Comparing Performance Between Night-Vison Goggles

Zander N. Miller, Elizabeth Merwin, Joseph O’Brien, Elijah Montano, & Haley Picciolo

Department of Human Factors Psychology and Behavioral Neurobiology, Embry-Riddle Aeronautical University

HFS 615: Sensation and Perception

Dr. Alex Chaparro

September 15, 2024

Proposal to Compare Visual Acuity Between Night-Vison Goggles

A review of existing literature was performed to establish critical aspects of visual acuity and to aid in determining the appropriate methodologies by which to evaluate two separate night-vision goggles (NVGs). The following sections will highlight the basics of visual acuity and how it can be determined, the importance of visual acuity in NVGs, and three psychophysical methodologies we will use to determine which of the two NVGs offers the most optimum performance in low-light conditions. The method sections of this proposal will detail the experimental design and methods we propose for the evaluation of both models of Night Vision Goggles (NVGs).

Visual Acuity

Visual acuity is the most commonly measured function when measuring an individual's visual abilities (Johnson & Casson, 1995). As defined by Goldstein & Cacciamani (2022, p. 53), "acuity refers to the ability to see details", which is incredibly important in high-stress, high-stakes environments such as a battlefield. While numerous tools exist to measure visual acuity, the most prevalent tool within a clinical setting is the Snellen chart (Zapparoli et al., 2009). The Snellen chart provides a portable solution to quickly assess visual acuity in both monocular and binocular vision (Tsui & Patel, 2020).

When assessing visual acuity using the Snellen chart (Figure A1), participants tell the observer the smallest letter they can see on which number line (1-11). This response is recorded as one of the fractions listed, with 20/20 being average or standard visual acuity. Any fraction smaller than 20/20 is "worse" visual acuity, while any fraction above 20/20 is considered better than average acuity (Vimont, 2022).

A key component to success on the battlefield is appropriately distinguishing between a target that poses a threat and one that does not—for instance, the ability to distinguish friend from foe or properly distinguish a combatant from a civilian. Visual acuity affects target discrimination and marksmanship (Hatch et al., 2009). Target discrimination becomes more challenging when using NVGs, mainly because of the display characteristics of the goggles (Parush et al., 2011). The image quality produced by NVGs directly affects visual acuity (Crowley, 1991). In fact, the effectiveness of NVGs is often gauged by the level of visual acuity they provide the user (Rash & McLean, 2001). As the illumination level of the night sky decreases, such as during a new moon, NVG-assisted acuity also decreases (Rabin & McLean, 1996). Due to the various factors that can affect visual acuity when wearing NVGs, correctly assessing this measure when testing will allow for safer working conditions and better battlefield performance.

Johnson Criteria

Johnson's Criteria, developed by John Johnson in 1958, is a metric used to calculate and predict the probability of target discrimination for objects imaged by an optical system (Sjaardema et al., 2015). Fundamentally, Johnson proposed that an observer's ability to detect, determine orientation, recognize targets, and identify targets when using an optical device depends on how well he or she can resolve bar patterns. These bar patterns vary in frequency at the same contrast for image intensifiers and television or the same temperature differences for

FLIR (Ratches, 1999). Johnson's Criteria provides metrics for night vision intensifier sensors and focuses on four areas: detection, orientation, recognition, and identification (DesAutels, 2022). For this reason, although there are various new factors and differing technologies from the time of development, Johnson's Criteria was a breakthrough in how visual devices were evaluated and served as a guide for the development of future technologies. Furthermore, it is still a commonly used method for evaluating visual systems. The proposed experiment considers this by following the framework and collecting data for detection, recognition, and identification to evaluate both night vision goggles comprehensively. This will be done by using the Johnson Criteria as a framework for what types of measurements to collect across our tests. The original detection aspect of this criteria will be used with the USAF 1951 Resolution Target, as the Johnson Criteria is commonly used with this target (Space and Naval Warfare Systems Center Atlantic, 2013). However, we will not be calculating any resolutions[1]. We have created a scoring system that will be discussed later. The USAF 1951 Resolution Target will be discussed in more detail in the following section.

