Parental reference photos do not always improve the accuracy of forensic age progressions

Provenzano, Andrew, Lampinen, James Michael, Erickson, William, Frowd, Charlie and Mahoney, Greg

Provenzano, Andrew, Lampinen, James Michael, Erickson, William, Frowd, Charlie ORCID: 0000-0002-5082-1259 and Mahoney, Greg (2020) Parental reference photos do not always improve the accuracy of forensic age progressions. Science and Justice, 60 (6). pp. 522-530. ISSN 1355-0306

Parental reference photos do not always improve the accuracy of forensic age progressions.

Abstract

During long-term missing children cases, forensic artists construct age-progressions to estimate the child's current appearance. It is commonly believed that incorporating information about the child's biological relatives is critical in accurately estimating the child's current appearance. However, some evidence suggests that predicting appearance based on inheritance of features may be error prone. The present studies examine whether age-progressions constructed with the aid of a biological reference photos led to better recognition than those constructed without a biological reference. We also investigated whether there would be any variation depending on the age-range of the age-progressions. Eight professional forensic artists created age-progressions based upon photographs provided by each of our eight targets. Half of their age progressions with the aid of parental reference photos and half without parental reference photos. Furthermore, half were age-progressed across a longer age-range (5-20 years) and half covered a shorter age-range (12-20 years). In Experiment 1 similarity scores were higher over shorter age-ranges. Further, across longer age-ranges age-progressions created with the aid of a parental reference were lower than those without a reference. In Experiment 2 recognition performance was higher across shorter age-ranges. Additionally, across longer age-ranges age-progressions created with the aid of a parental reference were recognized worse than those without a reference. These results suggest that in long-term missing person cases, forensic artists may benefit from not relying on biological references. Finally, consistent with previous research (e.g. Lampinen et al., 2012) age-progressions provided no benefit over using outdated photographs.

Keywords: Missing persons; Forensic imaging; Age progression; Parental reference

Funding: This work was supported by the National Science Foundation grant number 1155207.

1. Introduction

As of December 31, 2018, the National Crime Information Center (NCIC) reported there were approximately 85,459 active missing persons cases in the United States. Approximately half of those cases involved juveniles under 21 years of age [1]. While the vast majority of these cases are resolved within 24 hours; an important subset of cases involve children missing for several years [2]. In these cases, authorities may try to estimate the current appearance of the missing person. One common approach used is to authorize the construction of a forensic age-progression of the missing individual. Forensic age-progressions attempts to update the appearance of the missing person from the time they disappeared to their most likely current appearance.

Age-progressions are created through a variety of methods, including by forensic artists, computer software, or a combination of these approaches [3]. In the United States, the predominant approach is construction via forensic artists [4]. With this in mind, we will primarily discuss that approach to age-progressions. Furthermore, while understanding the processes of age-progressions across all age-ranges are important, we will primary focus on the approaches to the construction of juvenile age-progressions.

When constructing juvenile age-progression, forensic artists rely on their knowledge about typical facial growth norms and/or family resemblance to predict the missing person's appearance. The first method involves forensic artists using their extensive knowledge about typical craniofacial growth and development norms, including knowledge that faces typically lengthen from infancy to adolescence [5,6]. Most of the understanding about juvenile facial growth patterns, including cranial, nasomaxillary, and mandibular growth derive from work by dental surgeons [7]. A basic understanding about typical growth across these areas allow forensic artists to make educated decisions when predicting the growth of these critical facial features. For example, for European Caucasian males, the average width of a nose is 28.39 mm at age 5 and 32.57 mm at age 12. Using this information, a forensic artist may increase the width of the nose by 15% (i.e. (32.57-28.39)/28.39) when age-progressing a photograph from five to twelve years-old. While the use of growth norms is relatively non-controversial within the forensic artist community, there may be several limitations to exclusively relying on this approach.

One potential drawback to this approach is it fails to account for the wide variability in facial growth at an individual level [8]. In a longitudinal study investigating lower jaw growth, children in their early teens experienced average yearly growth of 0.21 cm/year and 0.22 cm/year for boys and girls respectably. However, there was a relatively wide standard deviation of 0.12 cm/year and 0.13 cm/year for boys and girls respectably. This same trend is consistent across other facial landmarks [9]. Additionally, even related face structures do not grow in a unitary fashion and thus creates a constant imbalance [7]. Therefore, a forensic artist relying solely on these averages, may greatly over or under-shoot the actual individual growth of these facial features. Another potential drawback to this approach is that individuals may display different development rates than the "average" child [10] For instance, a six-year-old boy may have the bone development of a seven-year-old which makes growth prediction more difficult. Given these limitations, growth predictions may allow for a fairly accurate population prediction, but this may not translate to a successful prediction on the individual level. This is especially true for facial features with a relatively wide standard deviation relative to its mean [10].

The second approach forensic artists may incorporate while constructing age progressions is termed genetic prediction. Typically, a forensic artist starts this process by conducting an extensive interview with family members of the missing persons, where they ask for input from the family on whether they see any particular familial resemblance between the features of the missing person and any biological relatives. Besides the interview, forensic artists base their genetic prediction on a set of reference photographs obtained from the biological relatives. A typical set contains photographs of both biological parents and siblings, but may also include photographs of grandparents, aunt, uncles. Forensic artists seek to acquire biological reference photographs which are of high resolution and frontal facing. Furthermore, the typical set of biological reference photographs include both those of the references at the age of the missing person's disappearance and current age. For a child who went missing at the age of 5 and been missing for 15 years, the artist would attempt to obtain photographs of family members at both age 5 and 20, or as close as possible to those desired ages. While constructing the age-progression, the forensic artist is looking for similar facial features between the child and biological relative. For instance, if at the age of disappearance, the child has a similar nose to their grandmother, ear of their father, and chin of their mother, the forensic artist would base their estimate on those respective features.

There is evidence that incorporating genetic prediction has some validity. First, there is some evidence that genetically encoded factors may affect craniofacial growth [11]. Additionally, a study by DeBruine et al. [12] suggest that biologically related individuals share more common features with each other than non-biological related individuals. When tasked with rating the appeared similarity between two adult photographs, participants rated siblings as more similar to each other than non-siblings regardless of gender [12]. Furthermore, children have a greater resemblance to both their biological father and mother than other similar looking non-related adults [13]. Further evidence of the importance that genetic factors in facial growth comes from the machine learning literature where machines have demonstrated an ability to identify shared kinship greatly above chance by relying on algorithms that compare certain critical facial features [14]. These findings provide evidence that basing age-progressions on family resemblance may be beneficial to increases the predictive ability of the age progression.

However, too much reliance on familiar resemblance may hinder the representativeness of the age-progression due to the level of unpredictability of facial similarity across age-ranges. In a study by Alvergne et al [13], parents rated newborns' facial features as more similar to their mother regardless of the child's sex. However, as the child aged, parents attributed more facial similarity to the father, especially for male faces. One potential reason is that males typically have more mature features that are less apparent until later in life. Hence, a forensic artist may misattribute features of the child at an early age.

