Chapter 3 – LABORATORY SET-UP AND RESEARCH

Gregory Craig, Todd Macuda, Sion Jennings
National Research Council of Canada
Ottawa, Ontario
CANADA

greg.craig@nrc-cnrc.gc.ca

3.1 LABORATORY TESTING

Flight testing, while more operationally valid, is typically more expensive to conduct than laboratory experiments. Typical balanced[1] experimental designs can require anywhere from 6 – 50 hours (or more) to conduct. Psychophysical testing is usually an efficient way of evaluating the operators/viewers response to various types of NVG imagery. The tests may involve simple evaluations of pilot visual ability through the goggles or performance of simulated flight tasks with NVG imagery or symbology. These tests are useful in answering specific functional questions about the NVGs under controlled conditions. Often these tests will be useful in determining what sorts of flight tests might be conducted and/or what future laboratory tests may be needed. Alternately, psychophysical tests may be used to further examine some phenomena observed during the flight tests. The following aspects of the psychophysical examination of NVGs are discussed:

- The requirement to develop facilities appropriate for NVGs, which will include well controlled lighting conditions and standard light sources;
- Scientific method or experimental method to evaluate equipment based on optometry, psychophysics, or cognitive research paradigms; and
- Simulator or flight test facilities to assess real operational function.

3.1.1 NVG Laboratory: Dark Room Facilities and Equipment

To conduct precise psychophysical testing, careful control of the testing environment is often required[2]. For NVG testing this entails careful control of the illumination in the test area so that a variety of outdoor night illumination conditions is attainable. By doing so, the performance of the goggles in the laboratory can readily be compared to, or used to predict, goggle performance in the aircraft.

The first step in controlling the lighting in a given area is to eliminate light leaks from outside of the room. This may require plugging small holes and the installation of curtains around doors and windows. Care is needed in choosing material for the curtains as most fabrics show up as “white” in NVGs, regardless of the actual colour of the material. Dark coloured materials such as leather, PVC, latex or similar materials will typically show up as dark under low ambient illumination. The next step is to reduce the amount of reflected light in the room, typically by painting the room with a flat black paint. The purpose of reducing the reflection is to precisely control the overall illumination level from the light source(s) and reflective surfaces will typically increase the ambient illumination level. While any dimmable light source could be used to test the NVGs, a light source which has similar radiant characteristics to those available at night is preferable. Typically the light source radiance characteristics will be matched to those of the moon and the stars, but may also be matched to urban night time lighting radiance characteristics. Fine control of the

[1] In a balanced experimental design all combinations of experimental conditions are tested and compared.

[2] There are a few instances where the ambient lighting, so long as it is consistent between test points, might not influence the experiment (e.g. direct comparison of the performance of different image intensifier tubes), but most of the time the lighting needs to be controlled so that the results of the experiment are relevant to conditions outside of the lab.

level of illumination is also desirable to replicate the wide range of illumination levels available at night. The radiance characteristics provided by the light source will need to be measured usually at the start and end of an experiment because radiant characteristics of the light source will change over time. The illumination level will also need to be measured, particularly if the experiment requires specific levels of illumination (e.g. quarter-moon light). To measure light levels several photometric and radiometric devices are available. However, the unique challenge of NVG research is to characterise light levels in terms of Night Vision Imaging System (NVIS) values. A good standard technology and corporate supplier of equipment capable in this measurement regard is the Hoffman Engineering Corporation. The authors recommend purchasing the Hoffman ANV 410 A and B, for measuring light levels in terms of NVIS and visible light levels, respectively.

The general recommendation of the authors of this paper is to enlist the support of a physicist knowledgeable about NVG function when developing light sources for laboratory use. When selecting light sources for your laboratory, it is important to balance cost with scientific rigour. In many cases funding infrastructure and limitations contribute to a need to use more cost efficient light sources. NRC has opted to use 2856K colour temperature Halogen lamps to be consistent with nominal moonlight and starlight conditions. The design of these systems is based on an original lighting system developed by Alan Pinkus and his research group at Wright Patterson Air Force Base. The main premise of these light sources is that they are a point light source using an aperture to denote illumination intensity without changing colour temperature (RCA Electro-Optics Handbook, 1974). Such a light source provides even lighting to NVGs, preventing variations in the NVG field of view. It can also be used to emulate light levels from moonlight to cloudy starlight. A second alternative approach towards lighting would be to select uniform light sources. These are somewhat more expensive than a point light source, but will provide even better controlled lighting conditions. The Australians have opted for this latter more expensive alternative (Hughes, 2001).