USAF 1951 Resolution Target

The USAF 1951 Resolution Target is a chart developed to test UAVs, night vision goggles, and other imaging systems consisting of triplets of lines on visual acuity (Messina & Evans, 2006). This chart is made using a three-bar test pattern, which is a standard assessment method in both government and industry work (Mericsko & Collier, 1970) and is still commonly used for resolution testing for imaging optical equipment (Gonçalves & Griffith, 2006; Orych, 2015). This chart has been used as a basis for the development of numerous bar target field tests (Orych, 2015). We have chosen to use this chart because it is easy to use, widely used across domains, and is a standardized, reliable, and valid chart.

Psychophysical Methodologies

The method of limits presents participants with stimuli in a sequence that either increases or decreases the size of the stimuli (Goldstein & Cacciamani, 2022). In the case of the Snellen chart, letters are presented line by line, with letter size decreasing as you move down the chart. The results of the Snellen chart test provide a quick way to determine a participant's visual acuity. The USAF 1951 Resolution Target primarily uses physical tasks and judgments as a methodology while also incorporating some aspects of the method of constant stimuli. One of the primary advantages of using physical tasks and judgments is the practical relevance of assessing performance, the adaptability inherent in this methodology, the control the experimenter has during the task, and the objective measure for both detections and discriminations (Pelli & Farell, 1995). Incorporating the method of constant stimuli allows for extreme precision in measurements, which is beneficial when evaluating NVGs (Goldstein & Cacciamani, 2022). The forced choice methodology will be used in the final task of our evaluation. Using forced choice will aid in reducing participants' bias and guessing, allow for objective measurements, and also

[1] Note for Alex: I was thinking about explaining some methodology on this, but I do not have access to most articles. I would like to talk to you about it (it seems interesting).

reduce the cognitive load on the participants during the task (Bogacz et al., 2006; Jenadeleh et al., 2023).

Research Objectives

The research objectives of this proposal are to assess the clarity, sharpness, and overall performance of both NVGs in various environmental and distance conditions and compare the findings to determine which NVGs should be utilized for military operations.

Methods

Participants

This study will utilize active-duty ground unit military personnel aged 18 to 45 with normal or corrected 20/20 vision. This specific sample was chosen to obtain a representative sample of the intended users for the NVGs (U.S. Naval Institute, 2023). We aim to obtain adequate statistical power to obtain 80 participants (40 for each NVG model).

Equipment

The equipment required for this experiment includes two NVG models (Model A and Model B), a Snellen Chart (Figure A1), a USAF 1951 Resolution Target (Figure A2), several white boards to present stimuli, and a room of at least 110 meters in length with the ability to manipulate lighting for low light ($1.38\text{x}10^{-2}$ lux) and no light ($8.61\text{x}10^{-4}$ lux) conditions (Pinkus et al., 1998). This room should also be capable of utilizing equipment to simulate the following environmental conditions: moderate smoke will be simulated using a smoke machine, and moderate rain will be simulated by using a ceiling sprinkler system.

Procedure

The experiment contains three tests. The first test will be a standardized Snellen chart (Figure A1) to assess baseline visual acuity using the limits method. The second test will use the USAF 1951 Resolution Target (Figure A2) and utilizes the psychophysical method of physical tasks and judgments, but also contains elements for the method of constant stimuli (Fechner, 1860; Space and Naval Warfare Systems Center Atlantic, 2013). Each participant will complete 36 trials (two trials per condition). The third test utilizes a psychophysical method of forced-choice through a whiteboard and image setup, and each participant will complete 36 trials (two trials per condition). In addition, we will administer a post-study survey to gather subjective information about the NVG, including comfort, ease of use, and perceived performance.

The experiment will be evaluated using Johnson's Criteria, which will test for detection, recognition, and identification. Test two will be for detection, while test three will cover recognition and identification. For tests two and three, each participant will undergo four environmental conditions and three distance conditions (See Table B1). Furthermore, the rain and smoke conditions will always be completed in low lighting. To avoid order effects, conditions will be presented randomly. To avoid fatigue, 10-minute breaks will be provided between each condition. For example, participant A will be tested on the low light condition at all three distance conditions for two trials. There will then be a 10-minute break. They will then complete the following condition at all distance conditions for two trials. This will be repeated until all condition combinations are completed. Baseline visual acuity measures with and without goggles will be obtained for each participant before tests two and three. These measures will be collected using a standardized Snellen Chart.