1.2. Comparison to Outdated Photographs

Another method that authorities implement in efforts to recover missing persons is the release of photographs at the last known appearance. There is evidence that outdated photographs produce recognition at greater than chance levels. Seamon [15] demonstrated in a series of experiments that humans may have what is termed, "bidirectional dynamic facial recognition". Participants studied a set of faculty photographs taken in 1974 and completed a recognition test with a set of photographs of the same faculty members from 1966. Impressively, participants recognized faculty members well above chance. In a second experiment, participants displayed similar performance after studying the older photographs and having a recognition test on the newer photographs.

A study conducted by Bruck, Cavanagh, and Ceci [16], investigated the ability of former classmates to recognize one another on their 25th reunion. The authors mailed participants a booklet asking them to match old yearbook photographs with current photographs of visually similar individuals in their mid-forties. Participants were either old classmates or from another country who were unfamiliar with the classmates. Former classmates matched 49% of the classmates correctly, while 33% of those who were unfamiliar matched the younger and older images. Based on the study design, the chance recognition rate was 10% and therefore indicates that even while unfamiliar that humans have enhanced facial recognition ability.

In order to suggest that implementing age-progressions in a particular missing person case is beneficial, we first need to understand its effectiveness in the context of alternative methods used to aid in the recovery of missing persons. A common alternative approach is using an outdated photograph of the missing person from the time right before the child went missing. While this approach is logical in short-term cases, over time outdated photograph may be less likely to accurately represent the appearance of the missing person. In a series of experiments comparing age-progression to outdated photographs, Lampinen, Miller, and Dehon [17] hired forensic artists to age-progress target photographs from the age of seven to twelve years old. Participants studied one of three types of missing person posters (age-progressed, outdated, and current images) and were asked to be on the lookout for the missing person on the subsequent team-sorting task. Recognition rates across all conditions was above chance, with those viewing current photographs outperforming both the age-progression and outdated photograph condition. Importantly, there was no difference in performance between those in the outdated and age-progression conditions. Furthermore, Lampinen, Miller, and Dehon [17] made use of an age progression used in the real-life recovery of Jaycee Dugard and found no significant differences between the age-progression and outdated photograph.

Research by Charman and Carol [18] suggest that when age-progressions are computer generated, they may provide a worse match to memory than outdated photographs. In this study, participants viewed either (1) outdated target photographs, (2) age-progressions created via APRIL’s online age-progression system (www.ageme.com), (3) or both the age-progression and outdated photograph. After viewing the photographs, participants were asked to report whether the targets appeared on a subsequent recognition task. Importantly, participants who studied the outdated photographs had better recognition rates than participants who studied both or participants who just studied age-progressions. One argument made is that age-progressions may provide extra cues that increases the number of plausible faces that can match the appearance of the missing person. It is important to note that these studies all focus on the recognition rates of unfamiliar faces. There may be some difference in how age-progressions of familiar faces are processed. Additionally, these studies compare age-progressions to outdated photographs across a relatively narrow age-range.

1.3 Current Study

Previous research described above suggest that age-progressions don’t provide much benefit to enhancing the likelihood of recovering the missing person. One way to better understand the effectiveness of age-progressions is to look at strategies that might improve age-progressions. One factor that may influence the quality of the age-progression is the strategies used by the forensic artists. One notable strategy is the emphasis that they place on the parental reference photographs.

Based on the evidence provided above, there is no clear evidence that familial resemblance helps or hinders the representativeness of the age-progression. Therefore, in the present experiments, we look to investigate this question in a number of ways. In Experiment 1, we tested whether having access to a biological photographic reference allows forensic artists to create age-progressions with that are more discriminable between targets than foils. In

Experiment 1, age-progressions are classified as more discriminable when the similarity between the age-progression and a target is greater from that of a description-matched foil. In Experiment 2, we tested whether age-progressions created using a biological reference produced a higher recognition rate using a forced-choice recognition paradigm.

2. Method

2.1 Participants

One hundred and two introductory psychology students (Age: *M* = 19.01 years) participated in the following study in exchange of fulfilling a course requirement. The majority of participants were female (n = 75, 73.35%). The majority of participants were Caucasian (81.3 %), followed by African American (5.6%), and Hispanic (5.4%).

2.2 Materials

We recruited eight Caucasian adults (four female) to provide images of themselves at ages 5, 12, and 20 years, as well as photographs of both of their biological parents at those same ages. All eight adult volunteers were paid $200 for consenting use of their photographs for a series of studies. To construct stimuli for our study, we recruited four forensic artists and had them create age-progressions of our targets (i.e. volunteers) across various age-ranges. Each artist received images of two 5-year-old males (one including parental reference) and two 5-year-old females (one including a parental reference) and were asked to age progress the image of the 5-year-old target to 20 years. Each artist also received images of two 12-year-old males (one including a parental reference) and two 12-year-old females (one including a parental reference) and were asked to age progress the image of the 12-year-old target to 20 years. Each of the four forensic artists constructed eight age-progression, one for each of the eight volunteers. The creation of the age-progressions was counterbalanced by age-range (5-20;12-20) and whether a parental reference was provided (Yes; No). For example, volunteer 'A' had one age progression constructed with a parental reference at age 5, one without a parental refence at age 5, one with a parental reference at age 12, and one without a parental refence at age 12. In the study, each volunteer had 4 age progressions, and each one of these age progressions was constructed out by a separate artist. In total, there was a total of 32 age-progressions created (four per target). To construct the age progressions with a parental refence, the forensic artists had access to photographs of the mother and father at both the age of the child went missing and the child's current age. In our study, each age-progression was paired with either a 20-year old photograph of the target or a description matched foil.

2.3 Procedure

Participants completed an online experiment on Qualtrics using their personal computer. Participants completed a demographic questionnaire and were provided instructions for the main task. Participants were instructed that they would see pairs of faces and asked to rate the similarity of the image on the left (the age progressions) to the image on the right (target or foil). Participants were not told any specific information about the pairings, merely to rate for similarity. Participants rated similarity on a scale from 1 (extremely dissimilar) to 7 (extremely similar) with 4 representing 'neither similar nor dissimilar'. Participants were told that some faces would appear more than once but go on their first instinct when rating similarity. Face pairings were presented in a random order for each person. Participants completed 64 trials (32 target trials, 32 foil trials) across all levels of the independent variables. The similarity-rating task took approximately 15 minutes, after which participants were debriefed.

3. Results

3.1 Model Comparison

Repeated-measure linear mixed modeling of similarity scores was conducted using the R package *lme4* [18]. All models were fit using Maximum Likelihood and the corresponding t-tests used Satterthwaite approximations to degrees of freedom. The initial base model included random intercepts for *Subjects*, *Artists* (4 artists), *Item*s (32 items), and *Order*. We examined the clustering of the random intercepts by calculating the ICC for Subjects (ICC = .12), Artists (ICC = .00), Items (ICC = .12), and Order (ICC = .00). The ICC suggests that the model benefitted by including *Subjects* and *Age-progressions* as random intercepts, but not *Artists* or *Order*. Therefore, our final base model included just *Subjects* and *Age-progressions* as random intercepts.

For the test model, we included three dummy-coded fixed factors in the model, (a) *Parental Reference* (Yes, No), (b) *Age-Range* (12-20, 5-20), (c) *Target vs Foil* (Target, Foil) and all corresponding interactions. The *Target vs Foil* fixed effect functioned as a difference score measure for our analyses. We treated *Subjects*, and *Age-progressions* (32) as random effects factors allowing their intercept the vary randomly. Compared to the base model, the model fit was significantly improved, χ^2 (6) = 78.113, p < .001 and reduced unexplained variance by 2.2%.