3.1.2 NVG Laboratory: General Experimental Methodology

The basic premise behind any psychophysical experiment is to determine how behaviour (human reaction) is influenced by a prescribed set of condition. In most cases vision scientists conducting NVG work will have a strong knowledge of psychophysical procedures. However, in some cases human factors engineers, while versed in the scientific method and experimental psychology, will be lacking in domain expertise in this area. The authors suggest that in these cases a good working knowledge of psychophysics and human performance measures can be developed through consulting Gescheider (1997) and the AIAA/ANSI Guide to Human Performance measures as an initial starting point. Further, it should emphasized that all NVG and HMD research groups consult the expertise of a trained psychophysicist when conducting laboratory and flight based psychophysical experiments. For NVG testing, the goal is to determine how the subject's visual perception (and the resultant behaviour) is affected by the goggles. Four general research paradigms can be applied directly to laboratory-based NVG research including:

- Optometry;
- Basic psychophysical methods;
- Cognitive research techniques; and
- Simulation and modelling.

These four approaches are usually blended when conducting NVG laboratory investigations. Frequently, NVG studies have utilized basic optometric assessment techniques such as acuity, contrast, and depth perception to characterise NVG performance[3]. Acuity testing can be used to determine the distance at which an object of a given size can be detected (but not necessarily identified). Performing the acuity tests

[3] Colour vision tests, astigmatism, perimetery and other basic vision tests are likely be of limited use in NVG testing.

at different illumination levels will identify how the performance of the goggles may decrease on flights conducted during different moon phases and during different weather conditions (overcast vs. clear). The limitation of basic acuity tests is that only high contrast stimuli are used. In flight, the pilot is often interpreting and attempting to identify lower contrast targets. As such, contrast testing is particularly relevant to NVG flight in determining which objects/obstructions may be difficult to discern. Stereo-depth perception tests may also be useful in determining any deficiencies or changes in depth perception due to NVG optics or due to poor set-up of the goggles by the user. One difficulty with most standard optometric tests is that they are designed for basic unaided visual ability at short distances. However, some of the acuity stimuli, such as Landolt rings (Figure 3-1) and grating acuity stimuli (Figure 3-2) are relatively easy to produce and rescale for larger test distances.

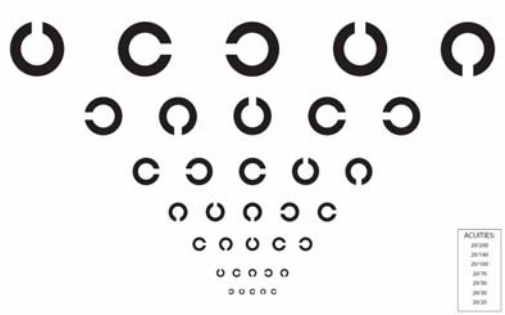

Figure 3-1: Landolt Ring Acuity Test.

