chi^2_{comb}(7) = 36.07$, $p < .001$; $\chi^2_{best}(7) = 39.66$, $p < .001$. Follow up analyses found significantly higher incorrect CNR rates to morphs than the combined individual composites for Actor B, with the opposite found for Actors E and G. For a fourth actor (C), incorrect CNR rates were higher to the best self-rated individual composite than the 4-morph (see Table 2). The overall chi-squared test on the combined actor data was also significant but for the overall individual composite comparison only, $\chi^2_{comb}(1, 1245) = 25.17$, $p < .001$, $\Phi = .142$, $\chi^2_{best}(1, 754) < 1$, $\Phi = .034$, with incorrect CNRs significantly higher to individual composites (28.1%) than 4-morphs (20.5%).

Table 2 about here

Discussion

Consistent with Experiment 1, within-witness 4-morph correct CNRs (54.3%) were higher than combined individual composites (35.8%), and creators' self-rated best composites (30.6%). With one target only, there was a suggestion of a morph-prototype effect in that incorrect naming rates were significantly higher to the 4-morph than to the combined individual composites. This target was also correctly named proportionally the least often, suggesting that the constructor creating this target may have possessed poor face recognition ability. The actual incorrect names provided by the participants were not formally recorded, although it was noted at the time that primarily two different actors - not in the target set of eight - were mistakenly named. These actors at the time were involved in a high profile story in the soap and like the target possessed balding hair. As such, it is possible that the participant-namers based their responses on this characteristic only. Despite the results from this anomalous single target, the comparison including all eight targets revealed no evidence of a morph-prototype effect. Indeed, the opposite effect was found in that significantly more incorrect names were provided to the combined individual composites (28.1%) than morphs (20.5%).

Experiment 3

At least two facial composites are required to create a morph, whereas, in contrast, vertical stretch can be performed on a single composite. Nevertheless, no previous research has applied this technique to EFIT-Vs. Reemploying the constructors' self-reported best EFIT-V composites[2] and the 4-morphs from Experiment 2, Experiment 3 therefore compared naming rates to within-witness 4-morphs, individual composites, and vertically stretched composites,

[2] The self-reported 'best' individual composites and not the individual composites correctly named the most often in Experiment 2 were reemployed. In a real investigation, the police could only rely on ratings provided by the constructor.

and also examined whether a combination of both (*morph-and-stretch*) would incur a cumulative benefit.

Experiment 3 occurred approximately five years after Experiment 2, and the images were consequently far older than would normally occur in a police investigation in which facial composites are often publicised as quickly as possible. However, the police sometimes release composites of 'historical' cases (see Lohr, 2013), and the design therefore allowed for a test of whether these post-production techniques would also enhance naming rates in these circumstances.

Based on previous research, correct naming rates were expected to be higher to morphed and vertically stretched images than individual composites. However, no predictions were made in relation to morph-and-stretch composites, as there was expected to be an upper limit as to the benefits of post-production averaging methods. No effects were predicted in relation to incorrect naming.

Method

Design

In a independent measures design, participants named images of eight target-actors in one of four randomly allocated conditions - individual composite, within-witness 4-morph, vertical stretch, and morph-and-stretch. The dependent variables were correct and incorrect CNR.

Participants

Participants (*n* = 159)[3] were recruited from the University of Greenwich or from adverts on social media fan pages for viewers of *Hollyoaks* and *Eastenders*.

Materials

The within-witness 4-morphs and the self-rated 'best' individual composites from Experiment 2 were reemployed in Experiment 3. Using the *perceptual-backdrop image* (PBI) technique (see Frowd, 2015), vertically stretched composites were created by 'pulling up and

[3] There was a high dropout rate in Experiment 3. 268 participants started the study; however, 109 provided no data. One explanation may be that many failed to notice and to click on a final arrow on the Qualtrics system that saves all responses.

down’ the top and bottom right hand corners of the originally square composites using Photoshop’s free transform tool, leaving an isosceles trapezium shape[4].

Procedure

Using the Qualtrics online survey system (www.qualtrics.com), participants completed four stages[5]. In the second cued naming stage, they were informed that the eight images of celebrities had appeared on *Eastenders* or *Hollyoaks* in the previous 4-5 years. They were asked to type a name or individuating information into a response box. In the final stage, participants provided a name to the target-actor photographs to calculate CNR.

