AD_____

Award Number: W81XWH-10-1-0736

TITLE: Objective Methods to Test Visual Dysfunction in the Presence of Cognitive Impairment

PRINCIPAL INVESTIGATOR: Randy Kardon, M.D., Ph.D.

CONTRACTING ORGANIZATION: University of Iowa
Iowa City, IA 52242-1316

REPORT DATE: October 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

1. REPORT DATE *(DD-MM-YYYY)* October 2013	2. REPORT TYPE Annual	3. DATES COVERED *(From - To)* 15 September 2012 - 14 September 2013
4. TITLE AND SUBTITLE Objective Methods to Test Visual Dysfunction in the Presence of Cognitive Impairment		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER W81XWH-10-1-0736
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Randy Kardon, M.D., Ph.D. Pieter Poolman, Ph.D. Email: randy-kardon@uiowa.edu		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS University of Iowa, The 105 Jessup Hall Iowa City, IA 52242-1316		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT (200 Words) Summary of most significant finding (s)
Purpose: to develop and validate objective tests to diagnose vision deficits in patients with cognitive impairment and ensure effective monitoring of their treatment.
Scope: objective methods to monitor visual function include 1) the pupil light reflex (PLR), 2) light evoked potentials (VEP) from the brain and from the eye and 3) purposeful eye movements to track moving targets that are resolved.
Major Findings (year 3): i) development of an objective test protocol and data analysis routines to derive metrics that correlate with clinical diagnoses of afferent and efferent pathway pathologies based on the PLR. This protocol is less than 22 seconds in duration. The Pearson correlation coefficient between the derived continuous RAPD metric and clinical RAPD measurements is 0.86; ii) implementation and testing of a user interface to quickly and efficiently detect and remove artifacts from PLR data; iii) comparison of the design matrix method for VEP stimuli with the maximum length sequences (m-sequences) method based on simulations and various statistical metrics.
Significance: objective tests of vision will greatly improve eye care by providing faster, lower cost testing that can be performed in remote settings and will provide a new tool for assessing innovative treatments being developed to save or restore vision.

15. SUBJECT TERMS
Traumatic brain injury, cognitive dysfunction, pupil light reflex, eye movements, evoked potentials, visual function

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	30	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

INTRODUCTION

Our project's research goal is to provide an objective and military relevant means for diagnosing and localizing the site of visual dysfunction in cognitively impaired patients. The successful attainment of this goal hinges on the development of a suite of objective tests designed to free the TBI patients from the cognitive demands placed on them during standard visual testing. As a necessary step for testing of cognitively impaired patients with suspected visual dysfunction, we will first optimize three objective tests of visual function: a) pupil contractions to light stimuli, b) evoked potentials elicited from the visual cortex in response to visual stimuli in the central and peripheral visual field locations, and c) eye position correlated with moving visual targets varying in spatial properties as a means of verifying that the patient was able to perceive the targets. Next, we will validate these objective tests against gold standard behavioral tests of visual field sensitivity in cognitively intact participants who are capable of performing these tests accurately. For the validation phase, participants will be selected who have either normal visual function or who have known dysfunction at different sites along their visual pathway. Using this strategy, normal eyes and eyes with well-defined damage to the retina, optic nerve, visual radiations or visual cortex will be used to study the sensitivity and specificity of the different modalities of objective testing being evaluated in this proposal. Once validated in these participants, these objective tests can then be rapidly implemented for use in cognitively impaired patients, specifically those who have suffered traumatic brain injury.

BODY

RESEARCH ACCOMPLISHMENTS ASSOCIATED WITH APPROVED STATEMENT OF WORK FOR YEAR 3:

Task 1. Implementation of novel product-ready hardware solutions that allow objective testing of the visual system (months 1-12):

1a. Hand-held portable pupillometer (Neuroptics, Inc.) (months 12-24).

The hand-held portable pupillometer (Neuroptics) which was able to deliver white light stimuli and record pupil movements in response to light stimuli for diagnosing retinal and optic nerve disorders has now been replaced by a small desktop unit that is binocular (delivers stimuli to each eye and collects pupil responses from both eyes simultaneously). This will allow afferent visual abnormalities to be differentiated from efferent disorders of the output pupil pathway. The new unit (Neuroptics Inc, Irvine, CA) also detaches for use at bedside exam. This unit also delivers chromatic light stimuli (blue, red, green or white) so that rod, cone and intrinsic melanopsin retinal ganglion cell activation can be derived for differentiating visual loss from photoreceptor disease from that of optic nerve/anterior visual pathway diseases. We are presently redesigning testing protocols to take advantage of this incorporated feature and to reduce testing time to a minimum (one minute):
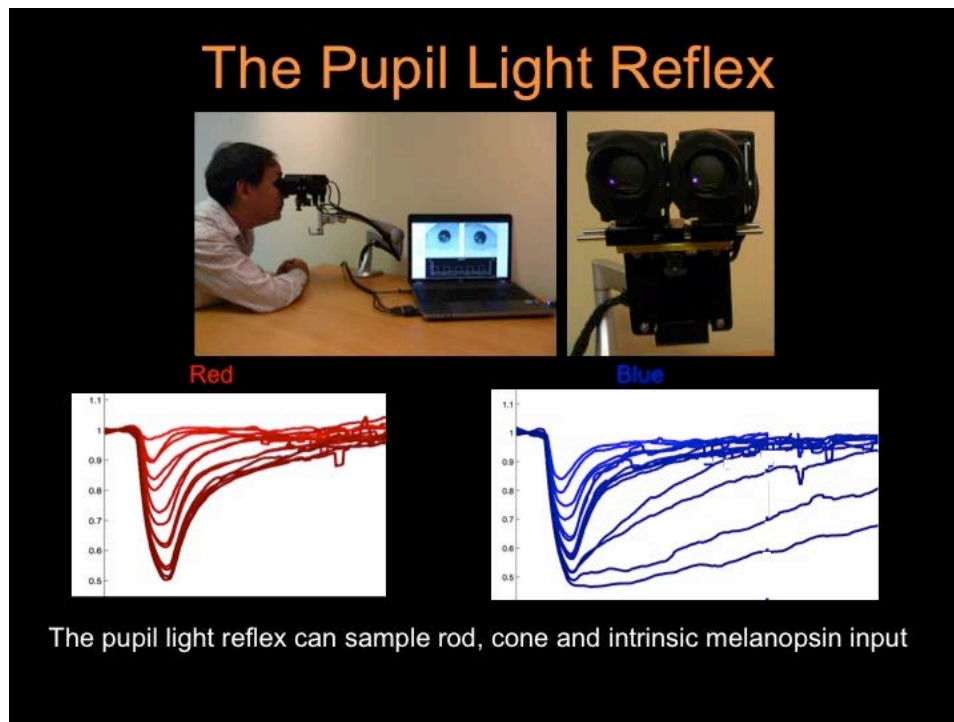

