14. RECOGNITION OF AGE-DIFFERENT FACIAL IMAGES OF UNFAMILIAR CHILDREN: WHAT ARE THE INFLUENTIAL FACTORS?

Zuzana Obertová[1*], Vera Lammel[1,2], Melanie Ratnayake[2], Daniele M. Gibelli[1], Debora Mazzarelli[1], Cristina Cattaneo[1]

[1]*LABANOF, Department of Biomedical and Health Sciences, University of Milan, Milan, Italy*
[2]*Institute of Zoology, University of Hamburg, Hamburg, Germany*
**obertovazuzana@yahoo.co.nz*

https://dx.doi.org/10.15496/publikation-66852

KEYWORDS | face recognition; age-different images; children; facial features; missing persons

ABSTRACT

Background: Each year 8 million children go missing worldwide. As time passes, physical changes associated with growth and development may be detrimental for a potential recognition by third parties. The objective of this study was to identify observer- and image-related factors that influence recognition of age-different unfamiliar faces of children.

Methods: Sixty frontal photographs of European males and females aged between 10 and 21 years were selected from a large data set of longitudinal images, along with one matching and three non-matching images and randomly ordered in a face recognition test catalogue. The image comparisons were equally distributed into three age gaps: 1-2 years, 3-5 years, and 6+ years. Twenty one observers (10 experienced and 11 non-experienced with facial identification) were asked to match the images, and to specify which facial features prompted their decision.

Results: Experts were more likely to correctly match child faces on age-different images compared with non-experts (82% v. 72%). The frequency of correct recognition decreased with an increasing age gap between the images. Conclusion: Experience with facial identification facilitated correct matches of unfamiliar faces of children on age-different images, partly due to the difference in the type of features noted as useful for recognition. Age gaps of six and more years between images rendered recognition more difficult regardless of the expert status. By identifying the factors that facilitate facial recognition on age-different images, the results of this study provide useful information for management of cases of missing children.

14.1 INTRODUCTION

Each year 8 million children go missing worldwide (National Center for Missing & Exploited Chil-

Michael Francken, Katerina Harvati (Hrsg.), Facetten der Osteologie. Festschrift zu Ehren von Prof. Dr. Joachim Wahl. Tübingen University Press, Tübingen 2023. https://dx.doi.org/10.15496/publikation-66865

dren, 2017). The majority of cases, mainly those involving teenage runaways, will be solved within hours or days. However, many children remain missing for more than a year. Missing Children Europe (2014) reported that only one third of the missing children reported to their hotlines were found in the same year. Although cases of longterm missing children are rarely recounted in the media after the initial missing child report is made public, the solving of the disappearances remains urgent even after years both for the families and the law enforcement (Findlay and Lowery, 2011). In some cases, missing children were recovered even after decades thanks to vigilance of members of the public or the police, who were able to recognize the children (Goldman, 2009).

The search for missing children usually includes posters with facial photographs being distributed nationally or even internationally by the media, social networks or as print-outs stuck to bulletin boards in supermarkets or pillars in transport stations (Findlay and Lowery, 2011). In all these cases, the 2D image from the poster is the initial stimulus for recognition.

To facilitate recognition of long-term missing children, organizations, such as National Center for Missing and Exploited Children (NCMEC) in the USA publish age-progressed images in certain intervals (National Center for Missing & Exploited Children, 2017). Long-term missing children are often very young when they disappear, so their face may undergo substantial changes due to growth and development (Findlay and Lowery, 2011). Since these changes are oftentimes difficult to predict, some authors argue that age-progressed images may not provide the right stimulus for recognition (Charman and Carol, 2012; Lampinen et al., 2012a, b).

Facial recognition or identification is commonly used as a tool for crime solving, since everyday activities (legal or illegal) are increasingly captured by surveillance systems (Davis et al., 2012; Davis and Valentine, 2015). The majority of the forensic methods for facial comparisons, either biometric or manual are developed for adults (Davis et al., 2010; Ritz-Timme et al., 2011; Arca et al., 2012; Davis and Valentine, 2015).