Results

Overall, 21.5% of CNRs were correct, and 9.2% were incorrect. As displayed in Table 3, CNRs to individual composites were 23.9% (31.0% correct), to 4-morphs they were 34.9% (79.7% correct), to vertically stretched composites they were 23.9% (78.7% correct) and finally to morph-and-stretched composites they were 39.2% (76.8% correct). Due to low case numbers, conducting analyses on the separate target data were not feasible and these data are polled in Table 3.

Table 3 about here

A 4 (condition: individual composite, 4-morph, stretch, morph-and-stretch) x 2 (correct CNR, other response) chi-squared test was significant, $\chi^2(3, 780) = 35.46$, $p < .001$, $\Phi = .213$. Correct CNRs were significantly lower to individual composites than to the other three conditions (all comparisons $p < .05$). The latter did not differ ($p > .1$).

A 4 (condition) x 2 (incorrect CNR, other response) analysis on the incorrect data was also significant, $\chi^2(3, 778) = 16.20$, $p = .001$, $\Phi = .144$. Incorrect CNRs were significantly higher to individual composites than both 4-morphs, and stretched composites ($p < .05$). No other comparison was significant ($p > .1$).

[4] Proportion width to left vertical = 1: 2.67; right vertical = 1: 3.40

[5] Note: First stage ‘un-cued naming’ data in which very few names were provided (M = < 0.2/8), and third stage ‘name-matching’ data to a longer list that included additional contemporary soap actors were collected and analysed - there were no differences between conditions ($p > .1$) For brevity, these are not reported as the conditions do not match the normal publication of composites in which some context would be provided.

Discussion

In Experiment 3, correct CNRs were higher in all three post-production enhancement conditions - within-witness 4-morphs, vertical stretch and morph-and-stretch than to individual composites. Incorrect naming rates were also higher to individual composites than to the 4-morphs and the vertically stretched images. The proportion of correct to incorrect naming rates to all post-production techniques were similar, demonstrating that morph-and-stretch did not provide an additional benefit, probably due to an averaging effect ceiling.

The composites were 5-years-old and unsurprising naming rates were lower in Experiment 3 than in Experiment 2, although a statistical comparison would not be valid as conditions differed across experiments. Indeed, some of the target actors had left the soaps in the interval, and the profile of others included in the stimuli set had altered substantially. Nevertheless, the police do release 'historical' composite images, and these results demonstrate that morphing or vertical stretch might enhance the likelihood of a suspect identification in these cases.

Considering composite age, the proportion of correct to incorrect responses in Experiment 3 was high. One explanation was that participants may have accessed internet information as to which actors were appearing in the soaps at the time, although as assignation to conditions was random this is unlikely to have confounded the results. Police officers and not the public are the primary identifiers of facial composites. It could be argued that this explanation matches the intelligence-gathering process by a police officer who might access police records after viewing a publicised composite if they believed it to be familiar, or if they were involved in investigations of a similar nature at the time.

Experiment 4

In some crimes, more than four witnesses may view the same offender, and yet no previous research has examined the influence of creating morphs made up of more than four composites. Experiment 4 examined the influence of morphing up to 16 facial composites on composite-target similarity ratings and on the perceived characteristics of composites. People make consistent unfamiliar face personality attributions (e.g., occupation, intelligence, trustworthiness, extraversion and criminality; Albright et al., 1988; Klatzy et al., 1982; Santos and Young, 2005; Zebrowitz et al., 2002), and perceived criminality is negatively related to attractiveness (MacLin and MacLin, 2004), unsurprising as criminals are stereotyped as unattractive (Saladin et al., 1988).

Hasel and Wells (2007) demonstrated that a large component of the morph-superiority effect is derived from a morph-attractiveness effect. Morphing induces an averaging effect reducing distinctiveness, and increasing attractiveness (e.g., Rhodes and Tremawan, 1996; Valentine et al., 2004). Indeed, typical faces are perceived as more attractive. These effects may partly be a consequence of facial asymmetry and skin blemishes being averaged out, as symmetry – a marker of healthy genes - is also associated with attractiveness (Rhodes et al., 1999). (Note however that Valentine et al., 2004 found that morphing profile view facial photos also increases attractiveness, demonstrating attractiveness-symmetry independence.)