Figure 1. Recent work with binocular chromatic (red and blue light stimuli) showing rod, cone and melanopsin mediated pupil responses. Top is a newly developed desktop unit that can record both pupils and stimulate either the right eye, left eye or both eyes with red, blue, green or white light. At bottom are a series of pupil light

5

reflex waveforms with increasing intensity of red light for cone responses (red tracings) and blue light responses (blue tracings) for rod mediated responses at low intensity and sustained pupil contractions from melanopsin mediated responses to blue light at the hightest intensities.

In the last half of Year 3, we have continued to perform pilot testing with the Neuroptics binocular desktop pupillometer using a number of different testing protocols on patients with and without afferent input disorders to refine the pupil testing protocol that best evaluates the presence of retinal and optic nerve disease. The refined protocol will be incorporated into Neuroptics' new portable binocular pupillometer, scheduled to be available in the last quarter of 2013. We will have the first prototype to evaluate and validate. We have recently performed another pilot study (not part of the DOD protocol or grant) in the Emergency Treatment Center at the University of Iowa Hospital on patients being evaluated in the ER for acute traumatic brain injury (TBI). In this study we researched the predictive power of the dynamics of the pupil light reflex using a hand held portable pupillometer in the trauma bays. We found very encouraging results. The neurological pupil index, a quantitative measure of pupil dynamics was statistically associated with whether the patient required admission, the presence of an abnormal head CT scan, and whether the patient needed a neurosurgical procedure. So this dynamic measure of pupil response appears to be very useful as an acute biomarker of TBI and we will continue to explore this and incorporate this dynamic measure into our DOD protocol for objective testing of visual function.

In response to our request to develop a more portable unit, Neuroptics has developed an even smaller binocular pupilllometer that is hand held (**Figure 1**).

Figure 2: New portable hand-held binocular pupillometer prototype instrument that was developed by Neuroptics Inc, based on our suggestions. This new pupillometer can deliver

red, blue, green or white light stimuli to either eye and record both pupil responses simultaneously to allow us to assess both afferent and efferent pupil disorders. Our recent work with chromatic red and blue light stimuli showed that we could differentiate rod, cone and melanopsin mediated pupil responses.

The next prototype due in 2013, will be smaller, lighter and will be able to be worn by the subject similar to a pair of virtual reality glasses, as depicted in **Figure 2**.

Figure 3: In the drawings above is the 3rd prototype conception, due in 2013 which will be a 9-oz lightweight pair of glasses that contain the miniature infrared video cameras and the LED stimulus source for red, blue, green, and white stimuli that will be comfortably worn by the subject during testing. The rectangular module shown below it will be worn around the patient's neck and contains the LCD monitor screen for viewing the pupils during the test as well as the self contained battery operated computer processor and data storage device.

1b. Pupil testing protocol and advanced analysis

During the second quarter of Year 3, we set out to develop a test protocol and associated data analysis routines to derive metrics that correlate with clinical diagnoses of afferent and efferent pathway pathologies based on the pupil light reflex. The test protocol consists of white light impulses across a range of intensities (1-400 lux) presented sequentially to each eye individually, while the pupil diameter of both eyes are recorded continuously. In order to minimize subject's testing time fatigue, the protocol is less than 22 seconds in

duration, but during this time period, each eye is stimulated 8 times (which includes the binocular stimulus at the start of the protocol to drive the pupil size down from its dark adaptation level closer to its steady-state size during the test). We have implemented the protocol on the Neuroptics DP2000 system and have started to test both normal subjects and patients in order to build up a database of pupil responses. We intend to adapt the protocol to suit our TV-based test platform.
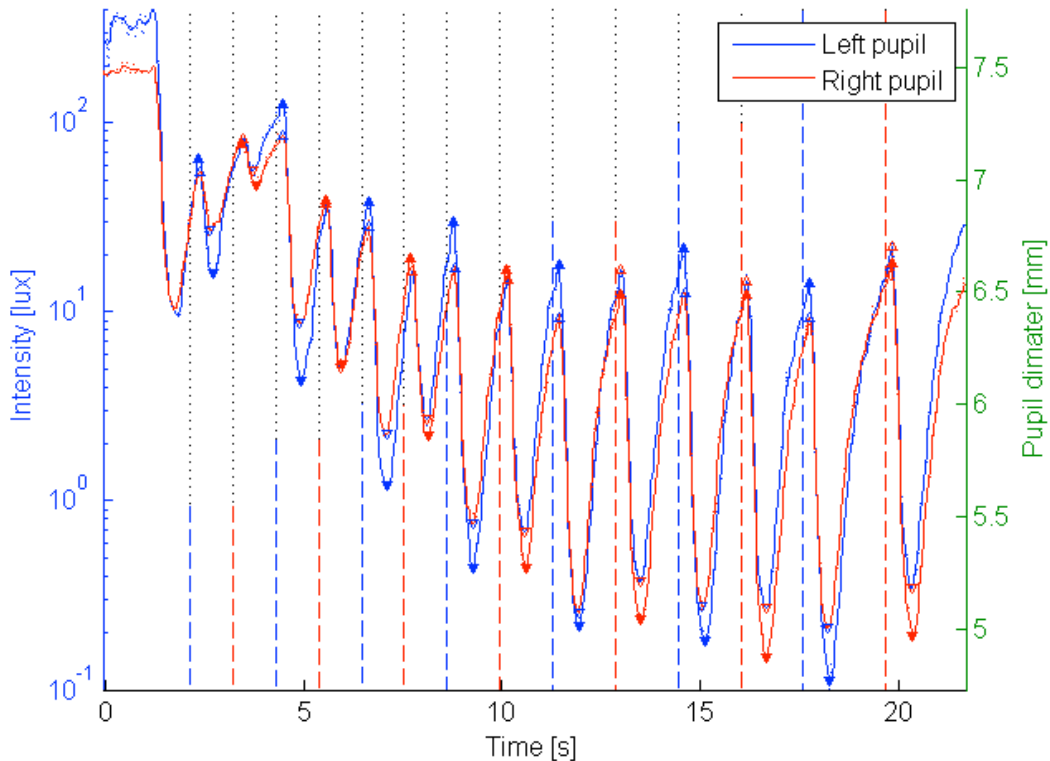

