Australian Government

Department of Defence
Science and Technology

Forensic (Facial) Comparison: Linking the Science and the Practice— Course Overview and Evaluation

Jessica O'Rielly and Rebecca Heyer

**National Security and ISR Division
Defence Science and Technology Group**

DST-Group-GD-1002

ABSTRACT

The Defence Science and Technology Group was commissioned by the Australian Federal Police to develop and deliver a course to familiarise participants with the psychological science underpinning forensic comparison, with specific focus on facial comparison, while fulfilling the requirements of six modules from the National Facial Training Framework. This paper contains an overview of the course content, and its delivery and evaluation. It concludes with some recommendations for the future, should the course be delivered again.

Produced by

National Security and ISR Division
Defence Science and Technology Group
PO Box 1500
Edinburgh SA 5111

Telephone: 1300 333 362

Copyright Commonwealth of Australia 2018
April 2018
AR-017-159

APPROVED FOR PUBLIC RELEASE

Forensic (Facial) Comparison: Linking the Science and the Practice— Course Overview and Evaluation

Executive Summary

The Australian Federal Police (AFP), Facial Identification - Specialist Operations Branch, commissioned the Defence Science and Technology (DST) Group to develop and deliver a training course for staff working in the forensic facial comparison domain. 'Forensic (Facial) Comparison: Linking the Science and the Practice' aimed to familiarise participants with the psychological science underpinning forensic comparison, with specific focus on facial image comparison, while fulfilling the requirements of six modules from the National Facial Training Framework.

The course was delivered across two days at the AFP's Forensic Facility Majura, Canberra. The first day of the course focused on the psychological theories relevant to forensic comparison, and their importance, as well as an overview of the facial comparison research. The second day focused on the sources of bias within the forensic comparison domain, and how to recognise and mitigate them. Attendees included nineteen individuals from the AFP from seven different disciplines (biological criminalistics/DNA, chemical criminalistics/ballistics, document examination, facial identification, fingerprint examination, intelligence, and quality assurance and training). Five individuals from other Australian policing and national security agency facial identification units also attended. The course was delivered in a traditional classroom-style format with content presented via PowerPoint™ slides interlaced with practical activities to consolidate learning. Opportunities for participant discussion and interaction were frequent.

Participants completed an evaluation at the conclusion of the course. Overall participants found the course enjoyable and engaging, generally rating the course to be of high educational value. While it was acknowledged that the course was commissioned by the Facial Identification Team, and as such was necessarily biased towards that discipline, it was still of value for the vast range of other disciplines represented. Participants appreciated the delivery style and the flexible timetabling of the sessions. That being said, some participants also indicated that the delivery of Section 1 of the course was not sufficiently in depth and did not provide clear linkages to their work in forensic comparison. Based on participant feedback two recommendations were made:

Recommendation 1: If delivering the course to a diverse audience in the future, content should be supplemented with research and case studies from other forensic comparison disciplines.

Recommendation 2: Section 1 of the course *An introduction to the psychology and human factors of forensic comparison* be reviewed to include more real world examples relevant to the field of forensic comparison.

This page is intentionally blank.

Contents

This page is intentionally blank.

1.   Introduction

The Australian Federal Police (AFP), Facial Identification - Specialist Operations Branch, commissioned the Defence Science and Technology (DST) Group to develop and deliver a training course for staff working in the forensic facial comparison domain. The course was developed and delivered under Project Arrangement S-010/2016, on a full cost recovery basis as required by DST's National Security Science and Technology business model. The course was designed to target specific training modules outlined in the Australian Government's National Facial Training Framework for facial comparison specialists, but was general enough to be applicable across a range of forensic comparison disciplines.

This paper provides an overview of the course content, its delivery and evaluation. It concludes with some recommendations for the future, should the course be delivered again.

2.   Course Content

The aim of the course, 'Forensic (Facial) Comparison: Linking the Science and the Practice' was to familiarise participants with the psychological science underpinning forensic comparison, with specific focus on facial image comparison, while fulfilling the requirements of six modules from the National Facial Training Framework. The course was delivered across two days (see Appendix A for course schedule). The first day focused on the psychological theories relevant to forensic comparison, as well as introduction to facial comparison research. The second day focused on the sources of bias within the forensic comparison domain, how to recognise and mitigate them.

The course was delivered in a traditional classroom-style format with content presented via PowerPoint™ slides interlaced with practical activities to consolidate learning. Opportunities for participant discussion and interaction were frequent. There were no formal pre-requisites for the course, however participants were provided with two articles for pre-reading (Dror, 2015; White, Kemp, Jenkins, Matheson & Burton, 2014). In addition, participants were provided with an extensive reference list of course material (see Appendix B) to consolidate learning across the four sections of the course.

2.1   Section 1: An introduction to the psychology and human factors of forensic comparison

Section 1 of the course fulfilled the requirements of Module 61 of the National Facial Training Framework to "identify the relevant psychological and other human factors research and their relevance to forensic comparison work". It included 86 PowerPoint™ slides and a practical exercise on cognitive ability and its measurement.

The lecture content provided a broad introduction to psychology and cognitive science. An overview of the biological mechanisms underpinning how we receive information, including brain architecture, visual processes and perception was provided. This was extended to demonstrate how this information is processed, stored by the memory system and then used by an individual for learning and decision making. An introduction to human factors was also provided with specific examples relevant to on-the-job performance. Interactive discussions were included throughout the session, which concluded with the practical exercise on cognitive ability – administration of the 80-item *Identical Pictures Test* (Ekstrom, French & Harman, 1979) and discussion of results.

2.2   Section 2: Facial image comparison: what does the research say?

Section 2 of the course fulfilled the requirements of Module 62 of the National Facial Training Framework to "describe the concepts and issues of familiar and unfamiliar observers for human facial comparison with reference to current research" and Module 66 to "describe the current face review and comparison research and testing on the performance of trained and untrained people". It included 120 PowerPoint slides and a practical exercise simulating face matching in a research context.

The lecture content provided an overview of the research regarding a range of issues relevant to forensic comparison. It began with an introduction to research terms and techniques, to equip participants to assess the quality and relevance of the research they encounter. Next a comprehensive overview of research comparing and contrasting unfamiliar and familiar faces, a range of variables that impact on forensic comparison performance, and the influence of expertise on job performance and learning was provided. Once again, interactive discussions were included throughout the session, which concluded with the practical exercise on face matching in the research context. During the exercise participants were asked to make a same or different judgement for ten pairs of faces, each of which was presented for 30 seconds on a screen at the front of the room. The ten pairs of faces included a mixture of challenging and easy pairs. While the exercise gave participants some insight into their own face matching ability, the aim of the exercise was to highlight the difference between research and on-the-job error rate testing: research tasks often do not reflect real world tasks and therefore care must be taken when extrapolating error rates from research to the real world context.

2.3   Section 3: Bias in forensic comparison

Section 3 of the course fulfilled the requirements of Module 65 of the National Facial Training Framework to "identify the issues and research on subjectivity, confidence, bias, peer and independent review as they affect conclusions". It included 112 PowerPoint™ slides and two practical discussion exercises: the first on identifying sources of bias in the early phases of a criminal investigation, based on the 2015 documentary series 'Making a Murderer' (Making a Murderer, 2015); and the second on applying a decision framework to evaluate the risk of bias in a criminal investigation based on the *R v Honeysett* case (Edmond, 2015).