3.2 Fixed Effects

Of particular interest to our hypotheses were the two-way and three-way interactions including the *Target vs Foil* factor. Testing these interactions allowed us to test whether there was any change in discriminability across the other predictor variables. There was a significant two-way interaction between Age-Range and the Target vs Foil factor on similarity, t (6391) = --2.79, p = .005 (See Figure 1). A follow-up simple effects test revealed that on 12-20 trials, targets (M = 4.22) were rated as more similar to the age-progression than foils (M = 3.80) (p <

.001). Meanwhile, on 5-20 trials, targets (*M* = 3.78) were rated as no more similar to the age-progression than foils (*M* = 3.90) (*p* = .058).

There was a non-significant two-way interaction between *Parental Reference* and *Target vs Foil* on similarity score, *t* (6391) = - 1.25, *p* = .25 (See Figure 1). However, a follow-up simple effects test revealed that on non-parental reference trials, targets *(M* = 4.22) were rated as more similar to the age-progression than foils (*M* = 4.01) (*p* = .001). Meanwhile, on 5-20 trials, targets (*M* = 3.78) were rated as no more similar to the age-progression than foils (*M* = 3.68) (*p* = .124).

The two-way interactions were qualified by a significant three-way interaction between Parental Reference, Age-Range, and Target vs Foil on similarity score, *t* (6391) = - 3.02, *p* = .003 (See Figure 2). A follow-up simple effects test revealed that for 5-20-year-old age-progressions with a parental reference, target photographs (*M* = 3.48) were rated as less similar to the age-progression than foils (*M* = 3.77) (*p* = .001). However, when constructing age-progressions from 12-20 with a parental reference, target photographs (*M* = 4.37) were rated as more similar to the age-progression than foils (*M* = 4.01) (p < .001). Meanwhile, for 5-20-year-old age-progressions without a parental reference, target photographs (*M* = 4.08) were rated comparable to foils (*M* = 4.03) (*p* = .554). However, when constructing age-progressions from 12-20 without a parental reference, target photographs (*M* = 4.08) were rated as more similar to the age-progression than foils (*M* = 3.60) (*p* < .001). All other predictors were not significant. For the full model refer to Table 1.

3.3 Artist Effects

We ran exploratory analyses investigating if similarity ratings varied by the forensic artist. For *target* photographs, there was a significant main effect of artist on similarity scores, *F* (3,3260) = 3.17, $\eta_p{}^2$ = .003, *p* = .023. However, a follow-up Tukey's HSD test revealed no differences between the artists. For *foil* photographs, there was a significant main effect of artist on similarity scores, F (3,3260) = 19.98, $\eta_p{}^2$ = .018, $p < .001$. Follow-up analyses revealed that Artist 2's similarity ratings were significantly higher than all other forensic artists. Overall, we feel that these differences accounted for by the natural variation in how forensic artists construct age-progressions. We do not believe that these differences harm the generalizability of our findings.[1]

4. Experiment 1 Discussion

We were interested in investigating whether forensic artists would benefit from having a parental reference while constructing an age-progression. The results from this study suggest that having a parental reference does not increase the discrimination between targets and foil compared to not having a parental reference. Furthermore, there may be a detrimental effect of using a biological reference when age-progressing across a wider age-range.

5. Experiment 2

Experiment 1 examined the effectiveness of various artist rendered age-progressions in the context of how similar they were to the target individual. In Experiment 2, we are looking to see if the results found in Experiment 1 would extend to a forced-choice recognition paradigm.

6. Method

6.1 Participants

Four-hundred and eighty-seven introductory psychology students (Age: M = 19.44 years) participated in the following study in exchange of fulfilling a course requirement. The majority

---

[1] A reviewer of this manuscript also suggested comparing similarity scores between each of volunteer targets. We found some differences between the similarity scores between the volunteers, but nothing that was outside what would be expected given the variability of our stimuli. Additionally, although such a comparison is interesting, it is not within the current scope of this article but would be an interesting future direction.

of participants were female (n = 294, 60.02%). The majority of participants were Caucasian (83.4 %), followed by and Hispanic (6.8%), and African American (4.3%).

6.2 Materials

We used the same materials as we did in experiment 1. This includes the age-progressions and photographs used in the study.

6.3 Procedure

Participants completed an online experiment on Qualtrics using their personal computer. Participants completed a demographic questionnaire and were then read instructions for the perceptual-recognition task. Participants viewed eight target individuals. For each participant, all eight targets were randomly displayed using one of seven presentation methods (current, 5-year old outdated, 12-year old outdated, 5 to 20-year-old age-progression created with parental reference, 5 to 20-year-old age-progression, 12 to 20-year-old age-progression created with parental reference, 12 to 20-year-old age-progression) and for each image, they were shown 2 photographs (one of the target and one of a description-matched foil) and were asked to pick the photograph that best matched the age progression. Participants were not told any information about the pairings, merely “Which of the following people do you think is the individual depicted in the age progression above?”. After selecting the photograph that they thought best matched age-progression, participants were asked, “How confident are you in your answer?”. Confidence was rated on a 5-point scale ranging from 1 (There is a 50-60% chance I am correct) to 5 (There is a 90-100% chance I am correct). We chose 50% as the starting point for confidence because given the design of the experiment, they had a minimum of 50% chance to get the questions correct. Face pairings were presented in a different random order for each person. The perceptual recognition task took approximately 10 minutes, after which participants were debriefed.

7. Results

7.1 ANOVA (3 X 2 Accuracy)

A 3 (Photo type: Outdated, Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on accuracy. Accuracy was calculated by the proportion of correct trials divided by total trials. This analysis revealed a significant main effect of photo type on accuracy, F (1,412) = 49.297, MSE = 1.098, $\eta_p{}^2 = .103$, $p < .001$. Post hoc tests using Tukey's HSD indicated that participants shown outdated photographs ($M = .842$, $SE = .013$) significantly outperformed participants shown age progressions with a parental reference ($M = .583$, $SE = .013$) and age progressions without a parental reference ($M = .604$, $SE = .013$), both $p < .001$. Participants' accuracy when shown age progressions with or without a parental reference did not differ from one another, $p = .257$. There was also a significant effect of age on accuracy, with participants shown 12-20-year-old ($M = .728$, $SE = .011$) photographs outperforming participants shown 5-20-year-old ($M = .625$, $SE = .011$) photographs, F (1,412) = 125.035, MSE = 2.902, $\eta_p{}^2 = .378$, $p < .001$. In addition to these main effects, there was a significant two-way interaction, F (2,412) = 4.602, MSE = .011, $\eta_p{}^2 = .022$, $p = .011$, suggesting that the effect of photo type depended on the age. A follow-up simple effects test revealed that having a parental reference for the 5-20-year-old age progressions hindered memory strength ($p = .017$), compared to not having a parental reference. However, this effect did not occur for the 12-20-year-old age progressions ($p = .756$).