In Experiment 4, participant-acquaintances rated E-FIT and EFIT-V individual composites and between-witness 2-, 4-, 8-, and 16-morphs for similarity to a target photograph. Unfamiliar participants provided ratings of perceived attractiveness, criminal demeanour, and distinctiveness. Age and intelligence were also included as controls. The aim was to examine which characteristics predicted similarity, and to examine how increasing morphing size influences these characteristics.

The hypotheses were based on the morph-superiority and morph-attractiveness effects in that increased morph-size was expected to be positively related to both target-composite similarity and attractiveness, and negatively related to criminality. Based on the morph-prototype effect, increased morph-size was expected to be negatively related to distinctiveness.

Design

In a 3 (system: E-FIT: greyscale, EFIT-V: colour, EFIT-V: greyscale) x 5 (individual composite, 2-morph, 4-morph, 8-morph, 16-morph) repeated-measures design, target-unfamiliar participants rated a randomly counterbalanced series of 186 images of a female target for either, - attractiveness, criminality, distinctiveness, intelligence or perceived age. Target-acquaintances provided composite-target similarity ratings. EFIT-Vs are constructed in colour, E-FITs in colour or greyscale. For the current research however, E-FITs were greyscale. To ensure no confounding effects, EFIT-Vs were rated in both colour and greyscale. The dependent variables were the mean scale ratings (1-10; in which a high score = high similarity, or high value on that perceived personality trait).

Participants

Constructors (*n* = 64)[6] and paired operators (*n* = 64) were forensic science students from the University of Kent, trained with the systems in order to create composites for an assessed assignment. Target-acquaintance raters (*n* = 29) provided composite-target similarity ratings. Target-unfamiliar raters (*n* = 175), visitors to the *Guerrilla Science* tent at *Secret Garden Festival*, Cambridgeshire (2010) were randomly assigned to one of the rating conditions.

Materials and procedure

The EFIT-Vs (colour) and E-FITs (greyscale) of a female target (aged 21 years) viewed in a video of 1 min 18 sec were originally created for research examining the influence of construction on line-up outcomes (see Davis et al., 2014) [7]. A different operator and paired constructor created each individual composite approximately 15 min after viewing the video – constructors were aware in advance they would be creating a composite. Morph-studio software produced between-witness 2-, 4-, 8-, and 16-morphs. Each contributed 50% to the successive morph (e.g., each individual composite contributed 0.0625 to the 16-morphs). Colour EFIT-V images were transformed to greyscale using Photoshop. All 186 images were transferred to *PowerPoint* for randomised and counterbalanced display with one image on each slide.

Target-unfamiliar raters rated the 186 images (16-morphs = 6, 8-morphs = 12, 4-morphs = 24, 2-morphs = 48, individual composites = 96), for either perceived age, or one of the characteristics (1-10). Target-acquaintances provided target-composite similarity ratings to a video still presented adjacent to the image.

Results

Table 4 displays the mean perceived age, attractiveness, criminality, distinctiveness, intelligence, and similarity ratings as a function of composite size and system. A 5 (composite size: individual composite, 2-morph, 4-morph, 8-morph, 16-morph) x 3

[6] Note: In the study by Davis et al. (2014; Experiment 1), 117 constructors created individual composites (E-FIT = 60, EFIT-V = 57). For the current study, 32 of each type were randomly selected to be morphed.

[7] Operator training in Experiment 4 did not include the Cognitive Interview and at six hours was not to the same level as the approximately 24 hours police operators would have encountered prior to contributing to Experiments 1 and 2.

(composite system: E-FIT, EFIT-V greyscale, EFIT-V colour) MANOVA conducted on the six dependent variables revealed significant effects of composite size, Wilk's Lamda, Λ = .398, $F(24, 580.32) = 7.31$, $p < .001$, $\eta = .206$, and composite system, $\Lambda = .520$, $F(12, 332) = 10.71$, $p < .001$, $\eta = .279$. The interaction was not significant, $\Lambda = .718$, $F(48, 820.85) = 1.18$, $p > .1$, $\eta = .054$.