Lecture material focused on providing participants with a broad introduction to bias with clear definitions of the various types of bias relevant to forensic comparison. The potential sources of bias were discussed and reiterated though the use of practical case studies from recent high-profile criminal cases here in Australia and internationally, including the major case study based on the 'Making a Murderer' documentary. Participants were walked through several aspects of the early stages of the investigation into the sexual assault of Penny Beernsten and the focus on suspect Steven Avery, the subject of the documentary, and asked to note down the sources and types of bias that were present. This was used as the basis for a discussion and reflection on their own work practices and involvement in investigative aspects of a case.

Research relevant to bias in forensic comparison was then reviewed with an emphasis on the evaluation of the risk of bias within the workplace. The session concluded with a practical exercise on applying a framework to evaluate the risk of bias in the forensic comparison work carried out by an expert during the *R v Honeysett* case. Again, participants were walked through the steps taken during the expert review of evidence (in this case a biological anthropologist conducting facial and body comparison from CCTV imagery), and then asked to assess the method against the risk framework. This was then used as a basis for discussion about how the analysis might have been conducted differently to minimise bias, which was a lead in to Section 4 of the course.

2.4   Section 4: Minimising bias and error in forensic comparison

Section 4 of the course fulfilled the requirements of Module 64 of the National Facial Training Framework to "describe the different conclusion opinion scales used in forensics and intelligence, and how they are applied for facial image comparison" and Module 63 to "describe the quality control, quality assurance, and audit processes that exist for facial image comparison to reduce human error". It included 71 PowerPoint™ slides and a group discussion exercise on mitigating bias.

The lecture content provided an overview of the methods empirically demonstrated to minimise bias and error with an emphasis on quality assurance, provision of feedback and proficiency testing within a range of forensic comparison domains. The session concluded with a group discussion exercise to integrate the learning outcomes and reflect on procedures already implemented in the workplace; what needs to change and how that change could be realised.

3.   Course Conduct

The course was conducted over two days at the AFP Forensic Facility Majura in Canberra.

3.1   Participants

Participants consisted of 24 attendees from a range of agencies and disciplines, including:

- Nineteen individuals from the AFP:
    - biological criminalistics/DNA (3)
    - chemical criminalistics/ballistics (1)
    - document examination (3)
    - facial identification (7)
    - fingerprint examination (3)
    - intelligence (1)
    - quality assurance and training (1)
- Two individuals from the Department of Foreign Affairs and Trade, Australian Passport Office (facial identification).
- Two individuals from the Queensland Police Photography Division (facial identification and imagery analysis).
- One individual from the Department of Immigration and Border Protection (facial identification).

4.   Course Evaluation

Feedback from participants was obtained in the form of an optional survey on the final day of the course (see Appendix C). Survey questions aimed to assess the extent to which the course addressed the modules outlined in the National Facial Training Framework, as well as the course delivery.

Participants were asked to rate, on a 7-point Likert scale, the extent to which they agreed with the statement in question. Responses ranged from 1-7, where 1 indicated 'strongly disagree/not endorse' and 7 indicated 'strongly agree/endorse'. Participants were asked to provide any feedback regarding their answer in the space provided and any additional feedback at the end of the survey.

Of the 24 course participants, 18 completed the survey.

4.1 Questions regarding fulfilment of the modules of the National Facial Training Framework

Question 1 assessed the fulfilment of Module 61 of the National Facial Training Framework. The mean participant response was 5.5 (SD = 0.85) with scores ranging from 4-7 (see Figure 1).
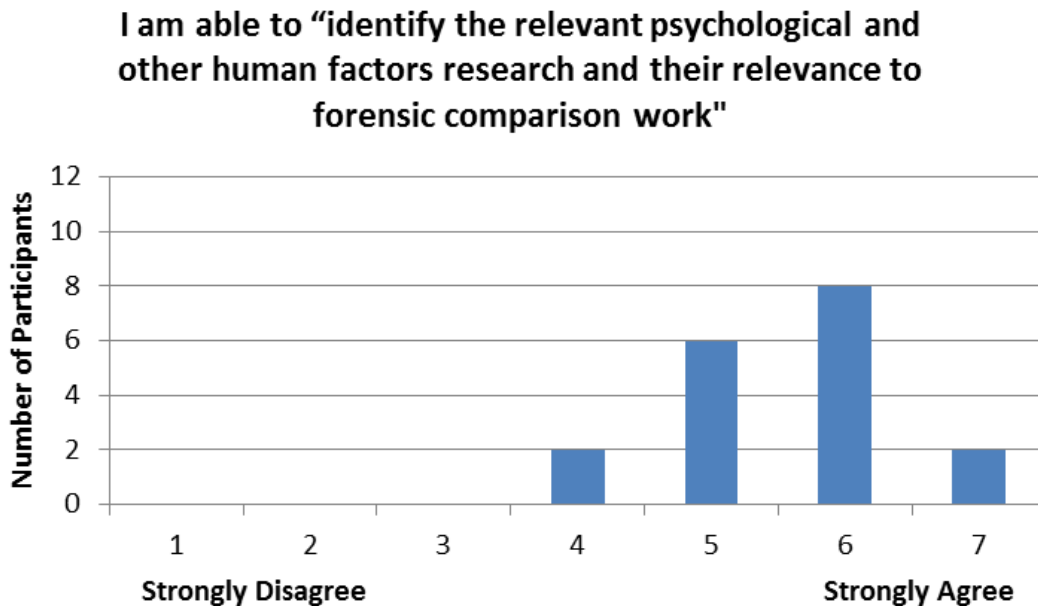

Figure 1 – Module 61 Assessment

Feedback from participants included that: "the psychology and human factors session felt a bit rushed. Feel it would work better to slow the section down to enable attendees to grasp the concepts more readily". Furthermore, some participants offered suggestions regarding how this section may be improved, for example: "have a look at the context of the first module and fit it into the real world examples for the layperson".

Question 2 assessed the fulfilment of Module 62 of the National Facial Training Framework. The mean participant response was 5.7 (SD = 0.95) with scores ranging from 4-7 (see Figure 2).

I am able to "describe the concepts and issues of familiar and unfamiliar observers for human facial comparison with reference to current research"
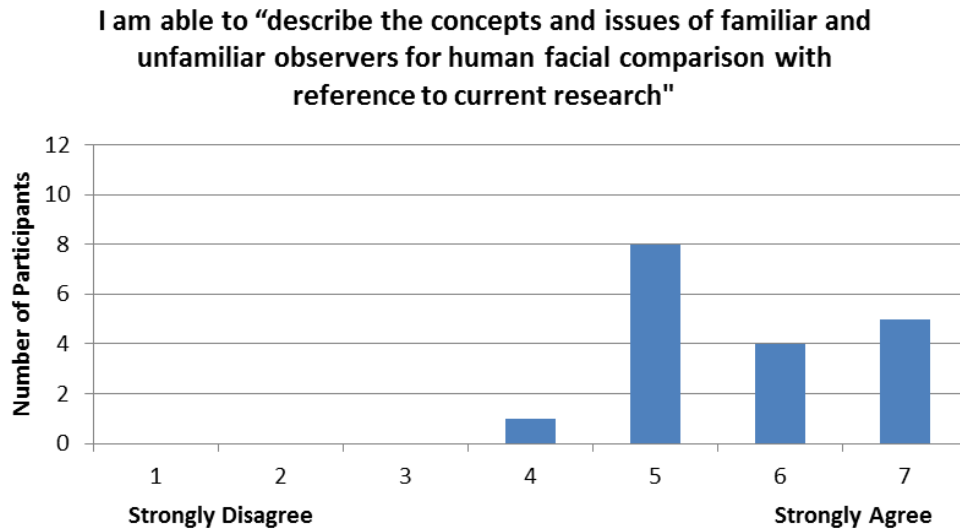

Figure 2 – Module 62 Assessment

Qualitative feedback provided by participants included that: "after reading some of the references supplied I hope to be able to do this better" and "I now have a better understanding of the complexity of information matching".