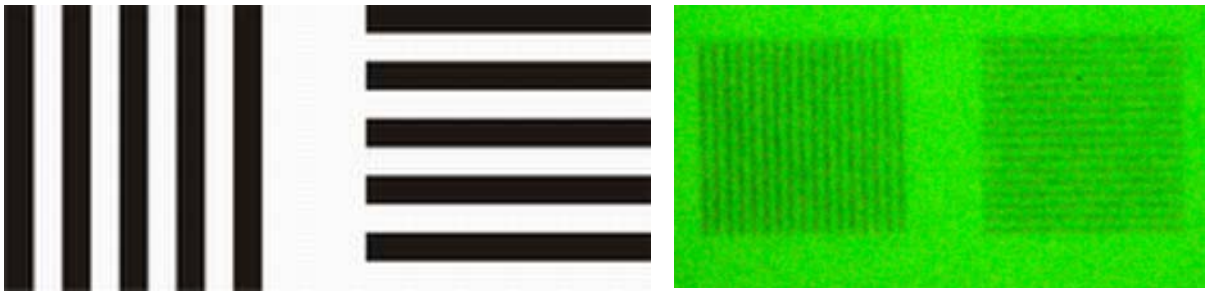

Figure 3-2: Grating Acuity Stimuli (left) and a Sample Seen through NVGs under High Level Illumination (right).

NRC employed some basic optometric tests in conjunction with psychophysical procedures to evaluate different candidate tube technologies for the Canadian Department of National Defence (Macuda, Allison, Thomas, Truong, Tang, Craig & Jennings, 2005). In this test, a comparison was made of subjects' acuity though the different image intensifier technologies at different levels of ambient illumination. The acuity of the subjects using the goggles was assessed via the standard Hoffman ANV-126 and by square wave grating stimuli (see Figure 3-2) that were developed for the test. While these are very basic tests of the NVGs, differences in the performance of the different image intensifier technologies were revealed by both systems. Small differences between the Hoffman ANV-126 and the grating tests were also detected which indicated that the gratings were more sensitive to small changes in acuity than was the Hoffman test[4]. While the optometric tests revealed differences between the tubes in the laboratory, a short flight test was also conducted to gather corroborating test pilot commentary on the performance of the goggles. Although only two pilots assessed the goggles, the salient points of their in-flight commentary mirrored the data that was obtained in the laboratory. This experiment, comprised of both a laboratory and a flight

[4] However, while it is less precise, the Hoffman test is much faster and easier to administer.

test of the NVGs, represents the preferred method of evaluating most cockpit display systems. While much of the precision investigation can be accomplished in the controlled environment of the laboratory, the diversity of conditions existing in real flight may uncover other aspects of the (NVG) system which warrant additional investigation.

Most basic optometric tests are merely simplified and standardized versions of more rigorous psychophysical assessments. Most psychophysical tests involve systematic presentation of a stimulus (e.g. an object that is to be detected) under various conditions such as the level of ambient illumination and/or the rate of motion of the target. Measurements of the subject's response typically consist of determining how accurate the response was or how long it took for the subject to make the response. For example, NRC examined the influence of Night Vision Goggle-produced noise on the perception of motion-defined form using synthetic imagery and standard psychophysical procedures. Synthetic image sequences incorporating synthetic noise were generated using a software model developed by the NRC research group. This model was based on the physical properties of the Aviator Night Vision Imaging System (ANVIS-9) image intensification tube. The image sequences either depicted a target that moved at a different speed than the background, or only depicted the background. For each trial, subjects were shown a pair of image sequences and required to indicate which sequence contained the target stimulus. Subjects were tested at a series of target speeds and at several realistic noise levels resulting from varying simulated illumination. The results showed that subjects had increased difficulty detecting the target with increased noise levels, particularly at slower target speeds. This study suggested that the capacity to detect motion-defined form was degraded at low levels of illumination. The findings were consistent with anecdotal reports of impaired motion perception in NVGs (Macuda et al., 2004, 2005).

There are also assessments of the user's cognitive processes when using the NVG systems. Workload is a general cognitive approach used in most evaluations to determine how easy or how difficult it is to use a given system or how much the pilot needs to compensate for the deficiencies of the NVG system. Workload assessments are discussed in more detail in Chapter 5. Situational awareness is a term for the concept of how and the degree to which the pilots integrate information from themselves, their aircraft and the environment to form an overall representation of their situation and take appropriate action(s) (Endsley, 1993; Endsley, 2000). Assessments of situational awareness may take the form of general questionnaires such as the China Lake Situational Awareness Rating Scale, the Situational Awareness Global Assessment Technique, the Situational Awareness Rating Technique, etc., or indirect measurements of significant aspects of situational awareness. Indirect assessments of situational awareness may gauge the pilot's ability to detect significant information or events, the time it took the pilot to become aware of significant information or how confident the pilot was in their response. The latter tests are typically done in laboratory-based simulation where significant events can easily be programmed to occur during a specified set of circumstances or at specific intervals. Situational awareness questionnaires are often used in both lab and flight settings as they are quick and easy to administer. Applied to NVG testing, the main goal of situational awareness assessments would be to compare the pilot's situational awareness with and without the goggles, with or without the NVG symbology (if applicable), or to determine if there are differences in situation awareness with different types of goggles (e.g. panoramic NVGs vs. normal field of view goggles).