Table 4 about here

For the similarity data, only the main effect of size was significant, $F(4, 171) = 45.26$, $p < .001$, $\eta = .514$. Paired comparisons revealed no significant differences between the 16- (M = 3.93) and 8-morphs ($M = 3.71$, $p > .1$, $r = .27$), or the 8- and 4-morphs ($M = 3.56$, $p > .1$, $r = .18$). However, 4-morphs received higher ratings than 2-morphs ($M = 3.04$, $p < .001$, $r = .46$), which were also higher than individual composites ($M = 2.34$, $p < .001$, $r = .49$).

For the age variable, only the size main effect was significant, $F(4, 171) = 11.87$, $p < .001$, $\eta = .217$. Paired comparisons revealed no significant differences between morph types ($p >.1$), although all morphs were rated as younger than individual composites ($p < .05$).

With attractiveness, the main effect of size was significant, $F(4, 171) = 34.69$, $p < .001$, $\eta = .448$. Paired comparisons revealed no differences between 16- , 8- , or 4-morphs, although all three types received higher ratings than 2-morphs, which were also higher than individual composites ($p < .05$). The main effect of system was significant, $F(2, 171) = 5.34$, p = .006, $\eta = .059$. Greyscale and colour EFIT-Vs were rated as more attractive than E-FITs ($p < .05$). The two EFIT-V types did not differ ($p > .05$).

With criminality, the main effect of size was significant, $F(4, 171) = 8.07$, $p < .001$, $\eta = .159$. Paired comparisons revealed that individual composites received higher ratings than 8- and 16-morphs ($p < .05$). No other differences were significant ($p > .05$). The main effect of system was also significant, $F(2, 171) = 28.21$, $p = .006$, $\eta = .248$. E-FITs received higher criminality ratings than both EFIT-V types ($p < .05$), which did not differ ($p > .05$).

With distinctiveness, the main effect of size was significant, $F(4, 171) = 22.38$, $p < .001$, $\eta = .344$. Paired comparisons revealed that individual composites received higher ratings than all morph types ($p < .05$). No other differences were significant ($p > .05$). The main effect of system was also significant, $F(2, 171) = 6.93$, $p = .001$, $\eta = .075$. E-FITs received higher ratings than both EFIT-V types ($p < .05$), which did not differ ($p > .05$).

With intelligence, the main effect of size was significant, $F(4, 171) = 12.67$, $p < .001$, $\eta = .229$. Paired comparisons revealed that 4-, 8- and 16-morphs received higher ratings than individual composites ($p < .05$). No other differences were significant ($p > .05$). The main effect of system was also significant, $F(2, 171) = 25.78$, $p = .001$, $\eta = .232$. E-FITs received lower ratings than both types of EFIT-V ($p < .05$), which did not differ ($p > .05$).

A multiple regression examined whether the characteristic ratings and morph-size (individual composite, 2-, 4-, 8-, 16-morph) predicted participant-acquaintance similarity ratings. The final model was significant, $F(6, 179) = 61.84$, $p < .001$ ($R^2 = .675$, adjusted $R^2 = .664$). The raw and standardised regression coefficients together with their bivariate correlations are displayed in Table 5. Attractiveness contributed the strongest weight to the model, followed by morph-size, and distinctiveness.

Table 5 about here

Discussion

In Experiment 4 there was a positive relationship between morph-size and target-composite similarity ratings. The strongest positive significant morph-superiority effects were found when comparing 2-morphs to individual composites, and 4-morphs to 2-morphs. Despite moderate effect sizes, there were no differences between 4-, 8- and 16-morphs. It is not possible to conclude though that increasing the number of composites making up a morph beyond four would have no effect on recognition rates. Indeed, the highest similarity ratings were provided to the 16-morphs in all conditions, suggesting that a combination of at least 16, and although not tested here, possibly more, might have the greatest likelihood of being identified. The lack of significant effects may partly be a consequence of the number of images assessed being halved at each successive level of morph-size, meaning statistical power was too low for determining whether increasing morph-size beyond four increases target-similarity.

As predicted, and consistent with proposals for morph-attractiveness and morph-prototype effects, increasing morph-size was positively related to attractiveness, and negatively related to criminal demeanour and distinctiveness. Morph-size was also negatively related to perceptions of age and, positively to intelligence. The latter effects can be explained as the morph-attractiveness effect endowing a generally positive disposition.