The challenges of facial recognition of unfamiliar faces (as opposed to our innate, but also not flawless ability to recognize familiar faces) have been well described in the literature (Ellis et al., 1979; Hancock et al., 2000; Johnston and Edmonds, 2009). Temporal differences between facial images of adults decreased the accuracy of correct matches by up to 20% (Megreya et al., 2013). Although adult faces change with age, these changes are considered to be less dramatic compared with those occurring in a growing child.

The objective of this study was to test the ability of experts and non-experts to recognize the face of an unfamiliar child on age-different images, and to examine which facial features are beneficial for recognition.

14.2 MATERIAL AND METHODS

The study was undertaken at the Institute of Forensic Medicine, The University of Milan (Italy). Twenty one observers – 10 males, and 11 females; mean age of 30 years (22-43 years) – completed the face recognition test. The observers were comprised of 11 non-experts (students of natural sciences, archaeology, biology and medical registrars), and 10 forensic experts (forensic pathologists and anthropologists) who regularly work on cases concerning personal identification based on various methods, including assessment of facial features. There were equal numbers of males and females by expert status.

The face recognition test included overall 300 frontal facial images of European males and females aged between 10 and 21 years collected in Italy and Germany. The photographs for the test catalogue were acquired during the implementation of the EU-funded project JLS/2007/ISEC/451. The project explored the possibility of extracting age-related information from the changes in facial proportions of children and juveniles. For this purpose, a large data set including sets of five images (as frontal as possible) of one person taken at different ages were acquired from personal albums of participants. The study was approved by the Ethics Committees of the participating countries. The subjects and their parents (in case of minors) signed an informed consent prior to their participation and image collection.

The face recognition test consisted of randomly ordered 60 cases (30 males, 30 females) of one target image and four comparative facial images, of which one was a match for the target individual. The images were equally distributed to represent three age gaps between the target and the comparative individuals: 1-2 years, 3-5 years, and 6+ years. The individuals on the target image were between 10 and 12 years of age. The largest age gap was 11 years. The comparative images were purposefully selected to show individuals who look as alike as possible (including eye and hair color) to the target and their matching individual and were at a similar age range of ±2 years to the individual on the matching image. The images were not modified in order to resemble real-life scenarios of recognizing missing persons from casual images provided by the family. Therefore the images differed in lighting, focus, background, and head orientation (although only slightly because of the prerequisite that the collected images shall depict the face in as frontal position as possible).

The observers were asked to choose whether the target facial image corresponds to one of the four comparative images in a hardcopy version of the face recognition test. None of the observers knew the depicted individuals, so the test assessed the recognition of unfamiliar faces. For each matching pair, the observers were asked to identify one or more facial features useful for recognition. The features were entered as free text; the observers were not offered a list of facial features to choose from. There were no time restrictions for completing the task.

After all observers completed the test, the facial features indicated as useful for recognition were grouped into nine categories: eyes (including eyelids), nose, mouth, ears, teeth, mandible/chin, facial shape, eyebrows, and other (including hair/hairline, cheeks, forehead, moles, freckles, smile/expression, and overall).

The statistical analysis was performed using IBM SPSS Statistics 22 software. The frequency of correct matches was calculated and compared by chi-squared test by expert status and sex of the observer, by sex of the individuals depicted on images and age gaps of the age-different images. The images matched correctly by less than half of the observers were assessed in more detail considering the facial features used for correct and incorrect recognition, sex of the depicted individual, age gap, and facial expression. The distribution of facial features useful for recognition was calculated from the total number of features mentioned and was assessed by expert status using the chi-squared test. Test probability level of <0.05 was considered to be statistically significant.

14.3 RESULTS

The twenty one observers achieved a correct match in 76.7% of the 60 images. Female observers performed better than male observers (79.1% v. 74.2%, chi^2 test $p = 0.04$). Experts performed better than non-experts (82.2% v. 71.8%, chi^2 test $p < 0.0001$).