Baseline Visual Acuity Testing

This test will be conducted using the Snellen eye chart to evaluate whether participants have normal or corrected vision with and without their assigned NVG. Participants will be in a low-light condition while wearing the NVGs. Any participants who score below the 20/20 threshold in either condition will be excluded from further analysis.

USAF 1951 Resolution Target

Each participant will complete a resolution chart test using the USAF 1951 Resolution Target for this test. We will determine the smallest set of lines/pattern each participant can detect. Each participant will complete this test with their assigned NVG and for the previously mentioned environmental and distance conditions. Participants will indicate whether or not they can detect a specific set of lines or patterns. The smallest detectable set will be recorded as their threshold for that trial. Each set will be assigned a point value, with that value increasing as the target set decreases in size (see Table B2 for scoring breakdown). The total number of points will be calculated for each participant, and the total scores will be analyzed using a 2 (goggle type) X 4 (environmental condition) X 3 (distance) mixed factorial ANOVA.

Forced Choice Task

This test will involve participants in a forced-choice scenario. Test 1 will include either an enemy figure holding an M4 Carbine on one side of a whiteboard or a friendly figure holding a medical kit, while the other side of the whiteboard will be blank, as seen in Figure A3 of the Appendix. The participant must determine which side of the whiteboard the figure is on, recognize whether it is a friend or foe, and identify what they are holding. Each participant will do this with their assigned NVG and for the previously mentioned environmental and distance conditions.

Before the evaluation test, participants will be given a brief familiarity trial where images will show either a friend or foe holding one of the following types of equipment (med kit, M4 Carbine, and a tactical shovel). Each unique combination will be shown once prior to the experimental trial. To evaluate their performance, participants will be scored on the following criteria, with higher scores representing better overall performance on this task. First, the participant will gain one point if they can correctly identify which side of the whiteboard the friend or foe figure is located. If they misidentify, then they will receive zero points for this subtask. Second, the participants will gain one point if they can correctly recognize if the target represents a friend or a foe based on images presented to them before the trial begins. They will receive zero points for this subtask if they fail to recognize the target accurately. Third, participants will receive one point if they can correctly identify the equipment the target is holding. They will receive zero points if they cannot correctly identify the target's equipment.

Scores on this task will be initially scored individually and then summed together to form a composite score. Composite scores will be compared between the two goggles through a 2 (goggle type) X 4 (environmental condition) X 3 (distance) mixed factorial design ANOVA to determine overall performance. In addition to the composite analysis, performance on the individual sub-tasks will be compared using a 2 (goggle type) X 4 (environmental condition) X 3 (distance) mixed factorial design MANOVA to look at individual subtask performances.

Subjective Measures of the NVGs

Perceived performance, comfort, and ease of use of the two NVGs will be evaluated. This will be done through post-test surveys after the experiment is completed, in which the NVG will be rated through seven-point Likert scales on the perceived comfort, how easy the goggles were to use, and how well the soldiers felt they performed while wearing them. The NVG will be given a score from one to seven for each of the three subjective measures, with a score of one representing extremely bad comfort/ease of use/performance and a score of seven representing extremely good comfort/ease of use/performance. Scores for each of the factors from both NVG will be compared using independent t-tests.