7.2 ANOVA (*3X2 All memory strength)*

A 3 (Photo type: Outdated, Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on the participants' memory strength across all trials. Memory strength was calculated based upon the participant's confidence and accuracy during each trial. For example, if participants had a low level of confidence (There is a 50-60% chance I am correct) and were accurate we would add 1 to their memory strength. However, if they were inaccurate and had moderate confidence (There is a 70-80% chance I am correct), we would subtract 3 from their memory strength score. Memory strength could range between - 40 and + 40.

This analysis revealed a significant main effect of photo type on memory strength, *F* (1,412) = 44.290, MSE = 115.856, $\eta_p{}^2 = .177$, $p < .001$. Post hoc tests using Tukey's indicated that participants had a significantly stronger memory strength when presented with outdated photographs (M = 2.320, SE =.136) than both age progressions with a parental reference (M = .68, SE =.137) and age progressions without a parental reference (M = .819, SE =.138), both p < .001.The participants' memory strength for age-progressions with or without a parental reference did not differ, $p = .100$. There was a significant effect of age-range on memory strength, with participants shown 12-20-year-old (M = 1.67, SE = .48) age-progressions outperforming those shown 5-20-year-old (*M* = .89, SE = .478) age-progressions, *F* (1,412) = 23.412, MSE = 61.243., $\eta_p{}^2 = .054$, $p < .001$. In addition to these main effects, there was also a marginally significant interaction, *F* (2,412) = 5.087, MSE = 13.307, $\eta_p{}^2 = .024$, $p = .007$. A follow-up simple effects test revealed that having a parental reference for the 5-20-year-old age progressions hindered memory strength ($p = .012$), but this effect did not occur for the 12-20-year-old age progressions ($p = .141$).

7.3 ANOVA (*3X2 Most Confident)*

A 3 (Photo type: Outdated, Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on the participants' memory strength for the trials that participants were most confident (i.e. There is a 90-100% chance I am correct). We rationalize that this analysis is important because those who are confident are most likely to report the missing person. Memory strength was summed by adding 5 for each correct response and subtracting 5 for each incorrect response. We chose to add or subtract 5 to keep it consistent with the overall memory strength score. This analysis revealed a significant main effect of photo type on memory strength, *F* (1,412) = 23.425, MSE = 1122.381, $\eta_p{}^2 = .102$, $p < .001$. Post hoc tests using Tukey's indicated that participants had a significantly stronger memory strength when shown outdated photographs (*M* = 7.05, SE =.583) than those shown either age progressions with a parental reference (*M* = 2.00, SE =.585) or age progressions without a parental reference (*M* = 2.312, SE =.592), both p < .001. The participants' memory strength for age progressions with or without a parental reference did not differ from one another, *p* = .100. Furthermore, there was a significant effect of age-range on memory strength, with participants shown 12-20-year-old (M = 5.19, SE = .48) age progressions outperforming those shown 5-20-year-old (M = .2.429, SE = .478) age progressions, *F* (1,412) = 16.095, MSE = 1122, partial eta = .378, $p < .001$. In addition to these main effects, there was also a marginally significant two-way interaction, *F* (2,412) = 2.367, MSE = 113.398, $\eta_p{}^2 = .011$, *p* = .095.

7.4 ANOVA (*2 X 2 accuracy)*

A 2 (Photo type: Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on accuracy. This analysis revealed a significant effect of age-range on accuracy, *F* (1, 273) = 35.486, MSE = .922, $\eta_p{}^2 = .115$, $p < .001$, with participants being more accurate when shown

12-20-year-old age progressions (*M* = .651, SE = .014) than 5-20-year-old age progressions (*M* = .536, SE = .014). There was no effect of parental reference on accuracy, *F* (1, 273) = 1.150, MSE = .030, partial eta = .004, *p* = .285. However, there was a significant two-way interaction, *F* (1, 273) = 6.835, MSE = .178, $\eta_p{}^2$ = .024, *p* = .009. A follow-up simple effects test revealed that having a parental reference for the 5-20-year-old age progressions hindered performance (*p* = .009), but this effect did not occur for the 12-20-year-old age progressions (*p* = .279).

7.4 *2 X 2 (All memory strength)*

A 2 (Photo type: Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on the participants' memory strength across all trials. There was a significant effect of age-range on memory strength, with participants provided with 12-20-year-old (M = 1.22, SE = .135)) photographs outperforming those provided with 5-20-year-old (M = .279, SE = .134) photographs, *F* (1, 273) = 24.523, MSE = 61.486, $\eta_p{}^2$ = .082, *p* < .001.There was no effect of parental reference on accuracy, *F* (1, 273) = .519, MSE = 1.301, $\eta_p{}^2$ = .002, *p* = .472. There was a significant interaction, *F* (1, 273) = 7.936, MSE = 19.898, $\eta_p{}^2$ = .028, *p* = .005. A follow-up simple effects test revealed that 5-20-year-old age progressions constructed with the aid of a parental reference had worse memory strength (*p* = .012) than those created without the aid of a parental reference, but this effect did not occur for the 12-20-year-old age progressions (*p* = .141).

7.5 *2 X 2 (Most Confident)*

A 2 (Photo type: Age progression with parental reference, Age progression without parental reference) X 2 (Age Range: 5-20, 12-20) between-subject's ANOVA was conducted on the participants' memory strength for the trials participants were most confident. There was a significant effect of age-range on memory strength, for participants provided with 12-20-year-old (M = 3.634, SE = .585) photographs outperforming those provided with 5-20-year-old (M = .679, SE = .579) photographs, F (1, 273) = 12.880, MSE = 604.567, = .045, $p < .001$. There was no effect of parental reference on accuracy, F (1, 273) = .144, MSE = 6.754, $\eta_p{}^2 = .001$, $p = .705$. There was a significant two-way interaction, F (1, 273) = 6.835, MSE = .178, partial eta = .016, $p = .034$. A follow-up simple effects test revealed that 5-20-year-old age progressions constructed with the aid of a parental reference had marginally worse memory strength ($p = .075$) than those created without the aid of a parental reference, but this effect did not occur for the 12-20-year-old age progressions ($p = .218$).

7.6 *Comparison to Current Photographs*

An omnibus one-way ANOVA comparing each condition to proportion correct was significant, F (6, 481) = 86.073, MSE = 1.815, $\eta_p{}^2 = .518$, $p < .001$ (See Figure 3). Post hoc tests using Tukey's indicated that current photographs (M = .927, SE =.017) significantly outperformed all other conditions (p < .001) with the exception of the outdated 12-year-old photographs (M = .880, SE =.017).

7.7 *Comparison to Chance*

To investigate whether performance for each condition differed from chance, we conducted seven separate One-way ANOVA comparing each condition to chance performance (.50). All conditions ($p <.001$) differed from chance with the exception of the five-to-twenty age-progression created with a parental reference ($p = 1.00$).

8. General Discussion

In the present study, we manipulated the availability of familial reference images to forensic artists who then age-progressed children's photographs to adulthood. Use of family images, particularly parental images, is common practice among forensic artists who specialize in age-progression [20,21]. Use of parental references in this way has intuitive appeal because children’s faces are more likely to develop similarly to genetic relatives than unrelated adults. However, this practice has never been systematically examined by comparing age-progressions produced with such reference images to those produced without them. We made just such a comparison in both perceptual similarity (Experiment 1) and recognition (Experiment 2) tasks.