Figure 4. Example data record of filtered pupil size tracings (Left pupil size in solid Blue line, and Right pupil size in solid Red line) and the onset time and intensity of the monocular stimuli (Left eye stimulus in dotted Blue line, and Right eye stimulus in dotted Red line). The intensity of the white light stimuli is indicated in lux on the y axis on the left hand side, and the y axis on the right hand side indicates pupil size in millimeter.

Subsequent to data collection, each subject's pupil data is pre-processed before it can be committed to our database of pupil responses. Pre-processing includes low-pass filtering of each eye's pupil size time series to smooth out high frequency noise components inherent to the frame-by-frame image analysis to find the best-fit circle matching the pupil border in the video frames, as well as locating and estimating the pre-response peak dilation pupil size and the minimum contracted pupil size for both eyes for calculating the maximum contraction amplitude in response to each stimulus. In the DP2000 system, stimulus onset timing is not synchronized to video frame shutter timing, and neither is the shutter timing of the left eye camera synchronized with the shutter timing of the right eye camera. However, the system timestamps every stimulus onset event, as well as each video

frame collected by the left and right eye cameras. To work around the lack of synchronization, the next pre-processing step involves isolating (i.e., cutting out) and resampling of the pupil time series for each stimulus-pupil response waveform based on the timestamp value of each stimulus and cubic spline interpolation. This step allows us to overlay and inspect the pupil response tracings by setting the stimulus onset time to zero for each stimulus-pupil response pairing as shown in Figure 5.
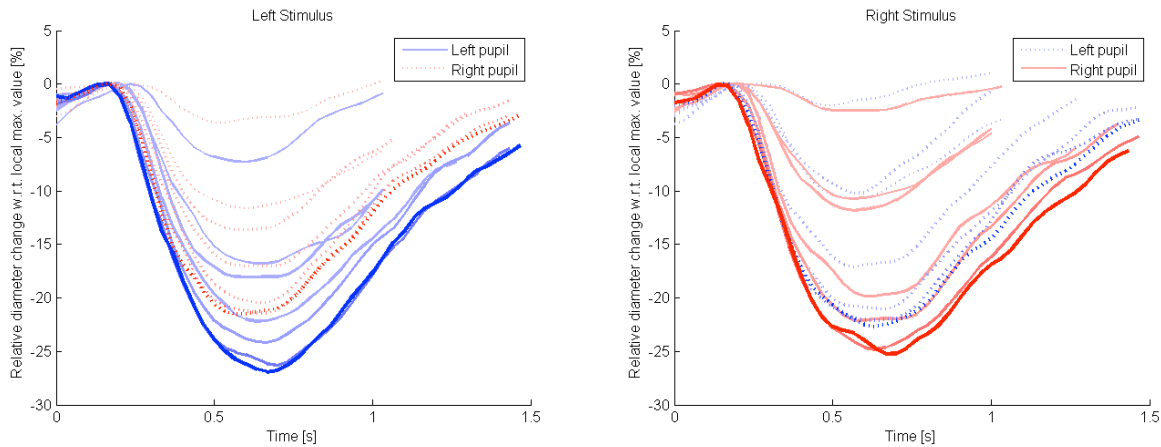

Figure 5. Example left and right pupil tracings for each of the left and right monocular stimuli. Color saturation of the different line series is determined by stimulus intensity, i.e., darkest Blue and Red colors correspond to the brightest stimulus conditions (which, in this case, is ~400 lux).

The short duration of the test protocol means that the pupil size at the start of each stimulus is not the same. Based on results published by Dr. Kardon previously, translation of each pupil response tracing into time series of percentage contraction values helps to alleviate the impact of varying pupil start sizes on subsequent statistical analysis. The next processing step therefore consists of normalizing the values of each stimulus-pupil waveform to represent the percentage change from the pre-response peak dilation pupil size, as calculated in an earlier step, for each individual stimulus. At this point, a subject's pupil data has reached a sufficient level of processing for comparison with data from other age-matched subjects. In future, we will commit a subject's data to a database (e.g., MySQL) at this stage for data warehousing purposes. A well-designed database will support many types of statistical analyses within specialized statistical packages (e.g., R, SPSS, etc.) by leveraging the ability to easily obtain data for analysis purposes via database queries.