There were no differences between the colour and greyscale EFIT-Vs, matching previous research finding no influence of colour on composite quality (Frowd et al., 2006). Nevertheless, in comparison to EFIT-Vs, E-FITS received higher distinctiveness and criminality, and lower intelligence and attractiveness ratings. These differences may relate to the differences in the holistic production process of EFIT-V, and the feature-based process of E-FIT. Nevertheless, consistent with Experiment 1, these effects did not impact on target-composite similarity ratings, as there were no between-system differences on this measure.

The results of Experiment 4 are limited by the use of a single 21-year-old target-actress. Hasel and Wells (2007) suggest that the morph-superiority effect is strongest with attractive faces. Further research should be directed at examining whether these findings would be replicated with faces of different ages, gender and attractiveness.

General Discussion

Consistent with previous research, the experiments reported here demonstrated that morphing EFIT-Vs and E-FITs constructed by the same or different witnesses, increases their likeness to targets, and correct recognition likelihood by someone familiar with the person depicted. Experiment 4 revealed that this morph-superiority effect is robust with 2-morphs, and that increasing morph-size to at least a 4-morph, and possibly up to a 16-morph has cumulative advantages. In addition, Experiment 3 demonstrated that vertical stretch enhances EFIT-V recognition rates, particularly useful if only one composite can be created. However, a combination of both morph and stretch provided no additional advantage, probably a consequence of there being a ceiling as to the benefit of averaging processes. As such, further research could examine whether an alternative post-production technique, not relying on averaging, would enhance correct naming in combination with morphing or stretch (e.g., dynamic animated caricatures; Frowd et al., 2012). Indeed, dynamic animation can also be presented on websites or TV which may provide greater opportunities for publication.

Experiment 4 also revealed that increasing morph-size *reduces* composite distinctiveness, consistent with a morph-prototype effect proposed by Hasel and Wells (2007), who asked raters to provide 4-morphs and individual composite similarity ratings to targets and foils. The authors suggested that this might also increase incorrect naming rates, although there was no evidence of this in Experiments 1-3. Indeed, in Experiments 2 and 3 individual composites were incorrectly named more often than the morphed images. One explanation is that in Hasel and Wells' study, and in Experiment 4, participant raters were u*nfamiliar* with the targets. Composites are designed to be type-likenesses and not exact portraits, and as such, morphing reduces distinctiveness, and thus in Hasel and Wells' study, similarity to the unfamiliar but physically similar foils. In contrast, in Experiments 1-3, participants were required to name composites of familiar targets from a small pool of actors. The reduction in distinctiveness that accompanies morphing was unlikely to have produced an image that becomes more similar to one of the alternative candidates. Nevertheless, these results suggest that morphed composites may help someone familiar with the suspect recognise their image and to rule out non-targets, although the results of Experiment 4, support Hasel and Wells' (2007) proposals that someone unfamiliar with the suspect might also be more likely to provide an incorrect name.

Previous research has found that holistic-system composites tend to be better recognised than feature-based composites (Frowd et al., 2007b). However, no between-system differences were found in Experiment 1 or 4. This is probably because construction occurred in time-limited training courses at which composite quality was not the primary aim.

The research reported in this paper does have limitations, most notably the number of low numbers of target actors used in each experiment, and that the viewing of videos before composite creation by constructors and the conditions of composite viewing by namers and raters did not match the normal experience of witnesses. Nevertheless, vertical stretch (Frowd *et al*., 2013) and morphing (Frowd *et al*., 2006) have now been shown to benefit facial composites produced by a number of different composite systems, and it is likely that they will provide a universal advantage with all systems. The highest target similarity ratings were given to 16-morphs, suggesting that this procedure could be followed in the probably rare occurrence of there being that many witnesses to a crime.

Morphing composites has become accepted police practice in the UK (Association of Chief Police Officers, 2009) as long as the investigating officer is certain that the composites depict the same suspect. This would never be a problem with within-witness morphs, and indeed witnesses often make more than one composite, if they believe the next would be of better quality. Vertical stretch has been used in police investigations, and as these results show the technique can improve composites created using different systems. Indeed, unlike with morphing, vertical stretch can be achieved with a single facial composite. However, as Frowd et al. (2013) have pointed out vertical stretch can make images appear 'silly,' which might reduce acceptance by the criminal justice system. Nevertheless, the authors note that if the image is accompanied by a message such as "viewing the composite sideways, to give a different perspective, may help you to recognise the face" (p.11), the natural scepticism of police officers to display images in the media in this manner is reduced.