In Experiment 1, age-progressions created with the aid of parental reference photographs were rated no more similar to targets than age-progressions created without this aid. Furthermore, we found that across the wider age-range of 15 years, age-progressions based on a parental reference were judged less similar to the targets they are intended to represent. In Experiment 2, we found the same pattern using a forced-choice recognition paradigm. Additionally, across both experiments we failed to find evidence that supports the idea that forensic age-progressions provide any advantage to facial recognition over the use of outdated photographs. Instead, performance during the recognition task improved when using outdated photographs, especially for those age-progressions made under the shorter seven-year age gap from pubescence to young adulthood. The finding regarding worse recognition performance for age-progressed images compared to outdated photographs extends upon similar findings by Lampinen et al. (2012) who found a similar pattern of results testing retrospective and prospective memory performance. Next we interpret these results in light of the extant literature on age progressions and the broader context of forensic imaging.

8.1 ***Artist-Based Age Progressions***

These results contribute to the extensive line of findings that fail to observe any aggregate benefit to age-progression over outdated photographs in laboratory perception and memory tasks.

These include images generated by artists [17, 21] and those made by computer algorithms [18] over short age ranges within childhood and longer age ranges from childhood to adulthood. In response to these consistent findings, the current experiments were designed to identify an optimal scenario for artist-rendered age-progressions. This is important because age-progressed images in real cases are predominantly created by trained forensic artists and this is likely to be so for the foreseeable future. For these experiments, we expected providing the forensic artist with a biological reference would increase the similarity of age-progressions to those they are intended to depict. Somewhat to our surprise, the age-progressions created with a biological reference performed worse across longer age ranges (age 5 to 20 years of age). A potential explanation for this result may be that forensic artists relied too much on a mistaken familiar resemblance that disappeared as the child aged. In future studies, it may be beneficial to have forensic artists self-report how much they relied on using a biological reference when constructing the age progression.

While these results may be discouraging to researchers and professionals alike, there is convergent evidence from the unfamiliar face matching literature that may help explain it. Namely, someone unfamiliar with a person is less accurate when matching a face to another image of the same face or recognizing the face later. In one field study, supermarket cashiers falsely accepted the ID card of a similar-looking foil (e.g. same gender, ethnicity, similar age and hairstyle) in more than 50% of trials [22]. Jenkins, White, Van Monfort, and Burton [23] tasked participants with sorting 40 unfamiliar face photos into different identities. While the true number of identities was only two, more than half the participants sorted the 40 photographs into seven or more piles. In a similar study by Megreya [24], participants matched more photographs taken a few minutes apart (87.5%) than photographs taken a month apart (67.5%). Furthermore, changes in angle of viewing, expression posed, or context disproportionally harm unfamiliar face recognition compared to familiar face recognition [24]. Together, these findings illustrate the difficulty people have accounting for within-person variability of unfamiliar faces under ideal situations involving actual photographs spanning a negligible age difference. Introducing predictive manipulation of facial appearance across an age gap of many years (whether via algorithm, artist, or a combination of both) as an intermediate step produces more error than outdated images do because of the wide array of morphological changes that might be incorrectly predicted for a particular person. People are generally good at making between-person identity judgments over age gaps spanning from infancy to adulthood [15], and this is likely because human observers base identity decisions on features that change little over the lifespan. Age-progressed images may feature alterations to such cues, leading members of the public to miss correct identity matches because the range of matching faces has shifted to a those with incorrectly rendered features that the true individual actually does not possess or by increasing the number of feasible false alarms [18]. The precise mechanism, like the many possible predicted facial appearances, may vary among methods, artists, and individuals whose faces are being age-progressed. For this reason, we wish to emphasize that although perceptual and memory experiments may fail to find an aggregate benefit of age-progressed images, individual age-progressions within these studies and in real cases can sometimes be very accurate resemblances of their target individuals. Moreover, we would not recommend removing human artists from these cases in response to the data reported here. Rather, we hope these results may be incorporated with other findings to formulate a set of “best practices” for training artists on constructing age-progressed images, which we will discuss after briefly exploring recent advances in facial aging algorithms.

8.2 *Algorithm-Based Image Manipulation*

If research shows limited efficacy for artist-based age progressions, then how do algorithm-based aging systems fare? Automated face recognition systems have seen advances and widespread implementation over the last two decades from social media platforms to border security checkpoints. The specific ways in which these systems operate vary from early multivariate eigenface frameworks [26], Fisherface systems, [27], to convolutional neural networks (CNNs;[28, 29]). Often such face recognition systems are developed and tested using high-quality two-dimensional standardized image sets featuring multiple images of multiple identities, and recent CNNs such as DeepFace [29] rival human performance when generalizing to new faces outside of training databases. However, many of these systems face similar challenges to human observers when tasked with recognizing faces in important situations including aging gaps, which decrease automated accuracy sometimes nearer to chance levels when exemplar images are many decades out of date [30]. Importantly, most literature presenting algorithm performance only reports accurate matches as markers of success and not mismatches. This means they do not measure the full scope of how their algorithms learn about within- and between-person similarities. Nonetheless, a recent systematic comparison between various algorithms and human observers (including super-recognizers) has found that the best automated methods and the best observers perform equivalently in certain tests, and the most accurate recognition rates were found in conditions where when humans and automated systems work together [31]. However, no single algorithm has found widespread implementation, and the newest, state-of-the-art systems may not see adoption until years after development.[2]

---

[2] A reviewer of this manuscript suggested comparing algorithm performance to our human observer data in a brief follow-up experiment. Although such a comparison would be of great interest, it is not within the current scope of this article. Importantly, due to the open-ended instructions we gave to artists, our small set of stimulus images features a wide variety of uncontrolled, non-trivial variability in pose, brightness, and contrast among images of the

Regardless of specific methods used, automated systems designed to age-progress facial images rarely incorporate knowledge of biological relatives when making identity decisions across age gaps. Instead, they make adjustments to stored exemplar images when presented with novel input photographs – in effect, age-progressing the images they have stored in memory to compare to newly encountered faces [32]. They can do this by digitally altering facial images in ways that are common to all people even if specific timing and magnitude of these morphological changes varies among individuals and prototypic groups such as sex, race, and ethnicity [33]. Resultant images can also be exported and disseminated to the public, although they employ general, "one size fits all" approaches that typically alter facial length, mandible size, and skin texture while retaining key features that remain stable over a lifespan (e.g., iris diameter, inter-canthial distance). Koudelova et al.[33] conclude that accounting for age-related facial appearance changes is still "difficult to get right, subjective, and variable" (p. 2) for individual targets no matter the precision of the given techniques in aggregate. However, after systematically photographing children volunteers from ages 7 to 17 years, the authors were able to highlight facial regions that change the most for male and female Caucasian children in the Czech Republic aging from childhood to young adulthood. Importantly, they were also able to determine the ages at which the most dramatic changes manifest, observing forehead growth and facial lengthening at around 12 years old for girls and brow pronouncement and jaw deepening at 14 years old for boys. Although their findings are noteworthy, their methods – using computer

---

same identities. Such image-level discrepancies produce a point of divergence between human and computer vision accuracy, which would make interpretations from comparisons using our image set problematic. Additionally, the aim of this study was to compare two approaches by forensic artists (with or without a parental reference) and we have no reason to believe the relative differences found would vary using an algorithm approach. There is evidence that suggest human and automatic face recognition systems relative performance (based on difficulty) are correlated with one another and therefore we anticipate that automatic face recognition systems would have a similar pattern of results [34].

models to take careful observations of individuals over a lifespan to highlight areas of most typical change – are an important possible future direction for artist-based age-progressions.