In order to compare pupil responses within a subject (i.e., one-to-one), or to compare a subject's (left and/or right) pupil response with a matched population-averaged response for diagnostic purposes (i.e., one-to-many), we have derived a novel computational method to support such data comparisons. The method is based on estimating a field of pupil responses across a range of stimulus intensities through a series of curves fitted to the normalized pupil responses for the discreetly-sample stimulus intensities at pre-determined time offsets from stimulus onset. A pupil response field can be used to describe the responses of a single pupil recorded at a finite number of stimulus intensities,

or be used to summarize the pupil responses from multiple sources (e.g., for all aged-matched normal subjects in a database). The method also enables data comparisons in the face of missing or mismatched stimulus intensities. Missing stimulus intensities would include the case for which a recorded response is rejected due to noise or artifact contamination. Several types of curves can be fitted, each having its own benefits and drawbacks. The available curves include the generalized hyperbolic, power, generalized logistic, and Naka-Rushton functions. Figure 6 shows an example of curves fitted at different time offsets. The method is also capable of computing the best-fitting response within any pupil response field (e.g., from a database) given a test pupil response waveform, and to report the matching stimulus intensity, as shown in Figure 7. We have extended this idea to calculate the continuous RAPD over a range of stimulus intensities as shown in Figure 8.
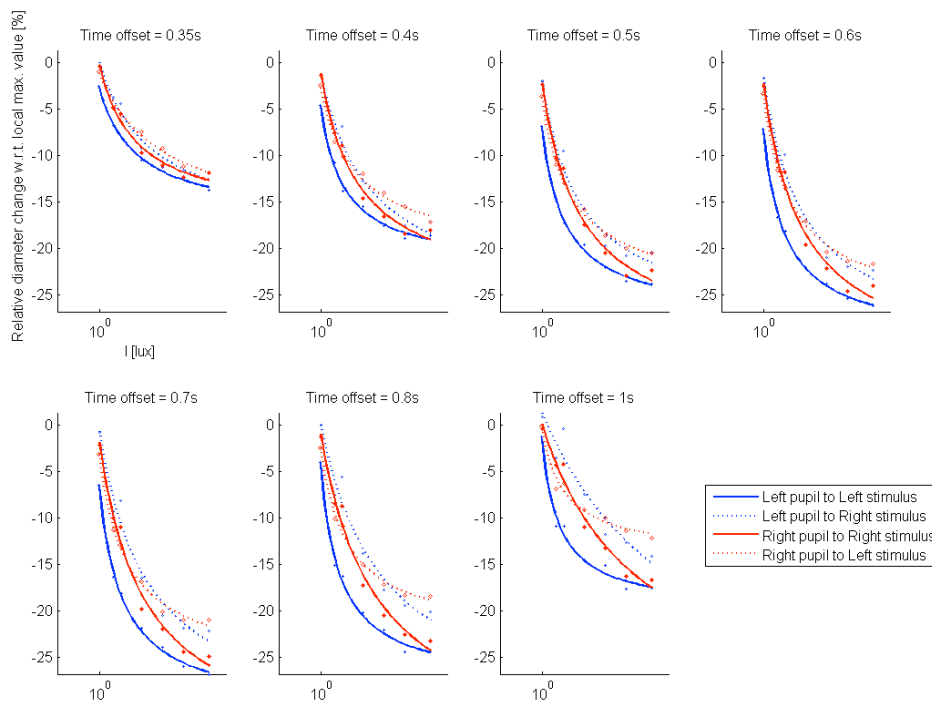