Implications for practice

• The four experiments described here demonstrate that the recognition likelihood of facial composites can be enhanced by the three post-production techniques – within-witness and between-witness morphing, and vertical stretch.

• Police investigators should therefore consider using one of the techniques prior to distribution and publication.

• In investigations with more than one witness, between-witness morphing is recommended. If there is only one witness, a decision could be made to create a vertically-stretched composite, or to ask the witness to create two composites for a within-witness morph

• Investigators should also be aware that the publication of morphed composites may result in an increase in the number of innocent suspects named.

References

Association of Chief Police Officers, ACPO (2009). Facial Identification Guidance. Produced by the National Policing Improvement Agency. Available to download at www.acpo.police.uk/documents/crime/2009/200911CRIFIG01.pdf

Albright, L., Kenny, D.A. and Malloy, T.E. (1988). "Consensus in personality judgements at zero acquaintance", *Journal of Personality and Social Psychology*, Vol. 55, pp. 387-395.

Bruce, V., Ness, H., Hancock, P.J.B., Newman, C. and Rarity, J. (2002), “Four heads are better than one. Combining face composites yields improvements in face likeness”, *Journal of Applied Psychology,* Vol. 87, pp. 894-902.

Bruce, V. and Young, A. (1986), “Understanding face recognition”, *British Journal of Psychology,* Vol. 77, pp. 305-327.

Davies, G.M., and Valentine, T. (2007), “Facial composites: Forensic utility and psychological research”, In R.C.L. Lindsay, D.F. Ross, J.D. Read, and M.P. Toglia, *Handbook of eyewitness psychology: Volume* 2 - *Memory for people* (pp. 59-83), Mahwah: LEA.

Davis, J.P., Gibson, S. and Solomon, C. (2014). The positive influence of creating a holistic facial composite on video lineup identification. *Applied Cognitive Psychology*, Vol. 28, pp. 634-639.

Davis, J.P., Sulley, L. Solomon, C., & Gibson, S. (2010), “A Comparison of individual and morphed facial composites created using different systems”, in G. Howells, K. Sirlantzis, A. Stoica, T. Huntsberger, and A.T. Arslan (Eds.), *2010 IEEE International Conference on Emerging Security Technologies* (pp. 56 – 60), Canterbury: IEEE.

Ellis, H.D., Shepherd, J. and Davies, G.M. (1979), “Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition”, *Perception,* Vol. 8, pp. 431-439.

Fodarella, C., Kuivaniemi-Smith, H.J. and Frowd, C.D. (2015), “Detailed procedures for forensic face construction”, Journal of Forensic Practice.

Frowd, C. (2015), “Facial composites and techniques to improve image recognisability”, in T. Valentine, and J.P. Davis (Eds.), *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV,* Oxford: Wiley-Blackwell.

Frowd, C.D., Bruce, V., McIntyre, A. and Hancock, P.J.B. (2007a), “The relative importance of external and internal features of facial composites”, *British Journal of Psychology,* Vol. 98, pp. 61-77.

Frowd, C. D., Bruce, V., Ness, H., Bowie, L., Thomson-Bogner, C., Paterson, J., McIntyre, A. and Hancock, P.J.B. (2007b), “Parallel approaches to composite production”, *Ergonomics*, Vol. 50, pp. 562-585.

Frowd, C.D., Bruce, V., Plenderleith, Y., & Hancock, P.J.B. (2006). Improving target identification using pairs of composite faces constructed by the same person. *IEE Conference on Crime and Security* (pp. 386-395). London: IET.

Frowd, C., Carson, D., Ness, H., Richardson, J., Morrison, L., Mclanaghan, S. and Hancock, P. (2005), “A forensically valid comparison of facial composite systems”, *Psychology, Crime & Law,* Vol. 11, pp. 33-52.

Frowd, C.D., Hancock, P.J.B., Bruce, V., McIntyre, A.H., Pitchford, M., Atkins, R., Webster, A., Pollard, J., Hunt, B., Price, E., Morgan, S., Stoika, A., Dughila, R., Maftei, S. and Sendrea, G. (2010), “Giving crime the ‘evo’: catching criminals using EvoFIT facial composites”, *IEEE International Conference on Emerging Security Technologies (*pp. 36-43). Canterbury, IEEE.