The correct-match frequency was greater for female images (79.0%) compared with male ones (74.4%), but the difference was not statistically significant. The lowest frequency of correct recognition was found for the 6+ years age gap between the age-different images (68.8%) compared with 79.5% for 1-2 years age gap and 81.9% for 3-5 years age gap. There was a significant difference between the age gaps 1-2 and 6+ years age gap (chi^2 test $p = 0.0004$) and 3-5 and 6+ years age gap (chi^2 test $p < 0.0001$), but not between 1-2 and 3-5 years age gap (Table 1).

There was no statistically significant difference in the proportion of correct matches between experts and non-experts when the age gap of the age-different images was 1-2 years, while with an increasing

	Overall	Age gaps	Experts	Non-experts	Experts v. non-experts
	% (n/N)	P (chi^2 test)	% (n/N)	% (n/N)	P (chi^2 test)
1-2 yrs	79.5 (334/420)	ns (v. 3-5yrs)	82.5 (165/200)	76.8 (169/220)	ns
3-5 yrs	81.9 (344/420)	<0.0001 (v. 6+ yrs)	87.0 (174/200)	77.3 (170/220)	0.01
6+ yrs	68.8 (289/420)	0.0004 (v. 0-2 yrs)	77.0 (154/200)	61.4 (135/220)	0.0006

Table 1: The frequency of correct matches by age gap and expert status. n: number of correct matches, N: number of observations. ns: not significant

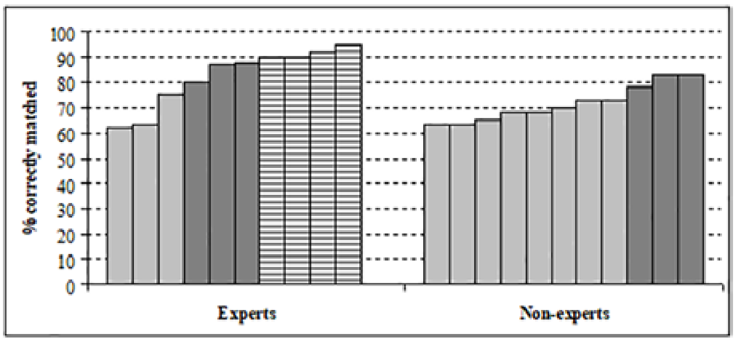

Figure 1: The proportion of correctly matched images for each expert and non-expert compared with total average of correct recognition (77%). Light grey: below total average, striped: above total average, dark grey: about average

age gap the performance of non-experts was progressively worse in comparison to experts (Table 1).

The proportion of correctly matched images for each expert and non-expert is shown in Figure 1 to highlight the performance of observers by expert status in relation to the total average frequency of correct matches (76.7%). Four out of 10 experts performed above average, while eight out of 11 non-experts performed below average.

The observers noted in total 2296 features useful for recognition. In 13 instances the space for specification of the features was left blank; of these 12 were by non-experts. Experts noted slightly more features (1240) than non-experts (1096). Significantly more features were reported when a correct match was achieved (1843) compared with 453 features noted for incorrect matches. Table 2 summarizes the distribution of the nine categories of features (eyes, nose, mouth, ears, teeth, mandible/chin, facial shape, eyebrows, other) noted as useful for recognition by expert status and correct/incorrect classification. The eyes were the most commonly noted feature useful for recognition in 44.4% (of the total number of cases), followed by the nose (39.1%) and the mouth (33.5%). Non-experts noted the nose most commonly (41.7%), followed by the eyes (35.9%), and the mouth (28.8%), while the most common features reported by experts were the eyes (53.8%) followed by the mouth (38.7%), and the nose (36.3%). Experts were significantly more likely to report eyes, mouth, ears, mandible/chin, eyebrows, and other features (including hair/hairline, moles/freckles, smile/expression) as features useful for recognition, while there was a tendency for non-experts to more likely note the nose as a useful feature.