Figure 6. Example of a generalized hyperbolic function fitted to normalized pupil responses across several stimulus intensities at different time slices from stimulus onset. At each time point, the graph represents the pupil size at that time point during the pupil waveform for seven different intensities. By fitting the function at each time point, the pupil size at that time point in the waveform can be interpolated for small increments of intensity. This allows us to reconstruct the waveform pupil response for an entire intensity series for a stimulus to each eye (see reconstructed waveforms for the contraction part of the pupil reaction shown in Figure 7).
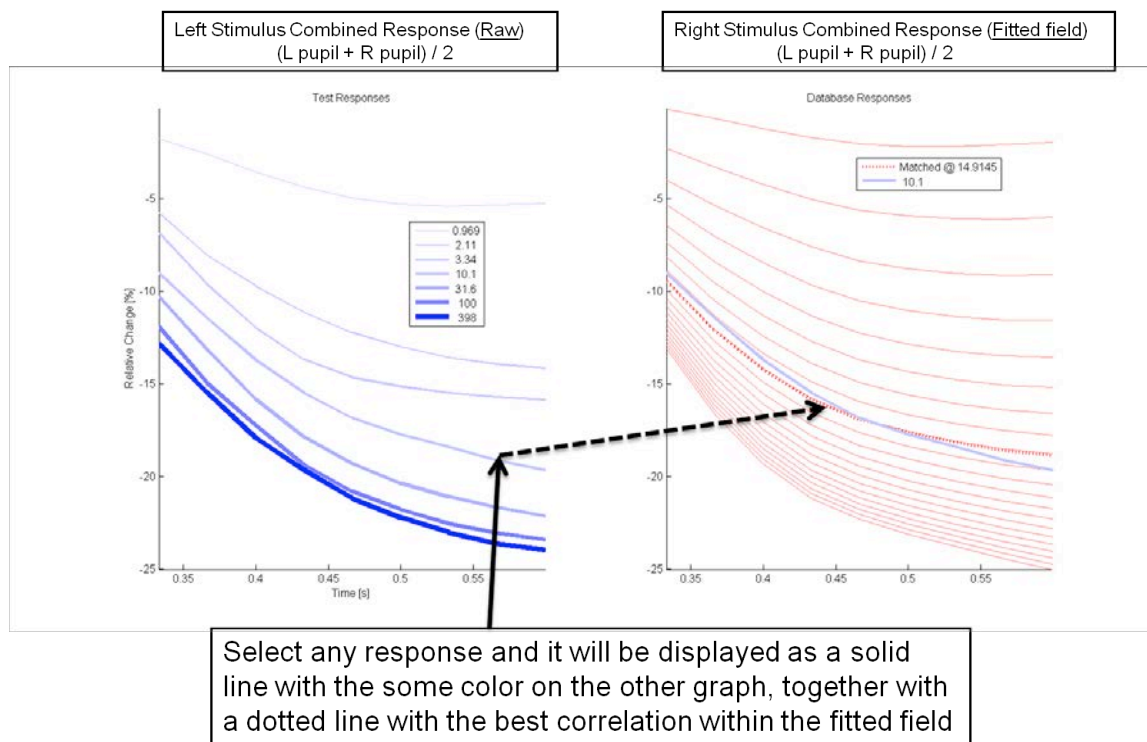
Figure 7. Comparison of waveforms from the left and right eye stimuli. Graph of waveforms is shown close up during the contraction time window of the pupil light reflex. Example of the best-fitting waveform response resulting from every 0.1 log unit increment of intensity within a pupil response field (graph on the right) given any test pupil response waveform selected from a set of measured responses shown in the graph on the left. The blue waveform resulting from a 10 lux stimulus to the left eye is superimposed on the different waveforms resulting from different stimulus intensities given to the right eye and the most correlated one is found (in this case, the right eye had to be stimulated about .25 log units brighter to give the same waveform). In this example, the response field was computed from the combined left and right pupil responses recorded for the right stimulus, while the mean left and right pupil responses recorded for the left stimulus were used as the test responses. In this particular case, the response at 10.1 lux was selected from the left graph, while the matching stimulus intensity from the response field on the right was computed to be 14.9 lux, which represents about a 0.25 log unit relative afferent pupil defect in the right eye.
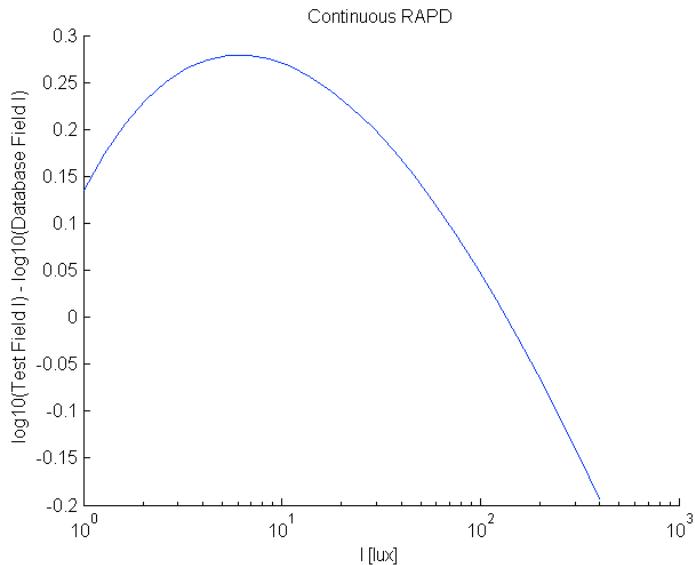
Figure 8. Example of the resultant calculated relative afferent pupil defect (RAPD) derived from the pupil waveforms (see close up of matching waveforms shown in Figure 7), as a function of light intensity used ("Continuous RAPD"). The one response field (in this case the right eye's response to light shined into it was compared against another response field (in this case the left eye's response to light shined into it). The graph for this example was calculated by first replacing the test responses in Figure 7 with a response field and then calculating the matching responses within the database response field in a continuous fashion. In other words, the derived pupil tracings from stimulus in the right eye were compared to the pupil tracings from stimulus to the left eye. The program finds the intensity of stimulus in one eye that gives the same pupil response at a fixed intensity in the other eye to derive the log unit RAPD. In the example shown there was a small right RAPD (positive numbers on the axis) that peaked at an intensity of around 10 lux (10^1).

The log unit RAPD−calculated in this manner as a function of how bright of a light is used – shows specific patterns of increasing or decreasing values for certain types of retinal and optic nerve disorders. Presently, in clinically derived RAPD measurements, one brightness level of light is used to compare the input between the two eyes. Our pilot data so far indicates that there may be more diagnostic information in deriving the RAPD as a function of brightness of the stimulus light.

During the last two quarters of Year 3, a significant effort was also spent to "clean" the pupil data of artifacts in the pupil tracings caused by loss of pupil border tracking that occurs in real time during a test. After poring over data collected from a number of subjects to support the implementation and populating of a database of age-matched pupil light reflexes, we have realized that additional pre-processing of the data is required to ensure that the database is populated with data of sufficient quality.

Specifically, we have noticed that the pupil tracking algorithms used in the Neuroptics

DP2000 system (and similarly, for the Smart Eye remote eye tracking system) consistently fail to locate and correctly estimate the size of the pupil after an eye blink, or even in case of a partly covered pupil due to droopy eye lids or squinting. The failure of the tracking algorithms in these cases are mainly a result of the real-time processing requirements placed on these commercial pupil and eye tracking systems, but nonetheless causes a significant number of subpar quality or "bad" data points during the course of a pupil light reflex, which we would like to either correct or delete from a data set before it is committed to the database.

Fortunately, in case of the DP2000 system, we have access to the raw video data, albeit at a quarter of the initial level of resolution (implemented to minimize storage requirements). Based on this need for clean data, we have spent a considerable amount of time to implement a user interface to allow an expert user to quickly and efficiently perform manual artifact detection and removal based on all available data -- that is, both video and pupil tracings -- and numerous tools and widgets. We envision that the outcome of this pre-processing step will be multifold: i) it will yield clean data for populating the database immediately; and ii) provide many instances of paired artifact and cleaned data segments that can be used to train future automated artifact detection/removal algorithms to minimize manual operations, etc.

In future, we seek to automate processing of a new patient's data before comparison against the database as much as possible, for say use in a clinical setting, and we recognize the fact that we might be able to lower the level of rigorous artifact detection/removal in case one wants to simply compare a patient's data against the database, but does not intend to add the patient's data to the database per se.