Question 3 assessed the fulfilment of Module 66 of the National Facial Training Framework. The mean participant response was 5.5 (SD = 0.98) with scores ranging from 4-7 (see Figure 3).

I am able to "describe the current face review and comparison research and testing on the performance of trained an untrained people"

Figure 3 – Module 66 Assessment

Participants found it "interesting that research does not necessarily replicate real world situations", which was a key outcome of this session.

Question 4 assessed the fulfilment of Module 65 of the National Facial Training Framework. The mean participant response was 5.7 (SD = 0.77) with scores ranging from 4.5-7 (see Figure 4).
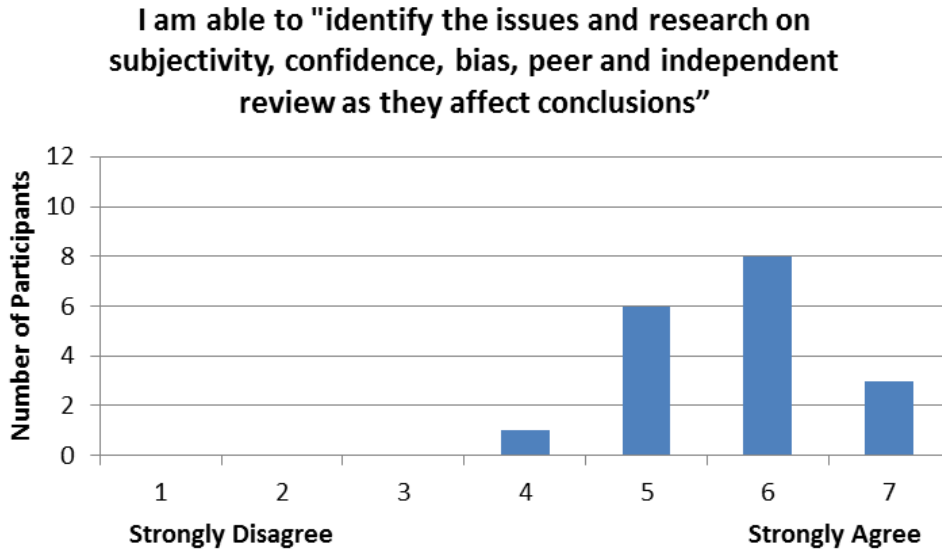

Figure 4 – Module 65 Assessment

A participant commented: "great information presented, will look forward to reviewing the notes".

Question 5 assessed the fulfilment of Module 64 of the National Facial Training Framework. The mean participant response was 5.5 (SD = 0.78) with scores ranging from 4.5-7 (see Figure 5).
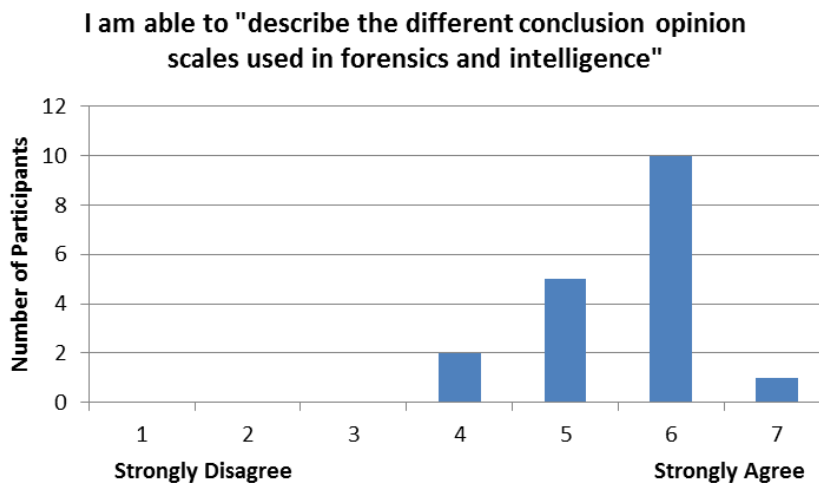

Figure 5 – Module 64 Assessment

A participant thought "the data on how jurors evaluate information we present to them was interesting".

Question 6 assessed the fulfilment of Module 63 of the National Facial Training Framework. The mean participant response was 5.4 (SD = 0.77) with scores ranging from 4-7 (see Figure 6).

Figure 6 – Module 63 Assessment

Participants provided no additional qualitative feedback for this item.

4.2   Questions regarding educational value and course presentation

Question 7 assessed the educational value of the course. The mean participant response was 6.7 (SD = 0.42) with scores ranging from 6-7 (see Figure 7).
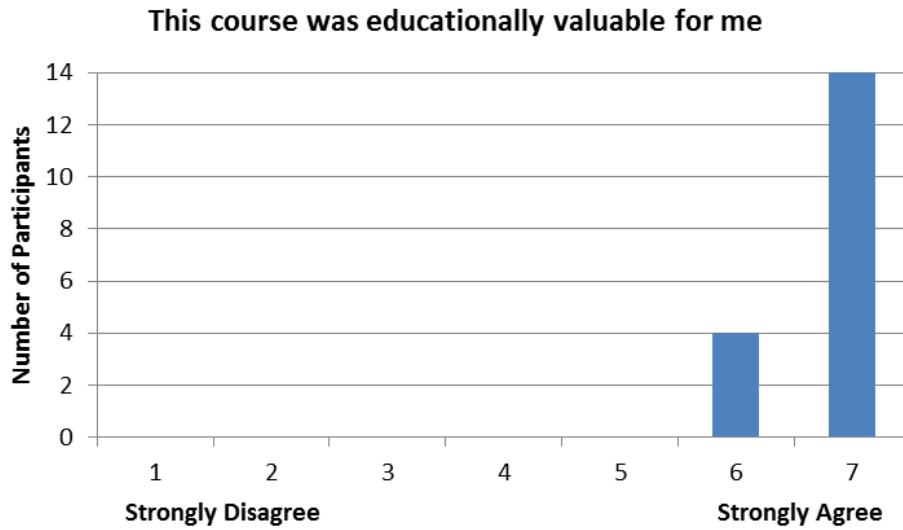
This course was educationally valuable for me

Figure 7 – Assessment of Educational Value

Participants commented that "bias has been debated in our section recently. It was valuable to see the potential outcomes/influences on bias after the decisions have been made" and that "the course provided a good platform to inform further discussion about the impact of bias on procedures". Furthermore a participant commented: "thank you for presenting this insightful course. I hope to be able to use this acquired knowledge to implement some small changes that will make a big difference for some positive gains".