The most commonly noted features for correct matches were the eyes (48.0%), followed by the nose (38.6%) and the mouth (36.3%), while for incorrect matches the most commonly reported feature was the nose (41.0%), followed by the eyes (32.8%) and the mouth (24.2%). The eyes, mouth,

	% (n/1260)	Experts (n/600)	Non-experts (n/660)	p (chi² test)	Correct matches (n/967)	Incorrect matches (n/293)	p (chi² test)
Eyes	44.4 (560)	53.8 (323)	35.9 (237)	<0.0001	48.0 (464)	32.8 (96)	<0.0001
Nose	39.1 (493)	36.3 (218)	41.7 (275)	ns	38.6 (373)	41.0 (120)	ns
Mouth	33.5 (422)	38.7 (232)	28.8 (190)	0.0002	36.3 (351)	24.2 (71)	0.0001
Ears*	14.0 (141)	16.9 (81)	11.4 (60)	0.01	14.3 (107)	13.1 (34)	ns
Teeth**	13.9 (76)	15.4 (40)	12.6 (36)	ns	14.2 (62)	12.8 (14)	ns
Mandible/chin	15.2 (191)	18.2 (109)	12.4 (82)	0.005	15.5 (150)	14.0 (41)	ns
Facial shape	10.4 (131)	10.3 (62)	10.5 (69)	ns	9.9 (96)	11.9 (35)	ns
Eyebrows	5.5 (69)	9.2 (55)	2.1 (14)	<0.0001	6.2 (60)	3.1 (9)	0.04
Other (incl. hair/ hairline, moles/ freckles, smile/ expression)	16.9 (213)	20.0 (120)	14.1 (93)	0.005	18.6 (180)	11.3 (33)	0.003

Table 2: Categories of features noted as useful for recognition overall, by expert status and correct/incorrect classification. ns: not significant

eyebrows and other features (including hair/hairline, moles/freckles, smile/expression) were significantly more likely noted when correct matches were made compared with incorrect matches.

Within the "other" category, the most commonly noted feature was hair/hairline (34.7%), followed by smile/expression (18.3%), and moles/freckles (16.4%).

Moles/freckles were noted in 7% of possible cases, but when they were reported as useful for recognition, the given individuals were matched correctly in 100% of cases.

For nine (out of 60) images less than half of the observers found the correct match (Table 3). The theoretical percentage of achieving a correct match by chance would be 12.5% for each image. There was one image, to which only one observer (an expert) assigned the matching image correctly, so the overall correct-match frequency was only 5%. For these "difficult" images, 77.4% of the correct matches were achieved by experts.

On six of the nine images, the sex of the depicted individual was male. Five of the images had an age gap of 6+ years, while two had an age gap of 1-2 years. On seven images the most commonly noted feature were the eyes when a correct match was made, while on six images the most commonly used feature in case of an incorrect match was the nose. For correct matches, ears were noted on three images as useful for recognition. When incorrect matches were made, teeth were noted as useful for recognition in two cases, in which teeth were visible on only one of the comparative images (obviously a non-match) and on the target image.

14.4 DISCUSSION AND CONCLUSION

This study aimed at identifying observer- and image-related factors that affect the ability to recognize unfamiliar children at age-different images in order to contribute to the forensic knowledge

Image code	% correct match	Sex of the person on image	Age gap (years)	The most commonly noted feature(s) for correct match (in descending order)	The most commonly noted feature(s) for incorrect match (in descending order)	Comment±
D1	5	Male	3-5	ears, facial shape	nose	
D2	14	Female	6	nose, mouth	eyes, nose, teeth	Only the person on the target image and on one of the comparative images (non-match) had visible teeth
D3	24	Male	6	eyes	mouth, nose, teeth	Only the person on the target image and on one of the comparative images (non-match) had visible teeth
D4	29	Male	6	eyes, mouth, facial shape	nose, ears, mouth, eyes	
D5	29	Female	1-2	eyes, mouth	eyes, mouth, nose	Due to lighting the eye colour seemed more similar between the target and a non-match (indicated in all incorrect answers)
D6	38	Female	6	eyes, mouth	nose, eyes	
D7	43	Male	6	eyes, ears, nose	nose	
D8	48	Male	1-2	eyes, ears	nose, mouth	
D9	48	Male	3-5	ears, nose	nose, eyes, mouth	

Table 3: Overview of the characteristics of the "difficult" images, which less than half of the observers matched correctly.

in relation to the management of cases of missing children.