Below we show several screenshots of various aspects of the pupil cleaning software interface. This will allow us to post-process the pupil data after the experiment to further improve the pupil measurements even when tracking fails or when excessive blinking occurs during a test.
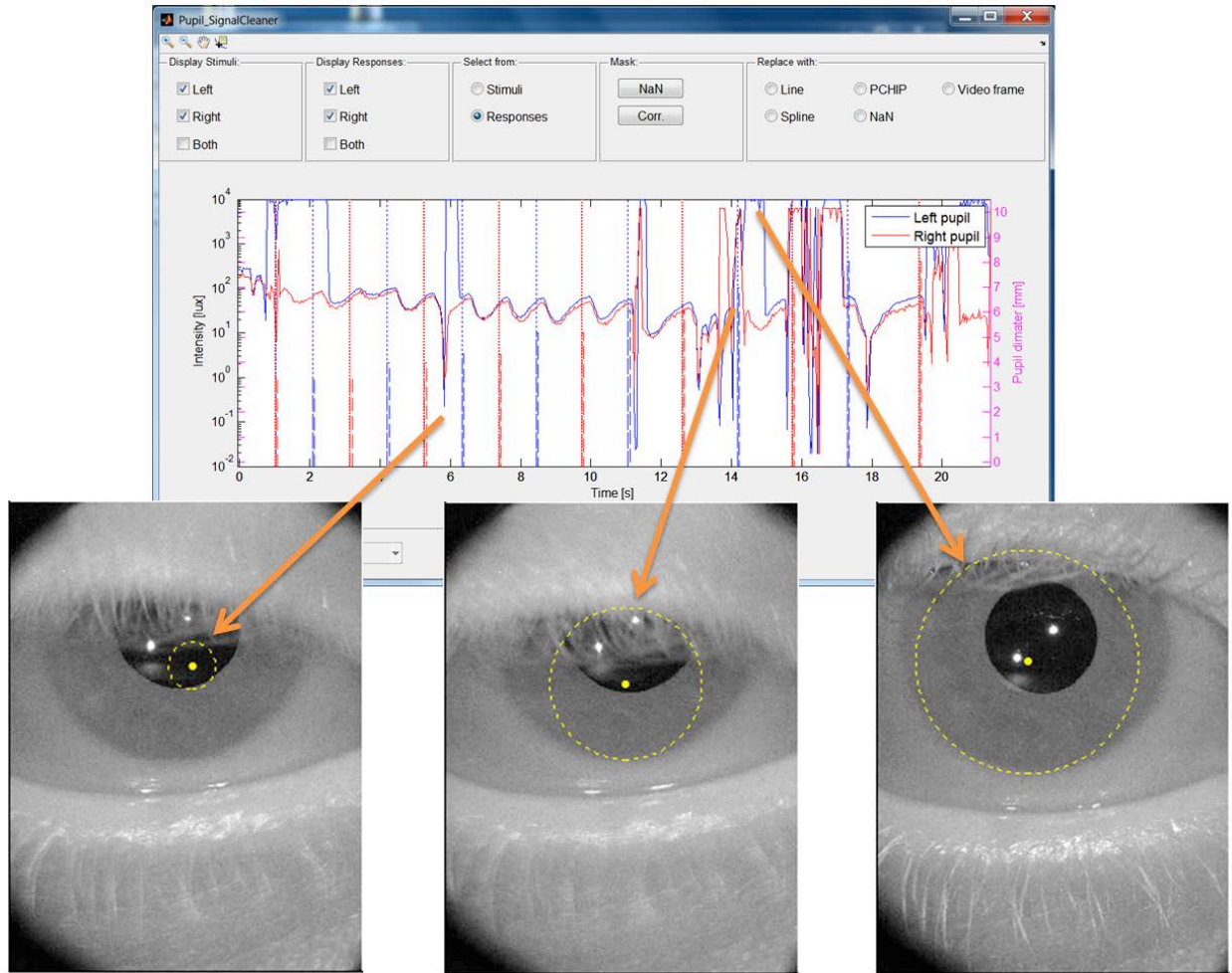
Figure 9. Three examples demonstrating the failure of the pupil tracking algorithm to locate and estimate the correct size of a pupil. It is important to note that without a video data record that is synchronized to the estimated pupil size values (as provided by the pupil/eye tracker), one would not be able to tell whether obviously "bad" pupil size data points could be salvaged (examples on the left and right), or whether a seemingly "clean" pupil size data point is correct by chance (example in the middle). If these points are not removed or corrected, subsequent data analysis steps might suffer, while the data quality of surrounding data points can be negatively impacted, such as would be the case if a digitally filter is applied to the tracing, etc. The top figure is the right(red) and left (blue) pupil waveforms (pupil diameter on the y-axis and time in seconds on the x-axis in response to a rapid repeating light stimulus varying in intensity. The video captures are shown below at different portions of the tracing, indicated by the arrows where the pupil tracking failed. Note that the camera sensor is oriented 180 degrees, so that the eye and eyelids appear "upside down". The dashed yellow circle in each captured frame of the video indicates where the Neuroptics software indicated the x-y position of the pupil and its size based on the real-time tracking algorithm that is used, which is stored in data for us to recover and superimpose on the captured video frames during the test.
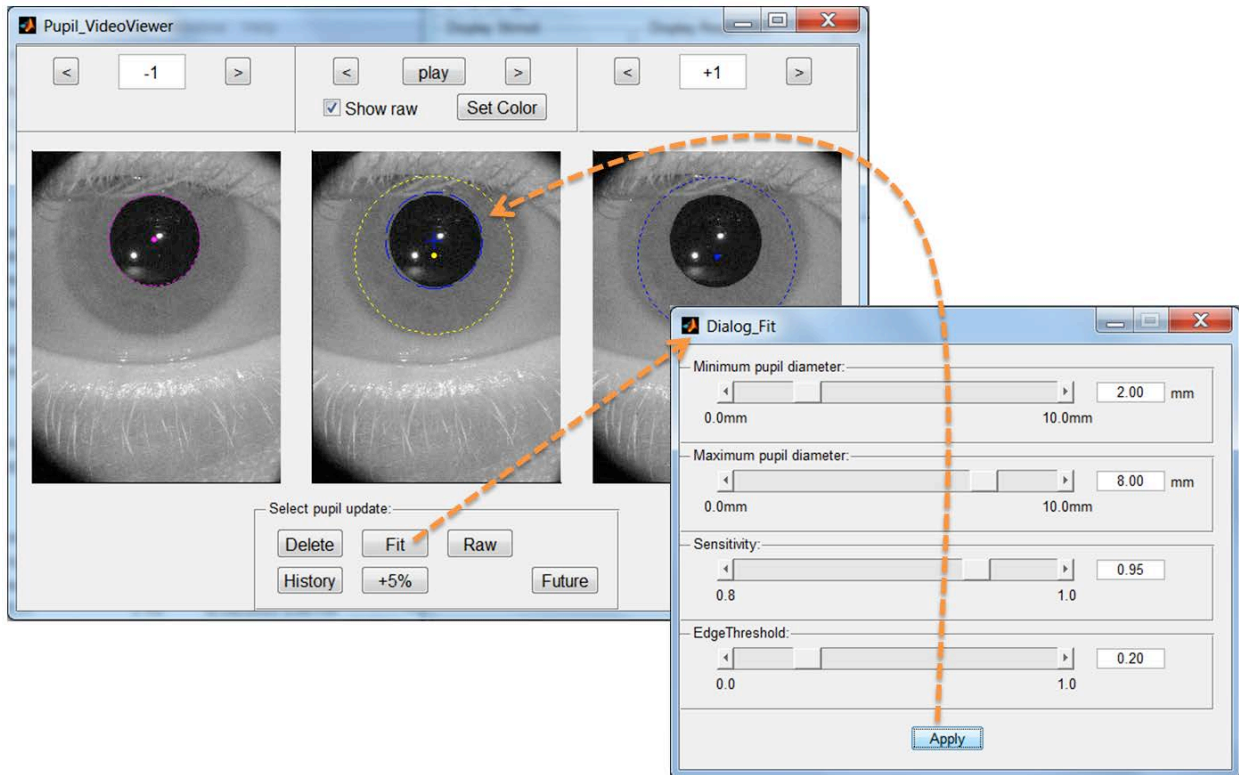
Figure 10. In case a synchronized video record is available, the expert user can launch a video editing interface to process any point on the pupil size tracing as selected from within the main interface (shown in Figure 9). Various widgets are currently implemented to change the size of the fitted pupil circle, or to drag the center of the pupil circle to a new location. For this example, the user has selected the "Fit" widget, which allows for adjusting several parameters before the Hough transformation -- for finding circles in an