Although many pilots report that NVGs affect situational awareness (Braithwaite et al., 1989, 1998) and spatial navigation few studies have used cognitive methods to address this issue. As an initial approach in this regard, NRC and Carleton University assessed the impact of NVGs on navigation and wayfinding performance and the acquisition of spatial knowledge. Spatial knowledge refers to a mental map or representation of the spatial layout of objects and landmarks in the environment. The methodological approach used two main phases: a learning phase followed by some tests of survey knowledge. Survey knowledge is an image-like representation of the entire geographical area that provides a "bird's eye view" of the environment. In the learning phase, one group of participants was required to navigate a walking maze consisting of 5 rooms with typical household items in each room and locate target objects while

wearing NVGs, while the control group navigated the maze without NVGs. Findings showed that navigation and wayfinding with NVGs appeared to be harder, with longer navigation times and more navigational errors compared to not using NVGs. The impact of NVGs was also evident in spatial memory tasks. Relative direction pointing to targets across rooms and to distracters in the same room was better for those who performed the search without NVGs compared to using NVGs. In a map drawing task, participants in the NVG group were more likely to position objects incorrectly. In addition, NVG users tended to draw maps with an extra room compared to the non NVG group. These results supported the notion that NVGs directly affect spatial navigation and wayfinding performance and the acquisition of spatial knowledge (Craig et al., 2006).

Simulation and modelling are also useful tools in assessing NVG performance under a variety of conditions. Simulation can take many forms ranging from relatively simple mathematical models of image intensifier reaction to stimuli and light levels or full scale motion-base NVG flight simulators. One aspect of NVG simulation that is fairly constant is that often it is the NVG image itself that is simulated. Most simulation displays/screens are not designed to operate at the low light intensity levels that would allow the goggles to be used. One way around this problem is to use a normal CRT or LCD display and put neutral density filters on the NVGs to reduce the amount of light coming into the goggles. The same method can be used to simulate night flight on an aircraft to allow daylight testing of NVG systems. This approach can be particularly useful when attempting to expand the flight envelope of the NVG/aircraft system as it allows the safety pilot to have daylight Good Visual Environment (GVE) visuals while the evaluator flies under NVG imagery. The problem with simulation is that the visuals are different in several ways from what is seen in the goggles during actual night flights. For example, the apparent size and intensity of the halo generated by the NVG in response to an incompatible light source will depend on the intensity and predominant wavelength of the light. In most simulations, the reaction of the NVGs to cultural lighting will be incorrect. While this is relatively unimportant for the effects of street lights and other sorts of urban lighting to be exact, the halos from simulated airport lighting should be accurate. Other aspects of the NVG image such as the noise in the goggles from low ambient illumination, the reaction of the goggles to precipitation, and the inaccurate representation of infra-red content in the image (some of the objects that might appear dark actually appear bright in the NVGs) might also affect the utility of the simulation. However, the greatest advantage of simulation is that the simulation can be changed to be more accurate. One must be careful in planning the type of experiment conducted in a simulation environment to ensure that the simulation itself does not have an undue impact on the results of the study. Finally, it is important to note the operational utility of developing advanced simulated NVG environments. Such simulation and modelling efforts can be used to develop effective NVG trainers. This will allow pilots to maintain perishable skills without the added cost of flight.