The total percentage of correctly matched age-different images was 77%. The overall percentage of correct matches was in accordance with earlier studies on recognition of unfamiliar faces of adults (Bruce et al., 1999; Megreya and Burton 2007; Burton et al., 2010; Megreya and Bindemann, 2015), although it was higher than the percentage of correct matches for age-different (by not more than 1 year) images of adults (Megreya et al., 2013). Sex, age, and ancestry of both the persons depicted on images and the observers were reported to affect the accuracy of face recognition (Light et al., 1979; Fulton and Bartlett, 1999; Hofmann et al., 2006; McBain et al., 2009; Megreya et al., 2011; Megreya and Bindemann, 2015). Females were shown to outperform males in facial recognition, which was also the case in the present study (Light et al., 1979; McBain et al., 2009).

In theory, the ability of recognizing human faces follows the Gaussian distribution in the general population (Burton et al., 2010), but experts in the present study showed above-average recognition ability while eight out of eleven of non-experts performed below average. Some studies showed that while there were so-called super-recognizers with a natural ability to quickly and accurately recognize unfamiliar faces (Russell et al., 2009), experience did not necessarily improve the recognition ability, especially when facial images were concerned (White et al., 2014c). In contrast, other studies emphasized that experience and feedback training were associated with increased accuracy in facial recognition in facial image analysts and students, respectively (Wilkinson and Evans, 2009; White et al., 2014b).

No difference was found between experts and non-experts in recognizing children on images with an age gap of 1-2 years but the larger the age gap the larger also the gap between the performance of experts and non-experts. Experience and training of the experts seemed to help them to look beyond the often dramatic growth and developmental changes of the face and use features that were stable through time to achieve a correct match.

Not surprisingly, the larger the age gap between images, the less likely were the children recognized. Notably, the recognition accuracy was not affected when comparing images with less than 6 years age gap, while age gaps of six and more years proved detrimental for recognition. In this study, the initial images represented children aged 10-12 years so images with an age gap of 6 or more years showed the children after the completion of pubertal changes. This major biological change may have contributed to the drop in recognition accuracy.

Previous studies showed that eyes, mouth, hairline, and ears were important for recognition of unfamiliar faces (Luria and Strauss, 1978; Bruce et al., 1999; Liu et al., 2013). Especially the eyes are the initial center of focus and recognition is hindered when eyes are not visible on an image (Janik et al., 1978; Burton and Bindemann, 2009). In addition, Light et al. (1979) reported that the presence of an unusual facial feature improved the recognition accuracy of unfamiliar faces. In the present study, the most commonly noted feature for correct matches were the eyes, while for incorrect matches the nose was the feature most reported useful for recognition. Apart from the eyes and the mouth, features categorized as "other", including hair/hairline, and moles/freckles were more likely noted when correct matches were made.

Experts reported using eyes, ears, and the other features more often than non-experts did, and, conversely, non-experts noted the nose more commonly. This finding is in accordance with Čaplová et al. (2017), who studied facial recognition of deceased and living individuals. The authors noted that individualizing features, such as scars, moles and dental morphology were rarely considered, although expert observers (forensic practitioners who routinely dealt with facial image comparisons and human identification) took such facial features more commonly into account than non-experts (students) did.

A study by Liu et al. (2013) showed that using solely the nose for facial recognition resulted in a considerably lower matching accuracy than using the eyes and the mouth. This is in accordance with the results of the present study. In addition, the high number of incorrect matches when using the nose as the feature most useful for recognition may be partly explained by the frontal orientation of the faces on the test images. It is likely that for profile images the proportion of incorrect matches using the nose as main recognition feature would decrease. However, in general the accuracy of recognition was shown to decrease when only profile images were available for comparisons (Burton and Bindemann, 2009). Notably, White et al. (2014a) showed that using multiple images from different perspectives facilitated correct recognition both of familiar and unfamiliar faces, while the ability of recognizing unfamiliar faces was negatively affected when matching different views of the same face (Hancock et al., 2000; Burton and Bindemann, 2009). Similarly, Sweeney and Lampinen (2012) reported that the presentation of multiple images of a missing person was beneficial for recognition, but they also pointed out that the number of false positives increased.