Question 8 assessed the presentation of the course. The mean participant response was 6.5 (SD = 0.61) with scores ranging from 5-7 (see Figure 8).
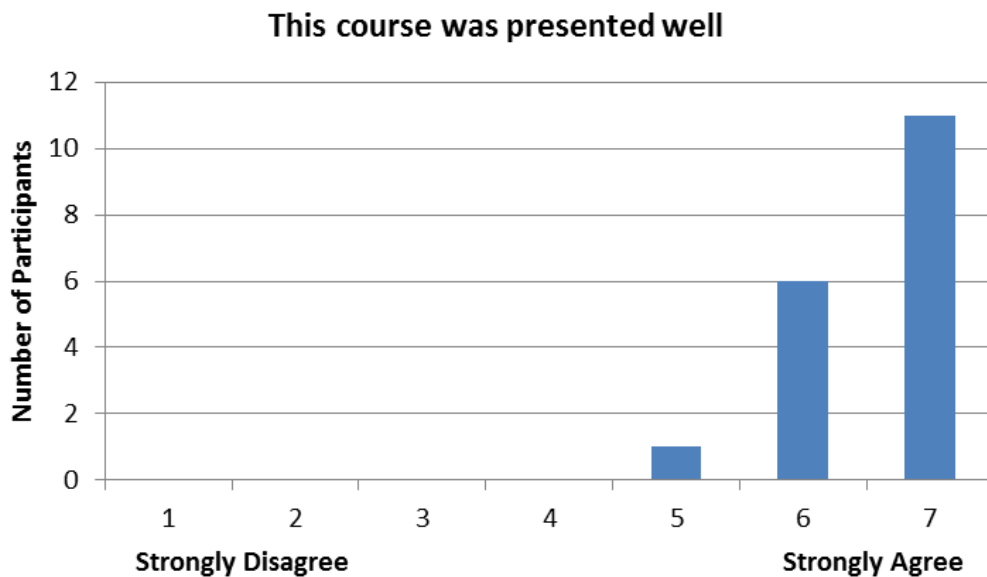
This course was presented well

Figure 8 – Assessment of Course Presentation

Participants noted that "[the presenter] was excellent" and the "time frames 9:30-3 recognised the student's ability to absorb the information". Others commented: "well done, the presenters knew their topics with confidence" and "congratulations to you both, completely interesting, I went home and watched Elizabeth Loftus's TED talk".

Question 9 assessed whether participants would use the information gained in the course in their job. The mean participant response was 6.5 (SD = 0.70) with scores ranging from 5-7 (see Figure 9).
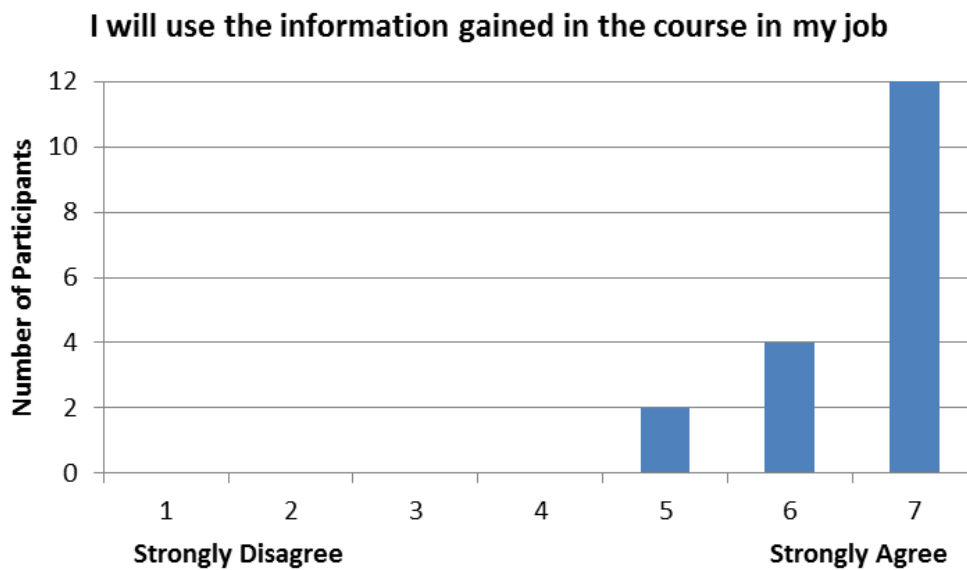
I will use the information gained in the course in my job

Figure 9 – Assessment of Applied Value

A participant noted that the uptake of information would "always depend on how teams and management view it. The ideas are strong but sometimes it may take a significant event to have it improved". However the course "provide(d) a good base to continue research and further reviews of process".

4.3   General feedback

Participants were provided with the opportunity to comment on the course as a whole. Feedback included: "all the information provided has increased my confidence in the training I am undertaking, confirming the practice we follow and the procedures in place. I have take-home information/studies I can use in the context of delivering training and also more perspective on some behavioural aspects to bring to the work place. The notes will be useful for me to review/read and share". Finally, one participant commented: "I've found this course fantastic, I've learnt a lot and have really gotten a lot out of it. Thank you".

5.   Conclusion and Recommendations

'Forensic (Facial) Comparison: Linking the Science and the Practice' aimed to familiarise participants with the psychological science underpinning forensic comparison, with specific focus on facial image comparison, while fulfilling the requirements of six modules from the National Facial Training Framework. The course was run over two days. The first day focused on the psychological theories relevant to forensic comparison and review of relevant facial comparison research. The second day focused on the sources of bias within the forensic comparison domain, how to recognise and mitigate them. Participants completed a course evaluation at the conclusion of the course. Overall participants found the course enjoyable and engaging, generally rating the course to be of high educational value. While it was acknowledged that the course was commissioned by the Facial Identification Team, and as such was necessarily biased towards that discipline, it was still of value for the vast range of other disciplines represented. Participants appreciated the delivery style and the flexible timetabling of the sessions. That being said, some participants also indicated that the delivery of Module 61 (Section 1 of the course) was not sufficiently in depth and did not provide clear linkages to their work in forensic comparison. Therefore, if the course was to be run again in the future it is recommended that:

Recommendation 1: If delivering the course to a diverse audience, content should be supplemented with research and case studies from other forensic comparison disciplines.

Recommendation 2: Section 1 of the course (*An introduction to the psychology and human factors of forensic comparison*) be reviewed to include more real world examples relevant to the field of forensic comparison.

6. References

Dror, I.E. (2015). Cognitive neuroscience in forensic science: understanding and utilizing the human element. *Phil. Trans.R. Soc. B*, 370: 20140255.

Edmond, G. (2015). A Closer Look at *Honeysett*: Enhancing our Forensic Science and Medicine Jurisprudence. *Flinders Law Journal* 17, 287-329.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*, *79*(2), 84.

Making a Murderer (2015), streaming video, viewed October 2016, Netflix, 18 December 2015.