Regardless of specific methods used, automated systems rarely incorporate knowledge of biological relatives when making identity decisions across age gaps. Instead, they make adjustments to stored exemplar images when presented with novel input photographs – in effect, age-progressing the images they have stored in memory to compare to newly encountered faces [32]. They can do this by digitally altering facial images in ways that are common to all people even if specific timing and magnitude of these morphological changes varies among individuals and prototypic groups such as sex, race, and ethnicity [33]. Resultant images can also be exported and disseminated to the general public, although they employ general, "one size fits all" approaches that typically alter facial length, mandible size, and skin texture while retaining key features that remain stable over a lifespan (e.g., iris diameter, inter-canthial distance). Koudelova et al.[33] conclude that accounting for age-related facial appearance changes is still "difficult to get right, subjective, and variable" (p. 2) for individual targets no matter the precision of the given techniques in aggregate. However, after systematically photographing children volunteers from ages 7 to 17 years, the authors were able to highlight facial regions that change the most for male and female Caucasian children in the Czech Republic aging from childhood to young adulthood. Importantly, they were also able to determine the ages at which the most dramatic changes manifest, observing forehead growth and facial lengthening at around 12 years old for girls and brow pronouncement and jaw deepening at 14 years old for boys. Although their findings are noteworthy, their methods – using computer models to take careful observations of individuals over a lifespan to highlight areas of most typical change – are an important possible future direction for artist-based age-progressions. Additionally, there has been recent work using algorithm-based aging models that uses a genetic bias to account for changes to the facial structure that may be biologically influenced [35]. Similar to forensic artists in our study, the genetic bias in the algorithm makes use of an age-matched biological relative. However, the algorithm-based approach uses systematic weighting based upon the similarity of the ‘missing’ person and the relative unlike the forensic artists in our study [35]. The growth and refinement of algorithms incorporating a genetic bias may be a step that aids in the recovery of missing persons.

8.3 *Recommendations and Future Directions*

The future best practices for forensic age progressions will increasingly require artists to work with computer assistance. For the human artist, a great deal of training and education goes into not only artistic technique and the processes of digital image manipulation but also learning fundamentals of human anatomy. The current experiments’ results illustrate a shortcoming that artists may experience when incorporating knowledge of the latter: Namely, trying to use parental images as a guide for better prediction ironically decreases the effectiveness of age-progressed images as tools for recognition, particularly over long age gaps. Perhaps quantification of timescale and magnitude of age-related appearance changes across any given age gap for any given subpopulation (e.g., sex, ethnicity) can give artists specially-tailored guides for what features to focus their attention on changing and – just as importantly – what features to leave unchanged when producing age-progressed images. This would mirror so-called “feature comparison” training found elsewhere in the forensic literature for trained facial comparison experts. Feature comparison encourages forensic face examiners to focus on specific features distinctive to wanted or missing persons, and also the features least likely to vary among different images of the same people. A recent systematic evaluation of facial examiner training programs has found that this type of training is the most effective for actually increasing facial matching abilities [36]. Rather than asking artists to take liberties with estimating future facial appearance incorporating some mixture of each biological parent, something like feature comparison training for forensic artists will help them focus on only the most age-stable features for individuals in comparison to the general population. This would require that artists train extensively by creating age-progressed images of individuals whose future appearances are known, although not to the artist. In real cases, artists rarely receive feedback for the images they create because many long-term missing persons remain missing. If artists are frequently challenged to create age-progressed images for training purposes and are blind to such tests until after the age-progressions are complete, they will be able to implement formative feedback by comparing their images to the true individuals at the target ages.

There are several limitations not only to the current study and the literature in general that prevent us from apprehending the full picture of age-progression creation and how human observers interact with them. One limitation of this is that we relied solely on unfamiliar age-progressions. As described above, many everyday challenges render unfamiliar face recognition error-prone, including recognizing faces across race or ethnicity, age group, and sex. An as-yet unexplored avenue in this literature would determine the magnitude to which such cross-categorical perceptions affect age-progression recognition. Moreover, no research has systematically examined the extent to which artists approach age-progression of individuals belonging to categories different than the artists themselves. Future research should also explore perception and recognition among those familiar with those whom age-progressions are intended to depict. Facial familiarity is marked by greater adaptability to within-person variability, so familiar observers may have better recognition rates than those who are unfamiliar by seeing past idiosyncratic liberties taken by artist. Additionally, researchers can increase familiarity with the targets during an initial learning phase at the start of the study. Finally, this study solely focused on whether parental references provided any benefit to increasing the validity of age-progressions. In real-life cases, forensic artists often use relatives besides the parents as a reference. Therefore, future studies should investigate whether having these additional references increases the age-progression's validity.

References

[1] Federal Bureau of Investigation, National Crime Information Center. (2018). NCIC Missing Person and Unidentified Person Statistics for 2018. Retrieved from https://www.fbi.gov/file-repository/2018-ncic-missing-person-and-unidentified-person-statistics.pdf/view

[2] A. Sedlak, D. Finkelhor, & H. Hammer, (2005). NISMART, National Estimates of Children Missing Involuntarily or for Benign Reasons.

[3] C. Wilkinson, A. Tillotson, Post-mortem prediction of facial appearance, in: C. Wilkinson, C. Rynn (Eds.), Craniofacial Identification, Cambridge University Press, Cambridge, 2012

[4] K.T. Taylor, Forensic Art and Illustration, CRC Press, 2001

[5] J.B. Pittenger, R.E. Shaw, & L.S. Mark, Perceptual information for the age level of faces as a higher order invariant of growth. Journal of Experimental Psychology: Human Perception and Performance, 5(1979), 478–493. https://doi.org/10.1037/0096-1523.5.3.478

[6] J. Mullins, Age progression and regression. Craniofacial identification, (2012), 68-75.

[7] T.J. Gillgrass, & R. Welbury, Craniofacial growth and development. Paediatric Dentistry, 1 (2012).

[8] P.H. Buschang, R. Tanguay, L. LaPalme, & A. Demirjian, Mandibular growth prediction: mean growth increments versus mathematical models. The European Journal of Orthodontics, 12(1990), 290-296.

[9] B. Solow, The dentoalveolar compensatory mechanism: background and clinical implications. British journal of orthodontics, 7(1980), 145-161.

[10] S.A. Feik, & J.E. Glover, Growth of children’s faces. Craniofacial identification in forensic medicine, (1998) 203-224.

[11] D.S. Carlson, (2005, December). Theories of craniofacial growth in the postgenomic era. In Seminars in Orthodontics(Vol. 11, No. 4, pp. 172-183). WB Saunders.

[12] L.M. DeBruine, F.G. Smith, B.C. Jones, S.C. Roberts, M. Petrie, & T.D. Spector, Kin recognition signals in adult faces. Vision research, 49(2009), 38-43.