In addition to the three blended approaches above, a new methodology for assessing the influence of NVGs on performance is to use aeromedical procedures for monitoring pilot performance. Such assessments include direct measurement of brain activity and other physiological indicators (e.g. eye movements, respiration, blood pressure) in an attempt to link these with behavioural correlates of performance. Until recently, the study of cognitive/perceptual functioning during flight has been limited mainly to simulation or monitoring of behavioural responses. To move towards an augmented cognitive environment in real flight platforms, it is necessary to develop viable airborne test facilities. The importance of developing a cognitive cockpit is that it provides a framework from which to understand the cognitive/neural mechanisms of flight. It will allow an understanding of how display systems and cockpit technologies affect workload and brain function to ultimately reduce the quantity of information presented to pilots. The University of Iowa and NRC are developing a broad research program that will establish airborne neurophysiological and psychophysiological recording facilities capable of monitoring pilot brain activity during flight. A 128 Channel, 128 Electrode, Electroencephalogram (EEG) system (manufactured by Electro-Geodesics Inc.) and related recording equipment are being integrated into fixed and rotary-wing test platforms. The utility of directly measuring pilot state (e.g. brain activity, heart rate) is that it provides an objective measurement of pilot workload and performance decoupled from his/her behavioural

responses. The value of this methodological approach is that it allows NVG evaluators to make inferences about performance and brain activity by directly assessing the human-in-the-loop.

3.1.3 NVG Laboratory: Basic NVG Image Assessments

The quality of the NVG image is a key concern and pilots can readily distinguish a good NVG image from a poor NVG image subjectively. In order to quantify the pilot's subjective assessments of the goggle imagery, a variety of tests can be performed. The fundamental parameters of image quality are assessed in the most basic of tests such as the following:

- Assessment of the signal to noise ratio in at least three light levels;
- Measurement of the Modulation Transfer Function (MTF);
- Checks for the magnitude of imaging defects (white spots, dark spots, shear, gross distortion, non-uniformity); and
- Side effects (halo, coma, veiling glare).

It is important to note that all of these tests should be performed at operationally representative light levels to ensure that the NVG system performs acceptably at each light level. For example, as the gain of the goggles increases, the noise in the goggles will increase, the contrast sensitivity may change, and the ability of the goggles to resolve detail may change.

3.1.4 Measuring Signal to Noise Ratio

The process by which NVGs amplify light, while highly effective, does generate noise in the image. In particular, as the gain of the goggles increases with decreases in the ambient light level, the noise (the number of random transient pixels) in the image increases. Signal to Noise Ratios are typically available from the manufacturer. The authors suggest that experimenters verse themselves with their requisite technological specifications. Alternatively, it is advised to consult a Physicist/Electro-Optics expert to measure these values in situ (see also Lupo, 1972).

3.1.5 Measuring Modulation Transfer Function

The modulation transfer function is the comparison of the contrast of the test stimulus to the contrast of the image on the eyepiece. The test can be done psychophysically by presenting sine-wave gratings (Figure 3-3) of various contrast levels to 3 – 4 observers and determining the contrast threshold level for the goggles. The contrast threshold is the contrast level which the observer can detect 50% of the time. By varying the frequency of the grating presented to the observer, the contrast and resolution of the goggles can be assessed at the same time to determine the "contrast sensitivity function" of the observer with the goggles. The contrast sensitivity function is a plot of the inverse of the contrast threshold by the spatial frequency of the grating. The difficulty with this method is that either a high quality printer or a specialized computer monitor is required to present the gratings without artefacts. There are two alternatives that can provide a somewhat coarser evaluation of the contrast resolving capabilities of the goggles. The first is the contrast ring (Figure 3-4) on the Hoffman Engineering ANV-20/20 can be used to identify the level of contrast visible in the goggles. The subject identifies the segments of contrast that can be seen with the goggles under the normal and low light settings on the ANV-20/20. A good NVG system should be able to resolve all the contrast segments at the normal lighting setting on the ANV-20/20, and only the first three or four segments under the low-light setting (Hoffman Engineering, 1997).