White, D., Kemp, R., Jenkins, R., Matheson, M., & Burton, A. (2014). Passport Officers' Errors in Face Matching. *PLoS ONE*. 9, e103510. doi:10.1371/journal. pone.0103510

Appendix A Course Schedule

Schedule

The course ran from 0930-1530 over two days, with breaks for lunch and morning/ afternoon tea. The schedule is below.

Day 1:

0930	Overview, introductions
0945	Introduction to the psychology and human factors of forensic comparison
1030	Break
1050	Practical exercise: test your perceptual ability Facial image comparison: what does the research say?
1230	Lunch
1330	Facial image comparison: what does the research say? cont.
1445	Break
1500	Practical exercise and group discussion: face matching under research conditions
1530	End Day 1

Day 2:

0930	Review
0945	Bias in forensic comparison Group discussion: 'Making a Murderer' case study
1030	Break
1050	Bias in forensic comparison cont. Minimising bias and error in forensic comparison
1230	Lunch
1330	Minimising bias and error in forensic comparison continued
1415	Small group exercise: what could you apply?
1445	Break
1500	Summary and evaluation

DST-Group-GD-1002

This page is intentionally blank.

Appendix B   Course References

Alenzi, H., & Bindermann, M. (2013). The Effect of Feedback on Face-Matching Accuracy. *Applied Cognitive Psychology*. DOI: 10.1002/acp.2968

Alenezi,H., Bindemann, M., Fysh, M., & Johnston, R. (2015). Face matching in a long task: enforced rest and, desk-switching cannot maintain identification accuracy. *PeerJ*. 3:e1184 https://doi.org/10.7717/peerj.1184

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. (2015). Telling faces together: Learning new faces through exposure to multiple instances, *The Quarterly Journal of Experimental Psychology*, 68, 2041-2050.

ANZFSS (2016). Code of Professional Practice. http://anzfss.org/code-of-professional-practice/

Bindermann, M., Attard, J., & Johnston, R. (2014). Perceived ability and actual recognition accuracy for unfamiliar and famous faces. *Cogent Psychology*. 1, 1-15.

Bindemann, M., Attard, J., Leach, A., & Johnston, R. (2013). The effect of image pixilation on unfamiliar face matching. *Applied Cognitive Psychology*, 27, 707-717.

Bindemann, M., Avetisyan, M., & Blackwekk, K. (2010). Finding Needles in Haystacks: Identity Mismatch Frequency and Facial Identity Verification. *Journal of Experimental Psychology:Applied*. 4, 378-386.

Bindemann, M., Fysh, M., Cross, K., & Watts, R. (2016). Matching Faces Against the Clock. *iPerception*, 7, doi: 10.1177/2041669516672219.

Bobak, A., Dowsett, A., & Bate, S. (2016). Solving the Border Control Problem: Evidence of Enhanced Face Matching in Individuals with Extraordinary Face Recognition Skills. *PLoS 11*: e0148148.

Bobak, A., Hancock, P., & Bate, S. (2015). Super-recognisers in Action: Evidence from Face-matching and Face Memory Tasks. *Applied Cognitive Psychology*. 30, 81-91.

Bressan P. & Dal Martello M. F. (2002). Talis pater, talis filius: perceived resemblance and the belief in genetic relatedness. *Psychological Science*, 13, 213–219.

Bruer, J. (2010). The Mind's Journey from Novice to Expert. Saint Louis, MO: James S. McDonnell Foundation.

Bruce, V., Henderson, Z., Newman, C., & Burton, A.M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, 7, 207-218.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology-Applied*. 5, 339-360.

Burton, A., Jenkins, R., Hancock, P., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*. 51, 256-284.

Burton, A. M., Miller, P., Bruce, V., Hancock, P. J. B., & Henderson, Z. (2001). Human and automatic face recognition: a comparison across image formats. *Vision Research*, 41, 3185-3195.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, 42, 286-291.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*. 10, 243- 248.

Butler, J.M. (2014). Validation overview. https://www.nist.gov/sites/default/files/documents/forensics/01_ValidationWebinar-Butler-Aug2014.pdf

Calleja, J. (2016). Human Identification at a Distance: The Impact of Image Quality and Image Restoration Techniques on Human Face Matching Performance. Unpublished Honours Thesis, School of Psychology, University of Adelaide.

Calic, D. (2012). From the laboratory to the real world: Evaluating the impact of impostors, expertise and individual differences on human face matching performance (Doctorial dissertation, The University of Adelaide). Retrieved from https://digital.library.adelaide.edu.au/dspace/bitstream/2440/91444/3/02whole.pdf

Charlton, D., Dror, I., & Fraser- Mackenzie, P. (2008). A qualitative study investigating the emotional rewards and motivating factors associated with forensic finger print analysis. Technical Report, University of Southampton, School of Psychology, Southampton, UK.

Clutterbuck, R., & Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology*. 17, 97-116.

Dowsett, A., & Burton, A. (2015). Unfamiliar face matching: Pairs out-perform individuals and provide a route to training. *British Journal of Psychology*, 3, 433-445.

Davies, G. (1997). Video surveillance: The impact of colour on person description and identification. *Proceedings of the British Psychological Society*, 5, 2–67.

Davis, J. P., Lander, K., Evans, R., & Neville, M. (2012, April). Facial identification from CCTV: investigating predictors of exceptional performance in police officers. Paper presented at the European Association of Psychology and Law, Cyprus.

Dick, D. (2015, July). An evolving natural experiment in the determination of human error in operational environments. Paper presented at the International Conference on Evidence Law and Forensic Science, Adelaide, Australia.

Dhamecha TI, Singh R, Vatsa M, Kumar A (2014) Recognizing Disguised Faces: Human and Machine Evaluation. *PLoS ONE* 9(7): e99212.

Doddington, G., Liggett, W., Martin, A., Przybocki, M. & Reynolds, D. (1998). *Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation* at http://www.itl.nist.gov/iad/mig//publications/storage_paper/icslp_98.pdf (accessed Sep 21 2009).

Dowsett, A., & Burton, A. (2015). Unfamiliar face matching: Pairs out-perform individuals and provide a route to training. *British Journal of Psychology*, 3, 433-445.

Dowsett, A., Sandford, A., & Burton, M. (2016) Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *The Quarterly Journal of Experimental Psychology*, 69, 1-10.

Dror, I. (2012). Combating bias: the next step in fighting cognitive and psychological contamination. *Journal of Forensic Sciences*, *57*(1): 276-277.

Dror, I. E. (2013). Practical solutions to cognitive and human factor challenges in forensic science. *Forensic Science Policy & Management: An International Journal*, *4*(3-4), 105-113.

Dror, I.E. (2015). Cognitive neuroscience in forensic science: understanding and utilizing the human element. *Phil. Trans.R. Soc. B*, 370: 20140255.

Dror, I.E. (2016). A Hierarchy of Expert Performance. Journal of Applied Research in Memory and Cognition, 5 (2), 121-127.

Dror, I., Busemeyer, J., & Basola, B. (1999). Decision making under time pressure: An independent test of sequential sampling models. Memory and Cognition. 27, 713-725.

Dror, I. E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., & Rosenthal, R. (2011). Cognitive issues in fingerprint analysis: Inter- and intra-expert consistency and the effect of a 'target' comparison. *Forensic Science International, 208(1-3), 10-17.* Dror, I. E., Thompson, W. C.,

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic science international*, *156*(1), 74-78.