image -- is invoked by hitting "Apply". Any changes in the size of the fitted pupil circle is immediately reflected in the pupil tracing on the main user interface. In this case, the "Fit" button calls the standard "imfindcircles" function from Matlab's image processing toolbox, but we can replace it with any other functional algorithm if the need arises, or add additional buttons to the interface to activate new image analysis functions. After the new fit is applied, the resulting pupil size and location is also shown (blue circle in the middle video frame displayed), which is much more accurate than the algorithm determination that was made in real time (yellow dashed circle; obviously not accurate).
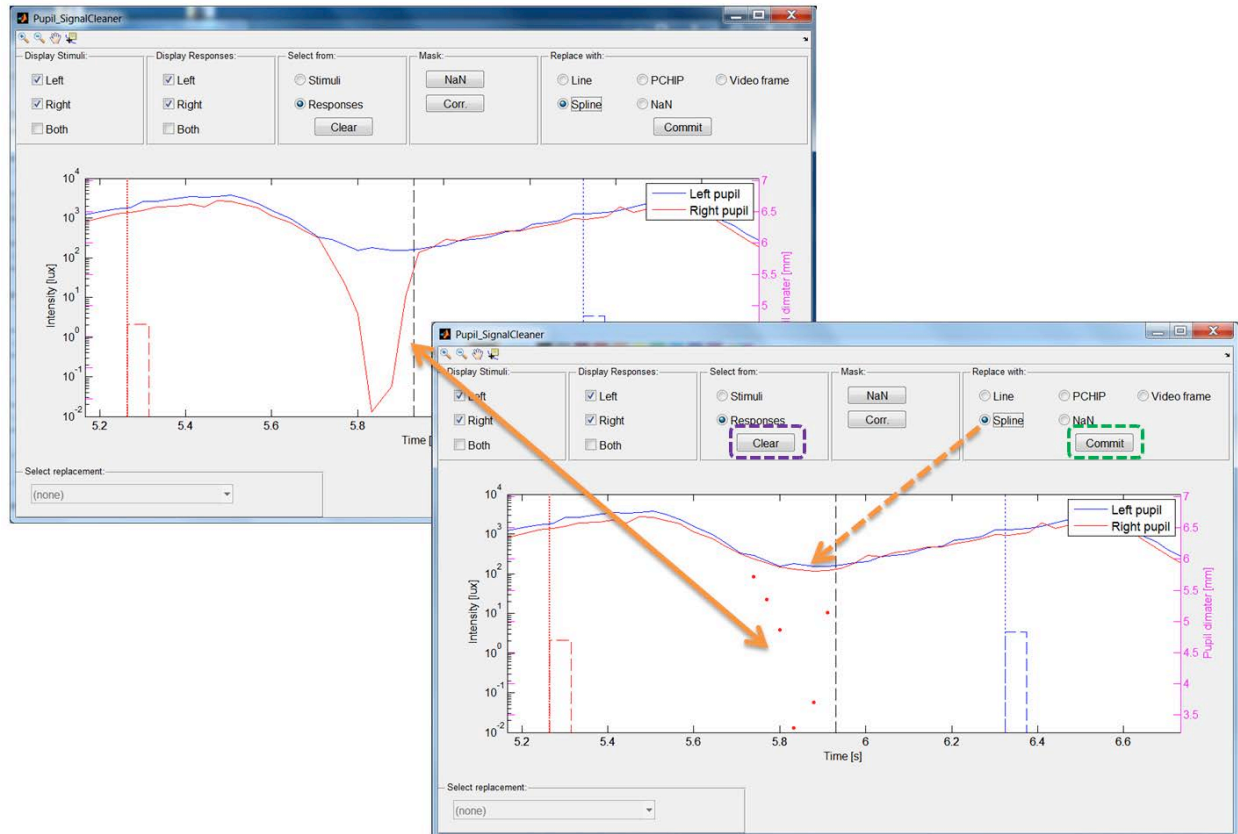
Figure 11. In case no synchronized video record is available, the expert user can use the main interface to select one or many points (shift + mouse click to select a region, or control + mouse click to toggle whether a single data point is to be included), and then apply different artifact removal algorithms, or mark the data to be completely ignored, i.e., replacing the selected pupil size data points with NaNs (not-a-number). In this example, the user has selected six (6) data points from the Right pupil tracing, and replaced it with a cubic spline interpolated values. The user can toggle through the radio buttons and inspect the quality of the different artifact removal procedures, and then has the option to either clear the selection (and select new points) by pressing the button marked with a purple box, or to commit the cleaned data points for further use by pressing the button marked with a green box.
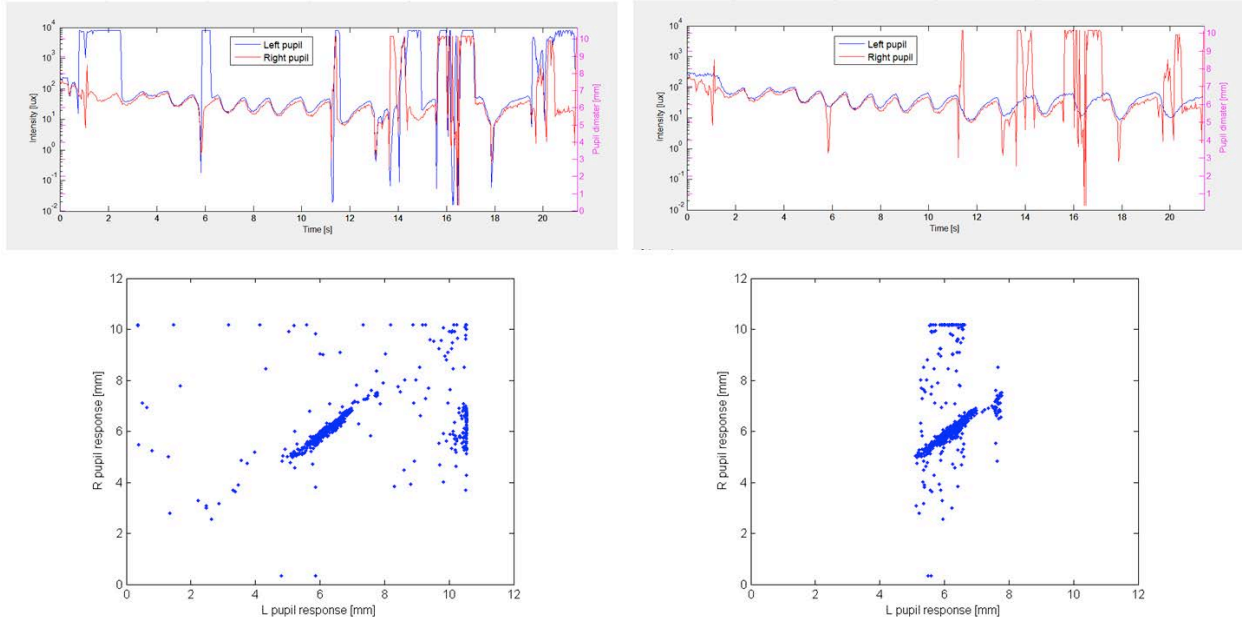
Figure 12. Correlation analysis can offer another potential approach to identify "bad" pupil size data points as part of automated data processing that we are in the process of implementing in the near future. In this example, the left hand scatter plot of Left versus Right pupil size data contains many outlier points, while the scatter plot on the right demonstrates the remaining spread of outliers due to "bad" Right pupil size data after the Left pupil tracing has been corrected through use of the video editing tool. Data points from normal bilateral pupil behavior are expected to cluster together in a tight band alongside a 1:1 straight line, with an intercept offset in case of anisocoria.