Conclusion

This proposal reviews the existing literature on the tests and methods utilized in this experiment and details the experimental design and methods we would employ to aid in the decision-making process for investing in NVGs. Both objective and subjective data will be collected for both NVG models. Objective performance measures will be provided through our tests, while subjective measures of perceived performance, comfort, and ease of use will be gathered through a post-study survey. The data analysis will determine which NVG is more suitable for investment.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700-765. https://doi.org/10.1037/0033-295X.113.4.700

Borissova, D. (2003). Methods for NVG visual acuity determination. *Cybernetics and Information Technologies*, *3*(2), 25-33. https://cit.iict.bas.bg/CIT_03/v3-2/25-33.pdf

Crowley, J. S. (1991). *Human factors of night vision devices: Anecdotes from the field concerning visual illusions and other effects* [Technical Report]. U.S. Army Aeromedical Research Laboratory. https://apps.dtic.mil/sti/tr/pdf/ADA237641.pdf

DesAutels, G. L. (2022). A modern review of the Johnson image resolution criterion. *Optik*, *249*, 168246.

Dyer, J. L., Young, K. M., Watson, S. A., & McClure, N. R. (1996). *Night vision goggle field-expedient visual acuity adjustment procedures: Initial experiment* [Research report 1692]. U.S. Army Research Institute for Behavioral and Social Sciences. https://apps.dtic.mil/sti/tr/pdf/ADA310099.pdf

Fechner, G. T. (1860). Elements of Psychophysics (Vol. 2). Breitkopf u. Härtel.

Goldstein, E. B., & Cacciamani, L. (2022). *Sensation & Perception* (11th ed.). Cengage.

Gonçalves, D. P., & Griffith, D. J. (2006). Estimating uncertainty in resolution tests. *Optical Engineering*, *45*(5), 053601-053601.

Hatch, B. C., Hilber, D. J., Elledge, J. B., Stout, J. W., & Lee, R. B. (2009). The effects of visual acuity on target discrimination and shooting performance. *Journal of the American Academy of Optometry*, *86*(12), e1359-e1367. https://doi.org/10.1097/OPX.0b013e3181be9740

Jenadeleh, M., Zagermann, J., Reiterer, H., Ulf-Dietrich, R., Hamzaoui, R., & Saupe, D. (2023, June 201-22). *Relaxed forced choice improves performance of visual quality assessment methods* [Paper presentation]. 15th Annual International Conference on Quality of Multimedia Experience, Ghent, Belgium. https://doi.org/10.48550/arxiv.2305.00220

Johnson, C. A., & Casson, E. J. (1995). Effects of luminance, contrast, and blur on visual acuity. *Optometry and Vision Science*, *72*(12), 864-869. https://www.researchgate.net/profile/Chris-Johnson-11/publication/14453909

Mericsko, R., & Collier, K. (1970). A Comparison Study of Resolving Power Targets.

Messina, E., & Evans, J. M. (2006). Standards for visual acuity. *ASTM International Task Group E54. 08.01 On Performance Measures for Robots for Urban Search and Rescue*.

Orych, A. (2015). Review of methods for determining the spatial resolution of UAV sensors. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*, 391-395.

Parush, A., Gauthier, M. S., Arseneau, L., & Tang, D. (2011). The human factors of night vision goggles: Perceptual, cognitive, and physical factors. *Reviews of Human Factors and Ergonomics*, *7*(1), 238-279. https://doi.org/10.1177/1557234X11410392

Pelli, D. G., & Farell, B. (1995). Psychophysical methods. In M. Bass, E. W. Van Stryland, D. R. Williams, & W. L Wolfe (Eds.), *Handbook of Optics* (2nd ed. pp. 29.1-29.13). McGraw-Hill. https://citeseerx.ist.psu.edu/document

Pinkus, A., & Task, H. L. (1998). Measuring observers' visual acuity through night vision goggles. *Air Force Research Laboratory*. https://apps.dtic.mil/sti/tr/pdf/ADA430646.pdf

Rabin, J., & McLean, W. (1996). A comparison between phosphors for aviator's night vision imaging system. *Aviation, Space, and Environmental Medicine*, *67*(5), 429-433. https://europepmc.org/article/med/8725468

Rash, C. E., & McLean, W. E. (2001). Optical performance. In C. E. Rash (Ed.), *Helmet mounted displays: Design issues for rotary-wing aircraft* (pp. 93-152). SPIE Press. https://doi.org/10.1117/3.397108.ch5

Ratches, J. A. (1999). *SPIE Proceedings*, *3701*, 2–12. https://doi.org/10.1117/12.352982