Dror, I., & Fraser-Mackenzie, P. (2008). Cognitive Biases in Human Perception, Judgement, and Decision Making: Bridging Theory and the Real World. In Rossmo, K (Ed.), Criminal Investigative Failures (pp. 53-67). Boca Raton, Florida: CRC Press.

Dror, I. & Mnookin, J.L. (2010). The use of technology in human expert domains: challenges and risks arising from the use of automated fingerprint identification systems in forensic science. *Law, Probability and Risk* **9**(1), 47−67.

Dror, I. E., Peron, A. E., Hind, S. L., & Charlton, D. (2005). When emotions get the better of us: the effect of contextual top down processing on matching fingerprints. *Applied Cognitive Psychology*, *19*(6), 799-809.

Dror, I. E., Wertheim, K., Fraser-Mackenzie, P., & Walajtys, J. (2012). The Impact of Human-Technology Cooperation and Distributed Cognition in Forensic Science: Biasing Effects of AFIS Contextual Information on Human Experts. *Journal of Forensic Sciences*, *57*(2), 343-352.

Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors, 44(1), 79-94.*

Edmond, G. (2015). 'A Closer Look at *Honeysett*: Enhancing our Forensic Science and Medicine Jurisprudence' *Flinders Law Journal* 17, 287-329.

Edmond, G & San Roque, M. (2014). 'Before the High Court - Honeysett v The Queen: Forensic science, "specialised knowledge" and the Uniform Evidence Law', *Sydney Law Review*, 36, 323-344.

Edmond, G., Found, B., Martire, K., Ballantyne, K., Hamer, D., Searston, R., Thompson, M., Cunliffe, E., Kemp, R., San Roque, M., Tangen, J., Dioso-Villa, R., Ligertwood, A., Hibbert, D., White, D., Ribeiro, G., Porter, G., Towler, A. & Roberts, A. (2016) Model forensic science, *Australian Journal of Forensic Sciences*, 48:5, 496-537.

Ertmer, P., & Newby, T. (1996). The expert learner; Strategic, self regulated and reflective. *Instructional Science*. 24, 1-24.

Ericcson, A., Krampe, R., & Tesch- Romer, C. (1996). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*. 100, 363-406

Federal Bureau of Investigation (FBI) (2016). PCAST Response. Retrieved from https://www.fbi.gov/file-repository/fbi-pcast-response.pdf/view

Garner, R., & Alexander, P. (1989). Metacognition: Answered and Unanswered questions. *Educational Psychologists*. 24, 143-158.

Fishburn, P. (1970). *Utility Theory for Decision Making*. New York: Wiley.

Forensic Science Regulator (2015). Cognitive bias effects relevant to forensic science examinations. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/510 147/217_FSR-G-217_Cognitive_bias_appendix.pdf.

Found, B. & Edmond, G. (2012). Reporting on the comparison and interpretation of pattern evidence. *Australian Journal of Forensic Science* 44(2):193–196.

Found, B., & Ganas, J. (2013). The management of domain irrelevant context information in forensic handwriting examination casework. *Science & Justice*, 53(2), 154-158.

Furl, N., Phillips, P. J., & O'Toole, A. J. (2002). Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. *Cognitive Science*, 26(6), 797-815.

Gigerenzer, G., Hertwig, R., Van Den Brock, E., Fasolo, B., & Katsikopoulous, K. (2005). A 30% chance of rain tomorrow": how does the public understand probabilistic weather forecasts? *Risk Analysis*, 25, 623- 629.

Grother, P. J., Quinn, G. W., & Phillips, P. J. (2011). Multiple Biometric Evaluation (MBE) 2010: Report on the Evaluation of 2D Still-Image Face Recognition Algorithms. Retrieved 23 September 2011, from http://www.nist.gov/customcf/get_pdf.cfm?pub_id=905968.

Hall, L., & Player, E. (2008). Will the introduction of emotional context affect fingerprint analysis in decision-making? *Forensic Science International*. 181, 36-39.

Hanson, J. (2015). The Admiralty Code: A Cognitive Tool for Self-Directed Learning. *International Journal of Learning, Teaching and Educational Research*. 14(10): 97-115.

Henderson, Z., Bruce, V., & Burton, A. M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology*, 15, 445-464.

Heuer, R. (1999) *The Psychology of Intelligence Analysis*. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/PsychofIntelNew.pdf

Heyer, R. (2013). *Understanding One-to-Many Unfamiliar Face Matching in the Operational Context: The Impact of Candidate List Size, Expertise, and Decision Aids on the Performance of Facial Recognition System Users* (Doctoral dissertation. The University of Adelaide).

Heyer, R. (2015). Technology and Cognitive Bias. *Wiley Encyclopedia of Forensic Science*, 1-6.

Heyer, R., Semmler, C., & McLindin, B. (2010, September). *Identification using an automated facial recognition system: towards an understanding of human operator decision making*. Paper presented at the International Symposium on the Forensic Sciences (ANZFSS'10), Sydney.

Hildebrandt, A., Sommer, W., Herzmann, G., & Wilhelm, O. (2011). Structural invariance and age-related performance differences in face cognition. *Psychology and Aging*, 25, 794-810.

Howes, L.M., Kirkbride, K.P., Kelty, S.F., Julian, R.Kemp, N. (2013). Forensic scientist's conclusions: how readable are they for non-scientist report-users? *Forensic Science International*. 231(1-3): 102-112.

DST-Group-GD-1002

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42-52.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology*, 11, 211-222.

Kirkpatrick, D. L. (1994) Evaluating Training Programs: The four levels. Berrett-Koehler, San Francisco.

Klontz, J.C.; & Jain, A.K. (2013). A Case Study of Automated Face Recognition: The Boston Marathon Bombings Suspects. *Computer*, 46(11), 91-94.

Knapp, K. (2012). Some differences between experts and novices. Saddle River, NJ.

Kramer, R., & Ritchie, K. (In Press). Distinguishing Superman: How Glasses Affect Unfamiliar Face Matching. *Applied Cognitive Psychology*.

Lee, W. J., Wilkinson, C., Memon, A., & Houston, K. A. (2009). Matching unfamiliar faces from poor quality CCTV footage: an evaluation of the effect of training on facial identification ability. *Axis*. 1, 19-28.

Leinsinger, G., Born, C., Meindl, T., Bokde, A., Britsch, S., Lopez-Bayo, P., et al. (2007). Age-dependent differences in human brain activity using a face- and location- matching task: An fMRI study. *Dementia and Geriatric Cognitive Disorders*, 24, 235-246.

Majamaa, H., & Ytti, A. (1996). A survey of the conclusions drawn of similar footwear cases in various crime laboratories. *Forensic Science International*, 82, 109–120

Matsuyoshi, D., Morita, T., Kochiyama, T., Tanabe, H. C., Sadato, N., & Kakigi, R. (2015). Dissociable cortical pathways for qualitative and quantitative mechanisms in the face inversion effect. *The Journal of Neuroscience*. 35, 4268-4279.

McBain, R., Norton, D., & Chen, Y. (2009). Females excel at basic face perception. *Acta Psychologica*, 130, 168-173.

McQuiston-Surrett, D., & Sacks, M. (2008).Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact. *Hastings Law Journal*, 59, 1159-1190.