[13] A. Alvergne, C. Faurie, & M. Raymond, Differential facial resemblance of young children to their parents: who do children look like more?. Evolution and Human behavior, 28(2007), 135-144.

[14] M.B. Lopez, A. Hadid, E. Boutellaa, J. Goncalves, V. Kostakos, & S. Hosio, Kinship verification from facial images and videos: human versus machine. Machine Vision and Applications, 29(2018), 873-890.

[15] D. Seamon, The phenomenological contribution to environmental psychology. Journal of environmental psychology, 2(1982), 119-140.

[16] M. Bruck, P. Cavanagh, & S.J. Ceci, Fortysomething: Recognizing faces at one’s 25th reunion. Memory & Cognition, 19(1991), 221-228.

[17] J.M. Lampinen, J.T. Miller, & H. Dehon, Depicting the missing: Prospective and retrospective person memory for age progressed images. Applied Cognitive Psychology, 26 (2012), 167-173.

[18] S.D. Charman, & R.N. Carol, Age-progressed images may harm recognition of missing children by increasing the number of plausible targets. Journal of Applied Research in Memory and Cognition, 1(2012), 171-178.

[19] D. Bates, M. Maechler, B. Bolker, S. Walker, R.H.B. Christensen, H. Singmann, ... & L. Rcpp, Package ‘lme4’. Convergence, 12(2015).

[20] W.B. Erickson, J.M. Lampinen, C.D. Frowd, & G. Mahoney, When age-progressed images are unreliable: The roles of external features and age range. Science & Justice, 57(2017), 136-143.

[21] J.M. Lampinen, W.B. Erickson, C.D. Frowd, C.D., G. Mahoney, Forensic Age Progression, in: S.J. Morewitz, C.S. Colls (Eds.), Handbook for Missing Persons, Springer, New York, 2016, pp. 251-269.

[22] R. Kemp, N. Towell, & G. Pike, When seeing should not be believing: Photographs, credit cards and fraud. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 11 (1997), 211-222.

[23] R. Jenkins, D. White, X. Van Montfort, & A.M. Burton, Variability in photos of the same face. Cogn. 121 (2011), 313-323.

[24] A.M. Megreya, A. Sandford, & A.M. Burton, Matching face images taken on the same day or months apart: The limitations of photo ID. Applied Cognitive Psychology, 27 (2013), 700-706.

[25] R.A. Johnston, & A.J. Edmonds, Familiar and unfamiliar face recognition: A review. Memory, 17(2009), 577-596.

[26] M. Turk, & A. Pentland, Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1991), 71-86.

[27] P.N. Belhumeur, J.P. Hespanha, & D.J. Kriegman, Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis & Machine Intelligence, (1997), 711-720.

[28] O.M. Parkhi, A. Vedaldi, & A. Zisserman, (2015, September). Deep face recognition. In bmvc (Vol. 1, No. 3, p. 6).

[29] Y. Taigman, M. Yang, M.A. Ranzato, & L. Wolf, (2014). DeepFace: Closing the gap to human-level performance in face verification. IEEE Conference on Computer Vision and Pattern Recognition (2008*)*. https://doi.org/10.1109/CVPR.2014.220

[30] K. Ricanek, & T. Tesafaye, (2006, April). Morph: A longitudinal image database of normal adult age-progression. In 7th International Conference on Automatic Face and Gesture Recognition (FGR06) (pp. 341-345). IEEE.

[31] P.J. Phillips, A.N. Yates, Y. Hu, C.A. Hahn, E. Noyes, K. Jackson, J.G. Cavazos, G. Jeckeln, R. Ranjan, S. Sankaranarayanan, J.C. Chen, C.D. Castillo, R. Chellappa, D. White, D., & A.J. O'Toole, Face recognition accuracy of forensic examiners, super-recognizers, and face recognition algorithms. Proceedings of the National Academy of Sciences of the United States of America, 24(2018), https://doi.org/10.1073/pnas.1721355115

[32] N. Ramanathan, R. Chellappa., & S. Biswas, Computational methods for modeling facial aging: A survey. Journal of Visual Languages & Computing, 20(2009), 131-144.

[33] J. Koudelová, E. Hoffmannová, J. Dupej, & J. Velemínská, Simulation of facial growth based on longitudinal data: Age progression and age regression between 7 and 17 years of age using 3D surface data. PloS one, 14(2019), e0212618.

[34] A.M. Burton, P. Miller, V. Bruce, P.J.B. Hancock, & Z. Henderson, (2001). Human and automatic face recognition: a comparison across image formats. Vision Research, *24* (2001), 3185-3195.

[35] S.J. Gibson, C.M. Scandrett, C.J. Solomon, M.I.S. Maylin, & C.M. Wilkinson, Computer assisted age progression. Forensic science, medicine, and pathology, 5 (2009), 174-181.

[36] A. Towler, R.I. Kemp, A.M. Burton, J.D. Dunn, T. Wayne, R. Moreton, & D. White, Do professional facial image comparison training courses work?. PloS one, 14(2019), e0211037.

Table 1. Regression Table of the Full Model

Mixed Linear Model predicting similarity (N=102)

Fixed Factors	B	SE	t
Intercept	4.01	0.24	16.57***
Age Range (Reference: 12-20)	-0.02	0.33	.067
Parent Reference (Reference: Yes)	-0.41	0.33	-1.25
Foil vs Target (Reference: Target)	0.36	0.08	4.62***
Age Range * Parent Reference	0.15	0.47	0.32
Age Range * Foil vs Target	0.31	0.11	-2.79**
Parent Reference * Foil vs Target	0.13	0.11	1.48
Age Range * Parent Reference * Foil vs Target	-0.47	0.16	-3.02**

Note: ***p < .001, ** p <.01, * p <.05, + p <.10

Figure Captions

Figure 1. Graph display the mean similarity ratings with standard error bars of target/foil by Parental reference interaction (left) and target/foil by age-range reference interaction (right)

Figure 2. Graph display the mean similarity ratings with standard error bars for three-way interaction between Age-Range, Parental Reference, and Target/Foil

Figure 3. Graph display the yes/no recognition performance with standard error bars for all seven conditions

Figure 1.