Frowd, C.D., Jones, S., Forarella, C., Skelton, F.C., Fields, S., Williams, A., Marsh, J., Thorley, R., Nelson, L., Greenwood, L., Date, L., Kearley, K., McIntyre, A. and Hancock, P.J.B. (2013), “Configural and featural information in facial-composite images”, *Science & Justice*, Vol. 54, pp. 215-227.

Frowd, C.D., Atherton, C., Skelton, F.C., Pitchford, M., Bruce, V., Atkins, R., Gannon, C., Ross, D., Young, F., Nelson, L., Hepton, G., McIntyre, A.H. and Hancock, P.J.B. (2012), “Understanding the animated caricature advantage for facial composite images”, *Visual Cognition,* Vol. 20, pp. 1215-1241.

Frowd, C.D., Bruce, V., Plenderleith, Y. and Hancock, P.J.B. (2006), “Improving target identification using pairs of composite faces constructed by the same person", *IEE Conference on Crime and Security* (pp. 386-395), IET: London.

Hasel, L.E. and Wells, G.L. (2007), "Catching the bad guy: Morphing composite faces helps", *Law and Human Behavior,* Vol. 31, pp. 193-207.

Hole, G.J., Eaves, G.K. and Rasek, A. (2002), "Effects of geometric distortions on face-recognition performance, *Perception,* Vol. 31, pp. 1221–1240.

Klatzky, R.L., Martin, G.L. and Kane, R. (1982), "Semantic interpretation effects on memory for faces", *Memory and Cognition,* Vol. 10, pp. 195-206.

Langlois, J.H. and Roggman, L.A. (1990), "Attractive faces are only average", *Psychological Science,* Vol. 1, pp. 115-121.

Lohr, D. (2013), "Madeleine McCann Update: New Images Released Of Man Sought In Girl's Disappearance", *Huffington Post,* 14 October 2013, downloaded 6 July 2014 from, http://www.huffingtonpost.com/2013/10/14/madeleine-mccann-update_n_4096496.html

MacLin, O.H. and MacLin, M.K. (2004), "The effect of criminality on face attractiveness, typicality, memorability and recognition", *The North American Journal of Psychology,* Vol. 6, pp. 145-154.

Memon, A., Meissner, C.A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, & Law, 16,* 340-372.

Rhodes, G., Roberts, J. and Simmons, L.W. (1999), "Reflections on symmetry and attractiveness", *Psychology, Evolution, & Gender,* Vol. 1, pp. 279-295.

Rhodes, G. and Tremewan, T. (1996), "Averageness, exaggeration and facial attractiveness", *Psychological Science,* Vol. 7, pp. 105–110.

Saladin, M., Saper Z. and Breen L. (1988), “Perceived attractiveness and attributions of criminality: What is beautiful is not criminal”, *Canadian Journal of Criminology*, Vol. 30, pp. 251-259.

Santos, I. and Young, A. (2005), “Exploring the perception of social characteristics in faces using the isolation effect”, *Visual Cognition*, Vol. 12, pp. 213-247.

Solomon, C.J., Gibson, S.J. and Maylin, M. (2012), “EFIT-V: Evolutionary algorithms and computer composites”. In C.Wilkinson and C.Rynn (Eds.), *Craniofacial identification* (pp.24-41), Cambridge: Cambridge University Press.

Tanaka, J.W. and Farah, M.J. (1993), “Parts and wholes in face recognition”, *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, Vol. 46A, pp. 225-245.

Valentine, T. (1991), “A unified account of the effects of distinctiveness, inversion and race in face recognition”, *Quarterly Journal of Experimental Psychology*, Vol. 43A, pp. 161-204.

Valentine, T., Darling, S. and Donnelly, M. (2004), “Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces”, *Psychonomic Bulletin & Review*, Vol. 11, pp. 482–487.

Valentine, T., Davis, J.P., Thorner, K., Solomon, C. and Gibson, S. (2010), “Evolving and combining facial composites: Between-witness and within-witness morphs compared”, *Journal of Experimental Psychology: Applied*, Vol. 16, pp. 72-86.

Zebrowitz, L.A., Hall, J.A., Murphy, N.A. and Rhodes, G. (2002), “Looking smart and looking good: Facial cues to intelligence and their origins”, *Personality and Social Psychology Bulletin*, Vol. 28, pp. 238-249.