Meissner, C. A., Kornfield, I., Krane, D., Saks, M., & Risinger, M. (2015). Context management toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of Forensic Sciences*, 60(4), 1111-1112.

Meissner, C. A., Susa, K. J., & Ross, A. B. (2013). Can I see your passport please? Perceptual discrimination of own-and other-race faces. *Visual Cognition*, 21, 1287-1305.

Megreya, A. M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: Evidence from matching tasks. *Acta Psychologica*, 137, 83-89.

Megreya, A., Sanford, A., & Burton, M. (2013). Matching Face Images Taken on the Same Day or Months Apart: the Limitations of Photo ID. *Applied Cognitive Psychology*, 27, 700-706.

Miller, L.S. (1987). Procedural bias in forensic science examinations of human hair. *Law and Human Behaviour*, 11: 157.

Moore, R., & Johnston, R. (2013). Motivational Incentives Improve Unfamiliar Face Matching Accuracy. *Applied Cognitive Psychology*.27, 754–760.

Mullen, C., Spence, D., Moxey, L., & Jamieson, A. (2013). Perception problems of the verbal scale. *Science and Justice*. 54, 154–158.

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. (2013). Are forensic experts biased by the side that retained them? *Psychological Science*, 24, 1889 –1897.

Nakhaeizadeh, S., Dror, I. E., & Morgan, R. M. (2014). Cognitive bias in forensic anthropology: visual assessment of skeletal remains is susceptible to confirmation bias. *Science & Justice*, 54(3), 208-214.

National Academy of Sciences (NAS) (2009*). Strengthening Forensic Science in the United States: A Path Forward*. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf.

Norell, K., Läthén, K., Bergström, P., Rice, A., Natu, V., & O'Toole, A. (2015). The effect of image quality and forensic expertise in facial image comparisons. *Journal of Forensic Science*, 60, 331–340.

Osborne, N. K., Zajac, R., & Taylor, M. C. (2015). Bloodstain Pattern Analysis and Contextual Bias. *Wiley Encyclopedia of Forensic Science*.

O'Toole, A. J., Phillips, P. J., Jiang, F., Ayyad, J., Pénard, N., & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces over changes in illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29, 1642-1646.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972-987.

Papesh, M., & Goldinger, S. (2014). Infrequent identity mismatches are frequently undetected. *Attention and Perceptual Psychophysics*. 76, 1335-1349.

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image Vision Computing*, 16(5), 295-306.

President's Council of Advisors on Science and Technology (PCAST) (2016). *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*. Retrieved from: https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf

Prince, J. (2013). *To examine emerging police use of facial recognition systems and facial image comparison procedures*. Retrieved from http://churchilltrust.com.au/site_media/fellows/2012_Prince_Jason.pdf

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, 11(1), 37–43

Pronin, E., Lin, D., & Ross, L. (2002): The Bias Blind Spot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin*, 28, 369-381.

Robertson, D., Noyes, E., Dowsett, A., Jenkins, R., & Burton, A. (2016). Face Recognition by Metropolitan Police Super-Recognisers. *PLoS ONE 11*, e0150036.

Ross, K. G., Phillips, J. K., Klein, G., & Cohn, J. (2005). *Creating expertise: A framework to guide technology-based training*. (Final Technical Report for Contract #M67854-04-C-8035 for the Marine Corps Systems Command/Program Manager for Training Systems). Fairborn, OH: Klein Associates.

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

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, 54, 35-54.

Schraw, G., & Dennison, S. (1994). Assessing metacognitive Awareness. *Contemporary Educational Psychology*. 19, 460-475.

Semmler, C., Ma-Wyatt, A., Heyer, R., & MacLeod, V. (2012a, September). *The Impact of Individual Differences and Eye Movements on Facial Comparison Performance*. Paper presented at the International Symposium on the Forensic Sciences (ANZFSS'12), Hobart.

Sinha, P., Balas, B. J., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*. 94, 1948-1962.

Sporer, S. L., Trinkl, B., & Guberova, E. (2007). Matching faces: Differences in processing speed of out-group faces by different ethnic groups. *Journal of Cross-Cultural Psychology*, 38, 398 412.

Stoel, R. D., Berger, C., Kerkhoff, W., Mattijssen, E. J. A. T., Dror, I. E., Hickman, M., & Strom, K. (2014). Minimizing contextual bias in forensic casework. *Forensic Science and the Administration of Justice: Critical Issues and Directions*, 67.

Stoel, R. D., Dror, I. E., & Miller, L. S. (2014). Bias among forensic document examiners: Still a need for procedural changes. *Australian Journal of Forensic Sciences*, 46(1), 91-97.

Sulner, A. (2015). Handwriting: Cognitive Bias. *Wiley Encyclopedia of Forensic Science*.

Schwaninger, A. (2006, November). *Airport security human factors: From the weakest to the strongest link in airport security screening*. In Proceedings of the 4th International Aviation Security Technology Symposium, Washington, DC, USA, November (pp. 265-270).

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect: Grotesque faces from relative spaces. *Perception*, 40, 628-630

Tetlock, P. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, 48, 227-236.

Towler, A., White, D., & Kemp, R. (2014). Evaluating training methods for facial image comparison: The face shape strategy does not work. *Perception*, 43, 214-218.

Vast, R., Lee, M., & Butavicus, M. (2004). *Fatigue, time pressure and distractions when doing face recognition task*. Biometrics Institute Conference, Sydney, New South Wales.

Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own - race effect and intergroup experience. *Applied Cognitive Psychology*, 20, 461-475.

Wells, G.L, Wilford, M.M. & Smalarz, L. (2013). Forensic Science Testing: The Filler-Control Method for Controlling Contextual Bias, Estimating Error Rates, and Calibrating Analysts' Reports. *Journal of Applied Research in Memory and Cognition* 2(1) : 53–55.

White, D., Burton, A., Jenkins, R., & Kemp, R. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied*, 20, 166-173.

White, D., Burton, A. M., Kemp, R. I., & Jenkins, R. (2013). Crowd effects in unfamiliar face matching. *Applied Cognitive Psychology*, 27, 769-777.

White, D., Kemp, R. I., Jenkins, R., & Burton, A. M. (2014). Feedback training for facial image comparison. *Psychonomic Bulletin & Review*, 21(1), 100– 106.

White, D., Kemp, R., Jenkins, R., Matheson, M., & Burton, A. (2014). Passport Officers' Errors in Face Matching. *PLoS ONE. 9*, e103510. doi:10.1371/journal. pone.0103510

White, D., Phillips, P. J., Hahn, C. A., Hill, M., & O'Toole, A. J. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the Royal Society of London B: Biological Sciences*. 282, 1814-1822.

Wilkinson, C., & Evans, R. (2009). Are facial image analysis experts any better than the general public at identifying individuals from CCTV images? *Science and Justice*. 49, 191-196.

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta psychologica*, 114(1), 101-114.

Appendix C   Course Feedback Form

Feedback Form
Forensic (Facial) Comparison: Linking the Science and the Practice

Thank you for participating in Forensic (Facial) Comparison: Linking the Science and the Practice. The course covered six modules from the Facial Biometrics centre of Expertise's National Facial Training Framework.