In conclusion, systematic observation of specific facial traits, expert knowledge about the stability of facial features during growth and development, and experience in facial recognition/identification of persons on images were beneficial for the recognition accuracy in cases of unfamiliar children on age-different images. Guidelines specific for facial image comparison in children in the style of the existing literature concerning adults (Buhmann et al., 1999; European Network of Forensic Science Institutes, 2018) would facilitate the forensic assessment, while practical evidence-based suggestions, such as using multiple images in different facial perspectives or highlighting unusual facial features may prove beneficial in designing posters used in public searches for missing children.

REFERENCES

Arca, S., Campadelli, P., Lanzarotti, R., Lipori, G., Cervelli, F., Mattei, A., 2012. Improving automatic face recognition with user interaction. Journal of Forensic Sciences 57, 765–771.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P., Burton, A. M., Miller, P., 1999. Verification of face identities from images captured on video. Journal of Experimental Psychology: Applied 5, 339–360.

Buhmann, D., Helmer, R.P., Jaeger, U., Jürgens, H.W., Knussmann, R., Rösing, F.W., Schmidt, H.D., Szilvassy, J., Ziegelmayer, G., 1999. Standards für die anthropologische Identifikation lebender Personen auf Grund von Bilddokumenten, Grundlagen, Kriterien und Verfahrensregeln für Gutachten. Anthropologischer Anzeiger 57, 185–191.

Burton, A.M., Bindemann, M., 2009. The role of view in human face detection. Vision Research 49, 2026–2036.

Burton, A.M., White, D., McNeill, A., 2010. The Glasgow Face Matching Test. Behavior Research Methods 42, 286–291.

Čaplová, Z., Obertová, Z., Gibelli, D.M., Mazzarelli, D., Fracasso, T., Vanezis, P., Sforza, C., Cattaneo, C., 2017. The reliability of facial recognition of deceased persons on photographs. Journal of Forensic Sciences 62, 1286–1291.

Charman, S.D., Carol, R. N., 2012. 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, 171–178.

Davis, J.P., Valentine, T., Davis, R.E., 2010. Computer assisted photo-anthropometric analyses of full-face and profile facial images. Forensic Science International 200, 165–176.

Davis, J.P., Valentine, T., Wilkinson, C., 2012. Facial image comparison, in: Wilkinson, C., Rynn, C. (Eds.), Craniofacial Identification. Cambridge University Press, Cambridge, pp. 136–153.

Davis, J.P., Valentine, T., 2015. Human verification of identity from photographic images, in: Valentine, T., Davis, J. P. (Eds.), Forensic Facial Identification. John Wiley & Sons Ltd., Chichester, pp. 211–238.

Ellis, H.D., Shepherd, J.W., Davies, G.M., 1979. Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. Perception 8, 431–439.

European Network of Forensic Science Institutes (ENFSI), 2018. Best practice manual for facial image comparison (ENFSI-BPM-DI-01). http://enfsi.eu/wp-content/uploads/2017/06/ENFSI-BPM-DI-01.pdf (Accessed 18 November 2018).

Findlay, P., Lowery, R.G., Jr. (Eds.), 2011. Missing and abducted children: A law-enforcement guide to case investi-gation and program management. 4th ed. National Center for Missing & Exploited Children, Alexandria.

Fulton, A., Bartlett, J.C., 1991. Young and old faces in young and old heads the factor of age in face recognition. Psychology and Aging 6, 623–630.

Goldman, R., 2009. Jaycee Dugard looks like the image forensic artists created to help find her. https://abcnews.go.com/Technology/AheadoftheCurve/hundreds-missing-children-found-age-progression-images/story?id=8830185 (Accessed 10 August 2018).

Hancock, P.J., Bruce, V. V., Burton, A. M., 2000. Recognition of unfamiliar faces. Trends in Cognitive Sciences 4, 330–337.