Sjaardema, T. A., Smith, C. S., & Birch, G. C. (2015). *History and Evolution of the Johnson Criteria* [Technical report]. Sandia National Lab. https://www.osti.gov/servlets/purl/1222446

Space and Naval Warfare Systems Center Atlantic. (2013, October). *Night vision technologies handbook*. U.S. Department of Homeland Security. https://www.dhs.gov/sites/default/files/publications/NV-Tech-HB_1013-508.pdf

Tsui, E., & Patel, P. (2020). Calculated decisions: Visual acuity testing (Snellen chart). *Emergency Medicine Practice*, *22*(4). https://pubmed.ncbi.nlm.nih.gov/32259420/

U.S. Naval Institute. (2023, November 29). 2022 Demographics Profile of the Military Community. https://s3.documentcloud.org/documents/24177791/2022-demographics-report.pdf

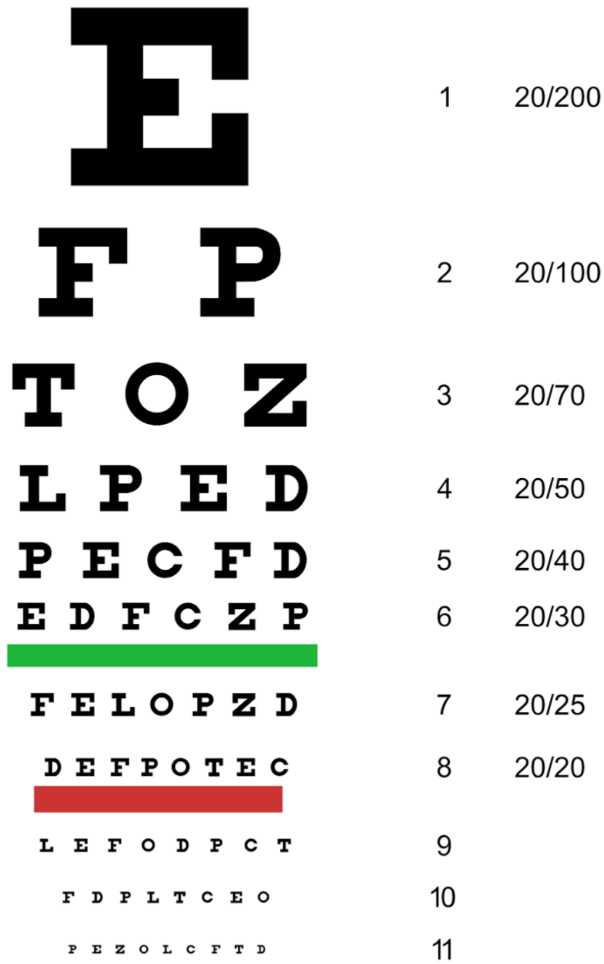
Vimot, C. (2022, January 28). *What does 20/20 vision mean?* American Academy of Ophthalmology. https://www.aao.org/eye-health/tips-prevention/what-does-20-20-vision-mean

Zapparoli, M., Klein, F., & Moreira, H. (2009). Snellen visual acuity evaluation. *Arquivos Brasileiros de Oftalmologia*, *72*(6), 783-788. https://doi.org/10.1590/s0004-27492009000600008