We would like to gain some feedback regarding your experiences with the course and material presented. Please fill in the feedback form and return it to one of the course facilitators. This feedback is not mandatory but it is greatly appreciated by the facilitators.

The scale provided consists of a rating system from 1-7. Where 1 indicates **strongly disagree/endorse** and 7 indicates **strongly agree/endorse.** Please provide any additional feedback regarding your answer in the space provided.

1. Please rate on the scale provided to what degree you feel able to "identify the relevant psychological and other human factors research and their relevance to forensic comparison work".

```
  |------+------+------+------+------+------|
  1      2      3      4      5      6      7
Strongly                                Strongly
disagree                                   agree
```

_____

_____

_____

2. Please rate on the scale provided to what degree you feel able to "describe the concepts and issues of familiar and unfamiliar observers for human facial comparison with reference to current research".

```
  |------+------+------+------+------+------|
  1      2      3      4      5      6      7
Strongly                                Strongly
disagree                                   agree
```

_____

_____

_____

DST-Group-GD-1002

3. Please rate on the scale provided to what degree you feel able to "describe the current face review and comparison research and testing on the performance of trained and untrained people".

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

4. Please rate on the scale provided to what degree you feel able to "identify the issues and research on subjectivity, confidence, bias, peer and independent review as they affect conclusions".

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

5. Please rate on the scale provided to what degree you feel able to "describe the different conclusion opinion scales used in forensics and intelligence, and how they are applied for facial image comparison".

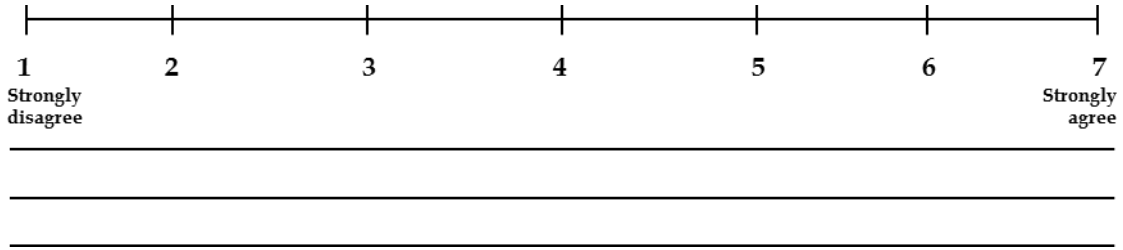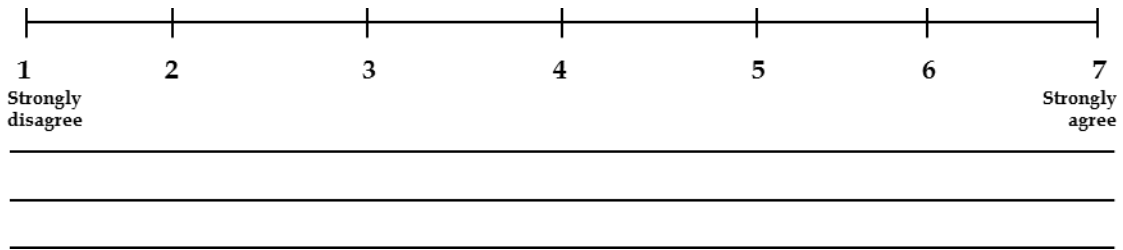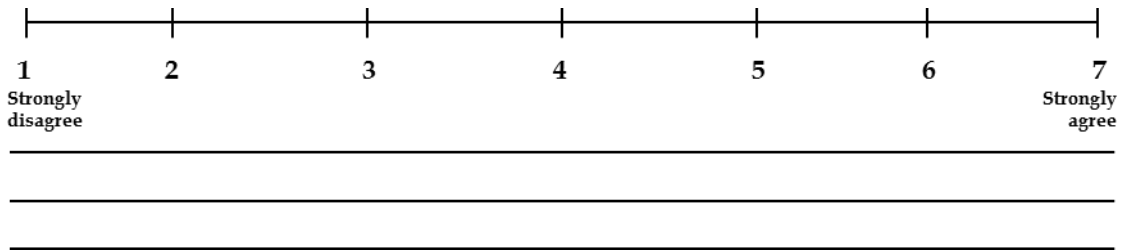| 1	2	3	4	5	6	7
Strongly disagree						Strongly agree

6. Please rate on the scale provided to what degree you feel able to "describe the quality control, quality assurance, and audit processes that exist for facial image comparison to reduce human errors".
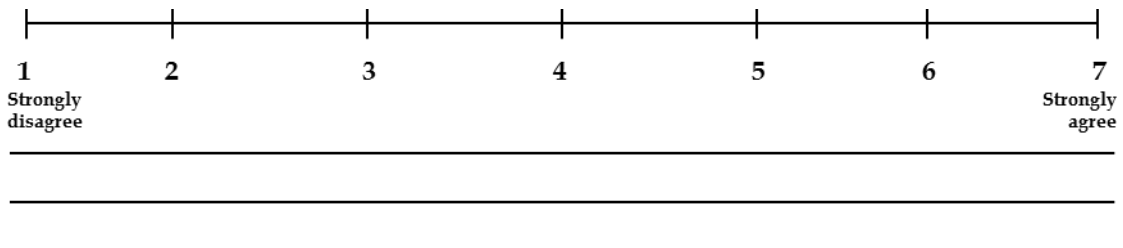
1	2	3	4	5	6	7
Strongly disagree						Strongly agree

_____

_____

_____

7. This course was educationally valuable for me.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

_____

_____

_____

8. This course was presented well.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

_____

_____

_____

9. I will use the information gained in the course in my job.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

_____

_____

_____

Any other feedback:

DEFENCE SCIENCE AND TECHNOLOGY GROUP DOCUMENT CONTROL DATA		1. DLM/CAVEAT (OF DOCUMENT)	
2. TITLE Forensic (Facial) Comparison: Linking the Science and the Practice - Course Overview and Evaluation		3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED LIMITED RELEASE USE (U/L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)	
4. AUTHOR(S) Jessica O'Rielly & Rebecca Heyer		5. CORPORATE AUTHOR Defence Science and Technology Group PO Box 1500 Edinburgh SA 5111	
6a. DST GROUP NUMBER DST-Group-GD-1002	6b. AR NUMBER AR-017-159	6c. TYPE OF REPORT General Document	7. DOCUMENT DATE April 2018
9.TASK NUMBER 17/548	10.TASK SPONSOR Australian Federal Police (PA S-010/2016)	10. RESEARCH DIVISION NSI Division	
11. MSTC Intelligence Analytics		12. STC Biometrics	
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT *Approved for public release* OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111			
16. DELIBERATE ANNOUNCEMENT No limitations			
17. CITATION IN OTHER DOCUMENTS Yes			
18. RESEARCH LIBRARY THESAURUS Face recognition, training, biometric identification, cogntive psychology			
19. ABSTRACT The Defence Science and Technology Group was commissioned by the Australian Federal Police to develop and deliver a course to familiarise participants with the psychological science underpinning forensic comparison, with specific focus on facial image comparison, while fulfilling the requirements of six modules from the National Facial Training Framework. This paper provides an overview of the course content, its delivery and evaluation. It concludes with some recommendations for the future, should the course be delivered again.			