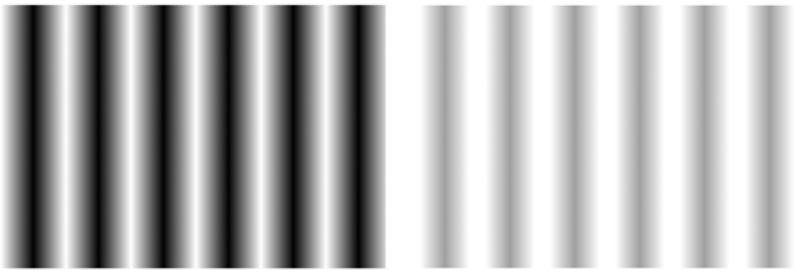

Figure 3-3: Typical Sine Wave Gratings for Testing Contrast.

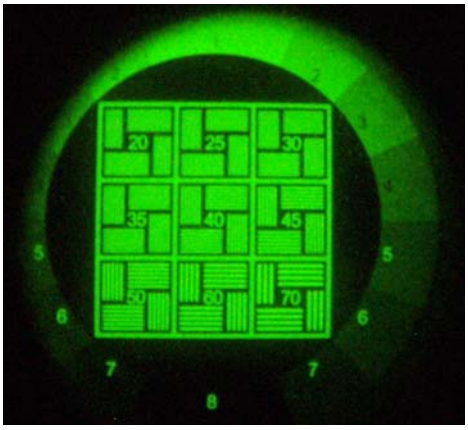

Figure 3-4: Hoffman ANV-20/20 Contrast Ring.

While the ANV-20/20 test provides a quick and easy way to check the level of contrast visible in the goggle, it may not be sensitive enough to identify lower contrast thresholds (i.e. smaller differences in contrast). An alternate test can be done by manufacturing Landolt rings of varying levels of contrast (see Figure 3-1, Figure 3-5). The Landolt rings follow a simple 1:5 ratio in that the diameter is five times the gap in the ring and five times the width of the ring. One simply needs to use different gray levels to draw the rings with different levels of contrast. The typical contrast value, also known as the Michaelson contrast ratio, is derived from the following equation:

$$\text{Contrast ratio} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

Figure 3-5: Low Contrast Landolt Ring Charts.

In the case of the Landolt ring, L_{max} (maximum luminance) is the luminance of the background and L_{min} (minimum luminance) is the luminance of the ring. Measurement of the contrast needs to be done on the printed copy of the stimuli since the contrast on the monitor and the contrast on paper are invariably different. The contrast measurement can be done with a basic photometer. The contrast can be measured directly from the chart itself[5] or through the NVGs. While measuring the luminance values from the chart through the NVGs is not the ideal method photometrically, it is appropriate in two senses. First, most visible-range photometers can be fitted with a filter designed to replicate the filtering of the human eye. The photometer then "sees" the luminance in the goggles in the same way that the pilot would, which is the primary concern when performing the contrast measurement tests. Second, certain types of printers or laminate coatings that might be used on the chart can affect the contrast of the stimuli as seen through the NVGs.

3.1.6 Checks for Imaging Defects

There are a variety of imaging defects that may appear as the result of NVG malfunctions. These defects may include the following (See Appendix 1 for more detailed descriptions of these defects):

- Shading (Dark edges along perimeter of the image);
- Edge Glow (Bright area along outer edge of the image);
- Bright Spots (Constant or intermittent spots in the image);
- Flashes/Flickering;
- Honeycombing (Dark honeycomb pattern over the NVG image);
- Distortion (Straight lines appear bent);
- Veiling Glare (Bright haze or bright artefacts in the NVG image); and
- Dark Spots (Constant dark spots in the NVG).

These defects, the diagnostics for each defect and the course of action are described in detail in the Hoffman Engineering ANV-202/20 Manual (Hoffman Engineering, 1997). Most of the defects listed above may require NVG maintenance action, if the problem(s) are of sufficient magnitude. Taking dark spots as an example, the loss of 1 – 2 pixels in the goggle phosphor in the middle of the field of view may not warrant goggle maintenance while a the loss of a 5 – 6° area would likely require maintenance.

[5] The stimuli on the chart may be too small to accurately capture the luminance value. Test patches of the gray levels used for the stimuli may be required in order to conduct contrast ratio measurements.