Appendix A

Figure 1

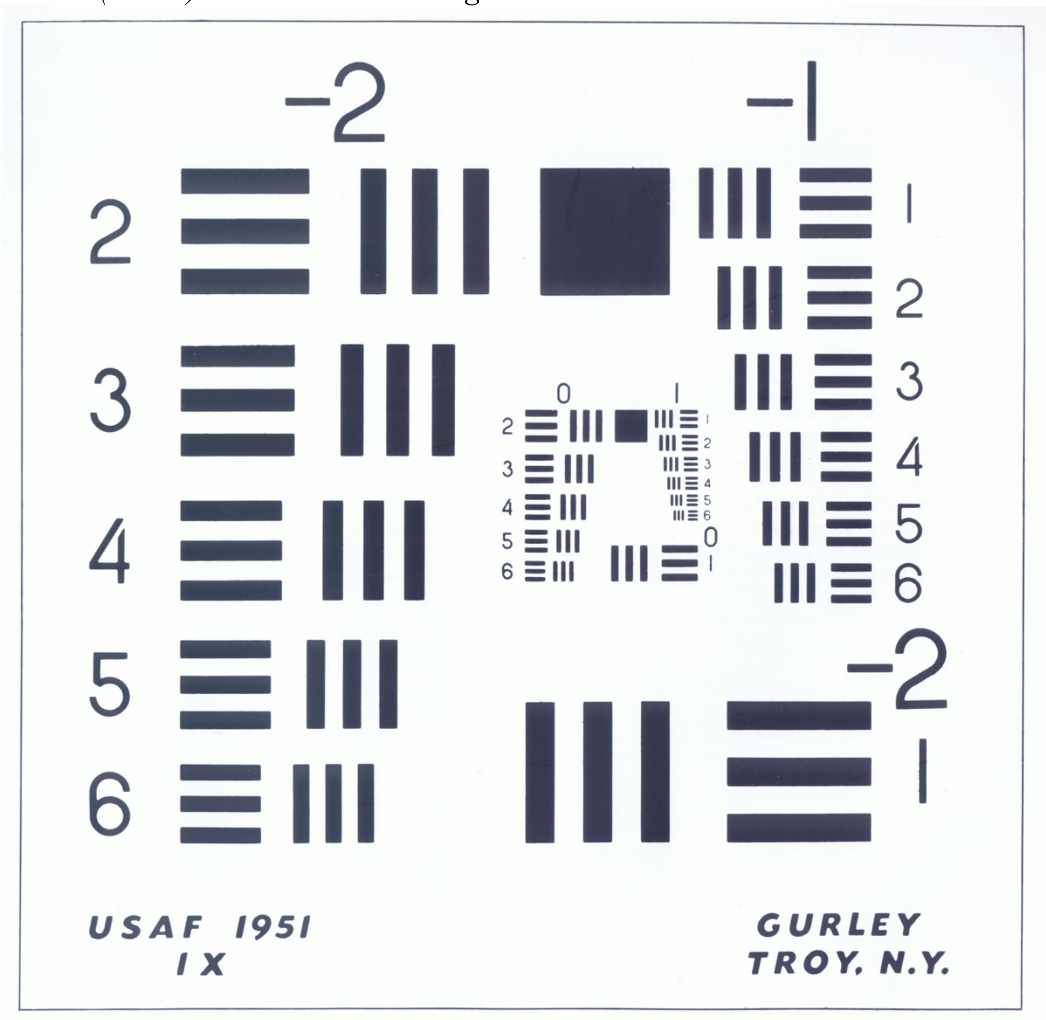
Snellen Chart



Example of a Snellen chart used to measure visual acuity. Participants that cannot correctly read the letters presented at the 20/20 line both with and without the NVGs will be excluded from further analysis.

