TEST-RETEST RELIABILITY OF COMPUTER BASED BINOCULAR FUNCTION MEASURES

Maria Gavrilescu[1], Renee Karas[2], Daniel White[3], Peter Gibbs[1], Amanda Douglass[4#], Larry Abel[4], Darren Doyle[5], Marc Winterbottom[6], Steven Hadley[6]

1. Aerospace Division, Defence Science and Technology Group, Victoria, Australia, 2.Advanced VTOL Technologies, Victoria, Australia, 3. ScienceFX, Victoria, Australia, 4. Department of Optometry & Vision Sciences, University of Melbourne, Australia, 5. Defence and Security Analysis Division, Human and Social Sciences Group, Defence Science Technology Laboratory, United Kingdom, 6. Operational Based Vision Assessment Laboratory, U.S. Air Force School of Aerospace Medicine Research Division, Wright-Patterson AFB, OH, USA. #Presenting author.

INTRODUCTION

- Military pilots experience visual fatigue when using optically misaligned Head Mounted Displays (HMD) and binocular Night Vision Devices (NVD).
- Individual optometric parameters play an important role in modelling the effects of visual fatigue:
  1. The individual variability in tolerance to visual fatigue is based on an individual's binocular system and optometric parameters (1,2)
  2. Optical misalignment can induce temporary changes in optometric parameters (4-7).

OBJECTIVES

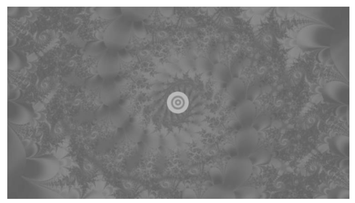
- To identify accurate and reproducible optometric predictors of degraded task performance that may be induced by visual fatigue when wearing misaligned NVD.
- To identify optometric parameters suitable to assess the temporary disruption in binocular function induced by optical misalignment.

METHODS

- Currently used methods to assess binocular function provide only coarse measures that are difficult to relate to individual variability in task performance.
- This study employed computer based vision tests developed by U.S. Air Force School of Aerospace Medicine's Operational Based Vision Assessment Laboratory as part of the Automated Vision Testing (AVT) system (3).
- Near and far stereoacuity and binocular fusion tests as implemented in AVT were investigated in terms of test–retest reliability and practice effects.

Data description and experimental set up

- 35 observers (mean age = 29.7 years) participated in two sessions of testing.
- Three computerised visual tests available on the AVT system; near stereoacuity, far stereoacuity and binocular fusion range.
- Same tests performed at two different sessions (approximately one week apart).
- Participants wore 3D glasses and viewed a 3D display.
- A subset of observers (n=4) participated in ten consecutive days of testing in order to determine any learning/training effects.

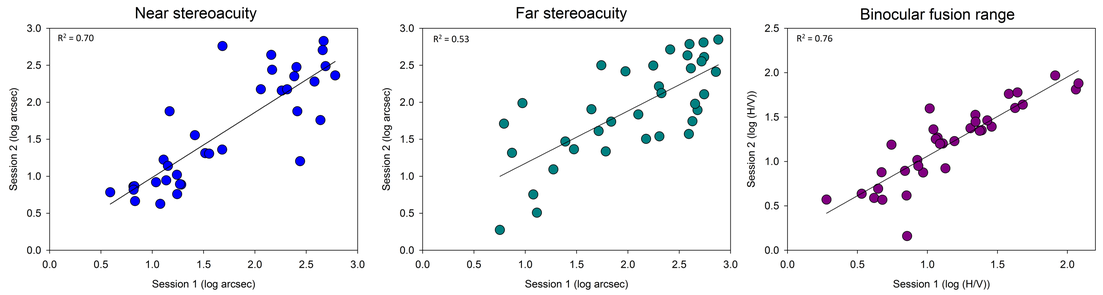


Stereoacuity test: two rings were presented on the centre of the screen for 2 seconds. In a two alternative forced choice, participants indicated whether the smaller centre ring was presented "in front" of the larger ring or "behind" it by button press. Auditory feedback was provided for correct and incorrect responses.

Binocular fusion task: participants indicated when a small white ring split into two clearly defined rings (break) and then again when the rings became one single ring (recovery). This was done four times horizontally (2 x uncrossed and 2 x crossed) as well as four times vertically (2x right up and 2 x right down).

RESULTS

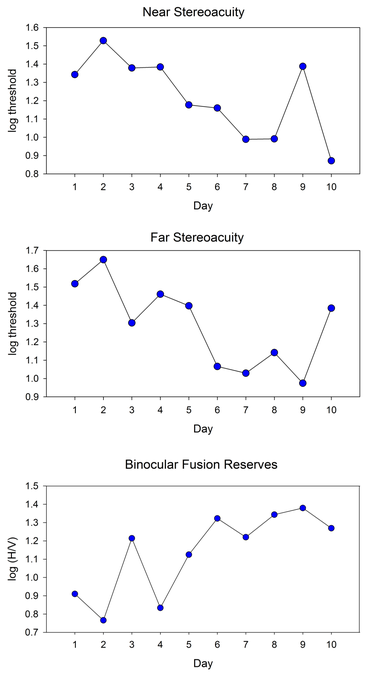
Test-retest reliability



Results from left to right for the near stereoacuity, far stereoacuity and binocular fusion tests.