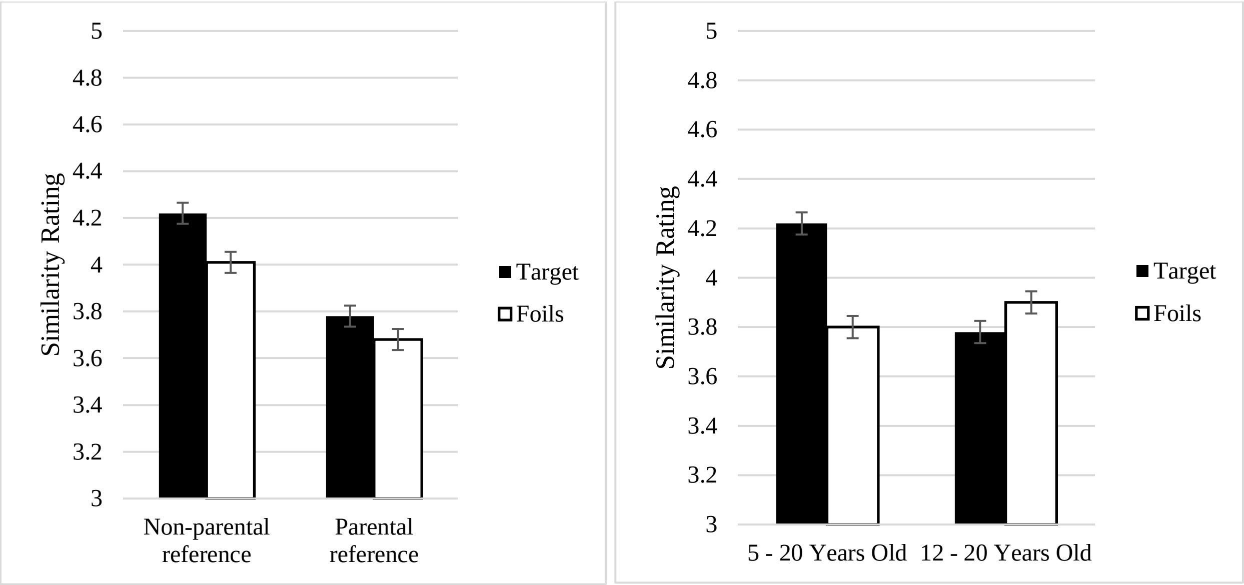

5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Similarity Rating
Non-parental reference
Parental reference
■ Target
□ Foils
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Similarity Rating
5 - 20 Years Old
12 - 20 Years Old
■ Target
□ Foils

Figure 2.

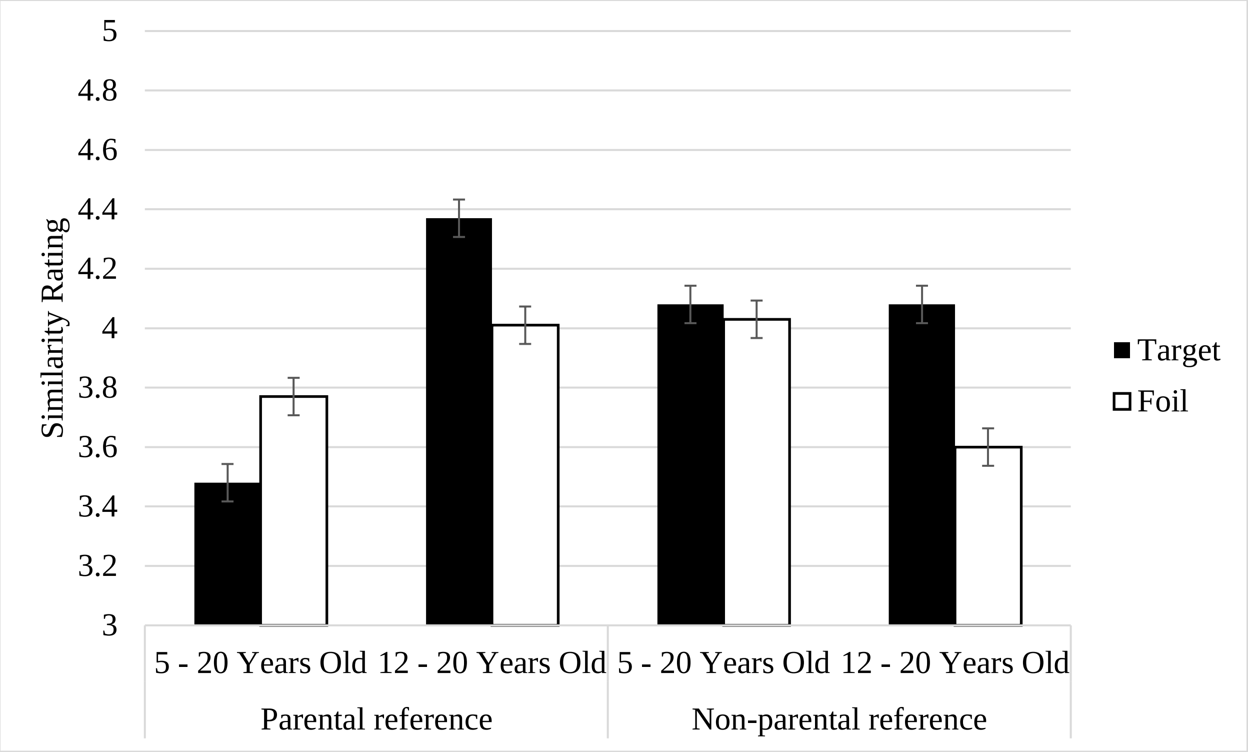
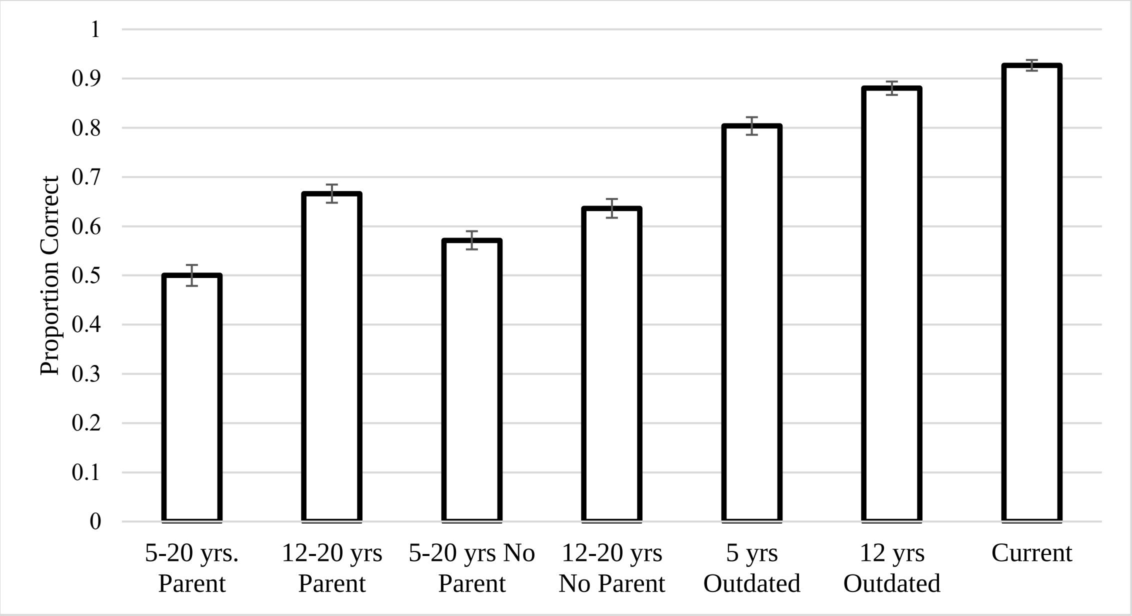

5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Similarity Rating
■ Target
☐ Foil
5 - 20 Years Old
12 - 20 Years Old
5 - 20 Years Old
12 - 20 Years Old
Parental reference
Non-parental reference

Figure 3.


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Proportion Correct
5-20 yrs. Parent
12-20 yrs Parent
5-20 yrs No Parent
12-20 yrs No Parent
5 yrs Outdated
12 yrs Outdated
Current