Hofmann, S.G., Suvak, M., Litz, B. T., 2006. Sex differences in face recognition and influence of facial affect. Personality and Individual Differences 40, 1683–1690.

Janik, S.W., Wellens, A.R., Goldberg, M.L., Dell'Osso, L.F., 1978. Eyes as the center of focus in the visual examination of human faces. Perceptual and Motor Skills 47, 857–858.

Johnston, R. A., Edmonds, A. J., 2009. Familiar and unfamiliar face recognition: a review. Memory 17, 577–596.

Lampinen, J., Arnal, J.D., Adams, J., Courtney, K., Hicks, J.L., 2012a. Forensic age progression and the search for missing children. Psychology, Crime & Law 18, 405–415.

Lampinen, J.M., Miller, J.T., Dehon, H., 2012b. Depicting the missing: Prospective and retrospective person memory for age progressed images. Applied Cognitive Psychology 26, 167–173.

Light, L., Kayra-Stuart, F., Hollander, S., 1979. Recognition memory for typical and unusual faces. Journal of Experi-mental Psychology Human Learning and Memory 5, 212–228.

Liu, S., Anzures, G., Ge, L., Quinn, P.C., Pascalis, O., Slater, A.M., Tanaka, J.W., Lee, K., 2013. Development of recognition of face parts from unfamiliar faces. Infant and Child Development 22, 165–179.

Luria, S.M., Strauss, M. S., 1978. Comparison of eye movements over faces in photographic positives and negatives. Perception 7, 349–358.

McBain, R., Norton, D., Chen, Y., 2009. Females excel at basic face perception. Acta Physiologica (Oxford, England) 130, 168–173.

Megreya, A.M., Bindemann, M., 2015. Developmental improvement and age-related decline in unfamiliar face matching. Perception 44, 5–22.

Megreya, A.M., Burton, A.M., 2007. Hits and false positives in face matching: A familiarity-based dissociation. Per-ception & Psychophysics 69, 1175–1184.

Megreya, A.M., Sandford, A., Burton, A.M., 2013. Matching face images taken on the same day or months apart: the limitations of photo ID. Applied Cognitive Psychology 27, 700–706.

Megreya, A. M., White, D., Burton, A. M., 2011. The other-race effect does not rely on memory: evidence from a matching task. The Quarterly Journal of Experimental Psychology A 64, 1473–1483.

Missing Children Europe, 2014. Annual Report 2014. Missing Children Europe, Brussels.

National Center for Missing & Exploited Children (NCMEC), 2017. http://www.missingkids.com/1in6 (Accessed 21 July 2017).

Ritz-Timme, S., Gabriel, P., Obertová, Z., Boguslawski, M., Mayer, F., Drabik, A., Poppa, P., De Angelis, D., Ciaffi, R., Zanotti, B., Gibelli, D., Cattaneo, C., 2011. A new atlas for the evaluation of facial features: advantages, limits, and applicability. International Journal of Legal Medicine 125, 301–306.

Russell, R., Duchaine, B., Nakayama, K., 2009. Super-recognizers: people with extraordinary face recognition ability. Psychonomic Bulletin & Review 16, 252–257.

Sweeney, L.N., Lampinen, J.M., 2012. The effect of presenting multiple images on prospective and retrospective person memory for missing children. Journal of Applied Research in Memory and Cognition 1, 235–241.

White, D., Burton, A.M., Jenkins, R., Kemp, R.I., 2014a. Redesigning photo-ID to improve unfamiliar face matching performance. Journal of Experimental Psychology: Applied 20, 166–173.

White, D., Kemp, R. I., Jenkins, R., Burton, A.M., 2014b. Feedback training for facial image comparison. Psychonomic Bulletin & Review 21, 100–106.

White, D., Kemp, R. I., Jenkins, R., Matheson, M., Burton, A. M., 2014c. Passport officers' errors in face matching. PLoS One 9, e103510.

Wilkinson, C., Evans, R., 2009. Are facial image analysis experts any better than the general public at identifying indi-viduals from CCTV images? Science & Justice 49, 191–196.