- Log thresholds for both stereoacuity tests were obtained and compared from one session to the other. Lower thresholds indicate good performance.
- For binocular fusion range, a log ratio of horizontal and vertical recoveries was calculated and used to compare performance between sessions 1 and 2 based on previous studies (3). A higher value indicates good binocular function.
- The binocular fusion test was the most reliable of the tests (R^2 = 0.76) followed by Near stereoacuity (R^2= 0.70).
- The far stereoacuity test did not perform as reliably as the other two (R^2 = 0.53).

Practice effects

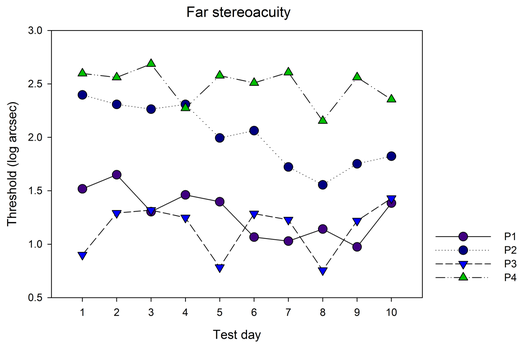


Example data from an observer for the 10 days of consecutive testing. Near and far stereoacuity results appear to be variable across the days but show a trend for increased performance with increasing number of days tested. The binocular fusion ratio increases over time also indicating improved performance.

Table 1. Percentage change between first and tenth sessions. Observers improved for every test (except one participant on one of the tests- far stereo).

Participant	Near Stereoacuity	Far Stereoacuity	Fusion range
1 (good acuity)	35.09	8.80	39.48
2 (medium acuity)	6.79	23.92	22.20
3 (good acuity)	34.87	-58.78	29.64
4 (poor acuity)	9.30	9.39	N/A*

*Participant 4 had no fusion reserve and showed no trend towards improvement. Values ranged from -0.08 to 0.03 over the 10 days.



Data from 10 days of testing for four observers for the far stereoacuity test. Overall, there appears to be some improvement on task over time, however, learning effects appear to be subject specific.

CONCLUSIONS

- Near stereoacuity and binocular fusion tests show good reproducibility.
- Near stereoacuity appears to show poorer reliability in people with medium levels of stereoacuity when compared to those at either extreme (very good or poor).
- The better reliability of the near stereoacuity and binocular fusion makes these tests suitable to be investigated as both predictors of visual fatigue and performance degradation as well as pre and post-task tests to investigate disruptions to binocular function after wearing NVDs.
- As practice effects might impact on test-retest reliability our experimental protocol will be modified in a way to account for these effects (e.g., allowing more time between test and retest).
- Practice effects appear to be subject specific with complex patterns and not monotonous improvements.
- Further investigations are required to fully characterise practice effects on measuring these parameters with the AVT system.
- The poor reliability of the far stereoacuity test makes it unsuitable for pre- and post-task use but may possibly be used to investigate optometric predictors of visual fatigue susceptibility.
- Future work aims to assess other optometric parameters such as phoria and accommodation.

REFERENCES

1. Rash CE, McLean WE, Mozo BT, Licina JR, McEntire BJ. (1998): Human Factors and Performance Concerns for the Design of Helmet-Mounted Displays RTO HFM Symposium on Current Aeromedical Issues in Rotary Wing Operations. San Diego, USA: Published in RTO MP-19.
2. Self HC. (1986): Optical tolerances for alignment and image differences for binocular helmet-mounted displays. Dayton, OH: Armstrong Aerospace Medical Research Laboratory. AAMRL-TR-86-019.
3. Winterbottom M, Lloyd C, Gaska J, Wright S, Hadley S. Stereoscopic Remote Vision System Aerial Refueling Visual Performance. Electronic Imaging. 2016 Feb 14;2016(5):1-0.
4. Sheehy JB, Wilkinson M. (1989): Depth perception after prolonged usage of night vision goggles. Aviat Space Environ Med 60(6):573-9.
5. Mon-Williams M, Wann JP, Rushton S. (1993): Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. Ophthalmic Physiol Opt 13(4):387-91.
6. Kalich ME, Lont LM, van de Pol C, Rash CE. (2004): Partial-overlap biocular image misalignment tolerance. Helmet- and Head-Mounted Displays IX: Technologies and Applications. 1 ed. Orlando, FL, USA: SPIE. p 261-272.
7. Kalich ME, Rash CE, van de Pol C, Rowe TL, Lont LM, Peterson RD. (2003): Biocular image misalignment tolerance. Helmet- and Head-Mounted Displays VIII: Technologies and Applications. 1 ed. Orlando, FL, USA: SPIE. p 284-295.

DISCLOSURE INFORMATION

I have the following financial relationships to disclose:

Grant/Research support from:
Defence Science Institute, Pilot Competitive Evaluation Research Agreement (CERA) Program 2016.

I will not discuss off-label use and/or investigational use in my presentation