Table 1: Naming rates in Experiment 1

	Correct Naming Rates				Incorrect Naming Rates			
	4-Morph		Ind. Comp		4-Morph		Ind. Comp	
Actor	E-FIT	EFIT-V	E-FIT	EFIT-V	E-FIT	EFIT-V	E-FIT	EFIT-V
A	21.5%	26.3%	10.8%	19.2%	26.2%	21.1%	26.2%	30.8%
B	40.0%	19.4%	30.5%	16.9%	17.8%	14.9%	6.8%	15.4%
C	43.3%	42.9%	17.5%	19.3%	16.7%	30.6%	21.1%	28.1%
D	33.3%	33.9%	23.1%	26.3%	3.9%	8.5%	3.8%	10.5%
Total	33.9%	29.7%	20.2%	20.3%	16.7%	18.1%	15.0%	20.8%

Table 2: Conditionalised Naming Rates (CNR) in Experiment 2

	Correct CNR (%)			Incorrect CNR (%)		
Actor	4-Morph	Combined	Best	4-Morph	Combined	Best
A	80.4	39.2*	30.0*	8.7	16.6	8.0
B	8.2	17.3	5.3	61.2	23.5*	42.1
C	80.0	44.4*	12.2*	12.0	15.9	42.9*
D	86.0	66.3*	72.0	6.0	14.0	12.0
E	73.5	35.6*	38.3*	4.1	18.1*	2.1
F	19.0	14.4	27.3	23.8	33.3	27.3
G	33.3	26.8	26.5	35.6	83.0*	28.6
H	48.0	40.0	26.1*	14.0	19.4	28.3
Total	54.3	35.8*	30.6*	20.5	28.1*	23.3

** Significant comparisons between 4-morphs and individual composites (p < .05)*

Table 3: Conditionalised Naming Rates (CNR) in Experiment 3

Actor	Correct CNR (%)				Incorrect CNR (%)			
	Ind. Comp	Morph	Stretch	Stretch & Morph	Ind. Comp	Morph	Stretch	Stretch & Morph
Total	7.4	27.8	18.8	30.1	16.5	7.1	5.1	9.1

Note: Data from individual targets are pooled as no actor-based effects were significant

Table 4: Mean ratings as a function of composite size and composite system (E = E-FIT, V = EFIT-V)

	Individual		2-morph		4-morph		8-morph		16-morph	
	E	V	E	V	E	V	E	V	E	V
Similarity	2.13	2.44	2.85	3.14	3.60	3.54	3.98	3.58	4.22	3.79
Age	25.0	24.7	24.6	23.5	23.4	22.3	23.5	21.8	22.9	22.4
Attractiveness	3.14	3.87	3.84	4.61	4.43	5.15	4.98	5.40	5.51	5.25
Criminality	4.87	5.91	5.03	6.19	5.15	6.14	5.31	6.73	6.38	6.11
Distinctiveness	5.85	5.11	5.14	4.57	4.73	4.08	4.74	4.32	4.59	4.41
Intelligence	3.96	5.47	4.45	5.89	4.69	6.40	4.85	6.29	6.04	5.70

Note: The data from colour and greyscale EFIT-Vs are pooled as no effects were significant

Table 5: Correlation coefficients and multiple regression with composite-target similarity ratings (similar) as the dependent variable and attractiveness (attract), composite size (size), distinctiveness (dist), intelligence (intel), Criminality (crim) and perceived age (age) as predictors

	Crim	Intel	Dist	Size	Attract	Similar	B	SE B	β	t	p
Constant							2.45	0.869			
Attract						.79*	0.490	0.071	0.599	6.93	< .001
Size					.49*	.54*	0.049	0.013	0.193	3.81	< .001
Dist				-.34*	-.72*	-.68*	-0.244	0.065	-0.246	-3.77	< .001
Intel			-.52*	.24*	.73*	.47*	-0.087	0.061	-0.114	-1.42	.159
Crim		.79*	-.41*	-.24*	-.65*	-.39*	0.067	0.076	0.063	0.88	.380
Age	.33*	-.35*	.54*	-.34*	-.49*	-.47*	-0.017	0.021	-0.042	-0.81	.421

*$p < .001$