Figure 2

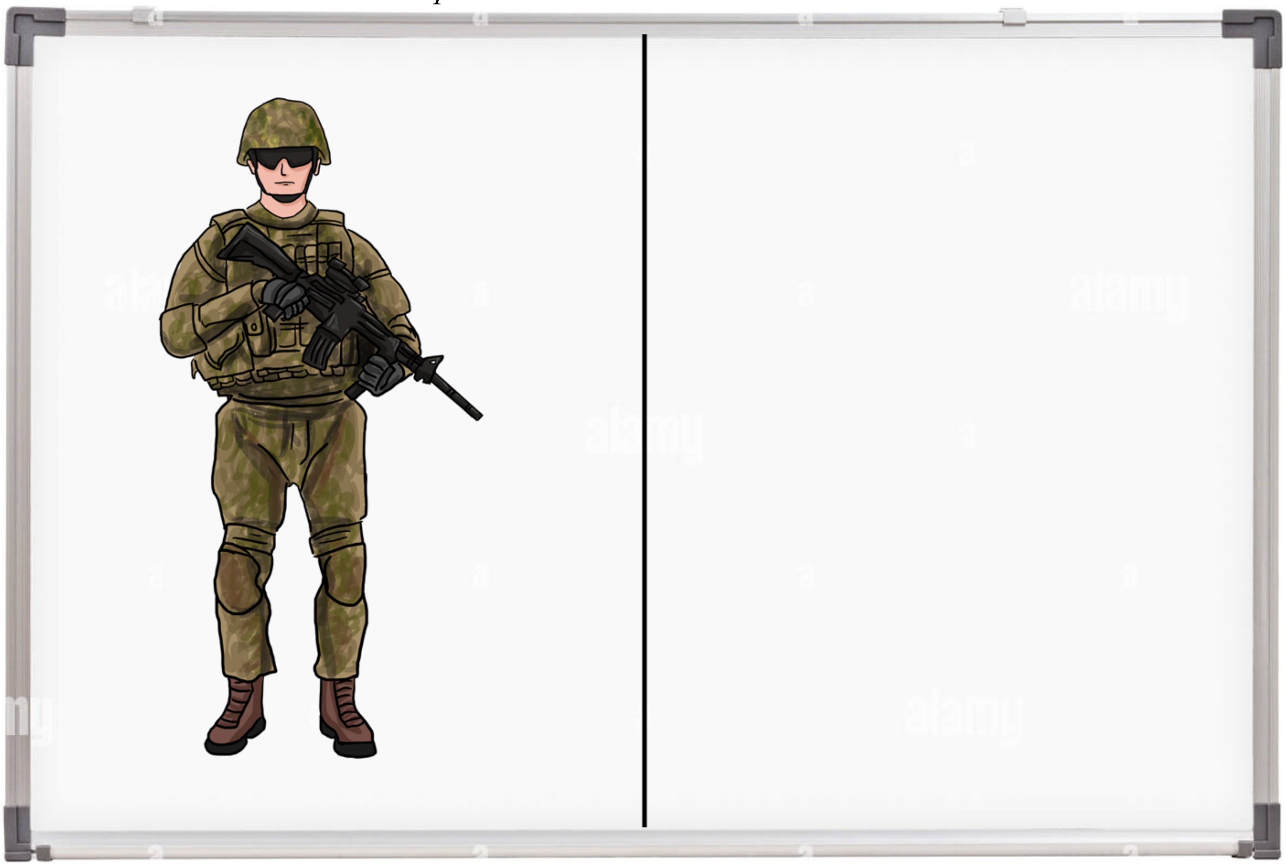
U.S. Air Force (USAF) 1951 Resolution Target



This target has assigned Groups and Elements used to measure visual performance. Groups are represented by the large numbers above the sets of lines (-2, -1. 0, and 1) to represent the size of the lines (Group -2 being the largest lines and Group 1 being the smallest lines). Elements are nestled inside of these groups, represented by the large numbers to the sides of the lines (1-6), and represent how thickly the lines are presented (Element 1 having the thickest lines and Element 6 having the thinnest lines). The smaller the sets are, the more points they are worth.

Figure 3

Forced Choice Test: Friend Example



Three points could be awarded in this task if the participant correctly identified the soldier on the left side of the whiteboard, correctly identified the soldier as a friend, and correctly identified that the soldier was holding an M4 Carbine.

Appendix B

Table 1
Condition Breakdown

Goggle Model	Environmental Condition	Distance (meters)
Model A	Rain	10
		50
		100
	Smoke	10
		50
		100
	Low Light	10
		50
		100
	No Light	10
		50
		100
Model B	Rain	10
		50
		100
	Smoke	10
		50
		100
	Low Light	10
		50
		100
	No Light	10
		50
		100

Visual representation of the experimental design layout with all conditions listed.

Table 2
USAF 1951 Resolution Target Scoring Breakdown

Group	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6
-2	1 point	2 points	3 points	4 points	5 points	6 points
-1	7 points	8 points	9 points	10 points	11 points	12 points
0	13 points	14 points	15 points	16 points	17 points	18 points
1	19 points	20 points	21 points	22 points	23 points	24 points

This table shows the breakdown of how points will be awarded for the USAF 1951 detection task.