1c. TrueField Analyzer visual field testing device (Seeing Machines Limited)

The TrueField Analyzer that has been developed by Dr. Ted Maddess at Canberra University in Australia for Seeing Machines Limited has still not been commercially released for patient testing at this time. Dr. Kardon visited their laboratory in September 2011 in Australia. As stated in previous progress report, Seeing Machines would not provide us with the multifocal pupil device for visual field objective perimetry due to continued problems that they have had with their prototype stimulus system. Because of this we have developed our own focal pupil perimetry device using the same stimulus platform as we are using for visual evoked responses taking advantage of an innovative stimulus paradigm that will allow us to record and analyze pupil and visual evoked responses regionally from the same subject, using our multifunctional platform.

1d. "Dry electrode" wireless EEG system (Sigmed, Inc.) (months 7-12) –acquired in Year 2 and being integrated into our testing device in Year 3:

The dry electrode system was obtained for recording EMG, ERG and VEP responses to light stimuli. During Year 3 we have begun to design a comfortable head band and face appliance that will allow us to mount the wireless electrodes so that they can be easily worn and

make contact with the skin on the scalp and face without the need for any skin preparation or solutions. As a starting point, we are using the head gear design from Emotiv who had manufactured a head mount with EEG electrodes for video gaming, as shown in Figure 13 below.

Figure 13. Emotiv's EEG electrode headset that we are using as a model for visual cortex electrode headset and flexible placement of electrodes for VEP measurements of regional visual field defects.

In the last quarter, we have formalized and started implementation of the fabrication of a headset that will hold the dry wireless electrodes on the skin overlying the orbicularis and procerus muscles for recording EMG responses to light (for our other DOD grant). We are incorporating into this same design, the portions of the headband that will be used to position the scalp electrodes over the visual area of the brain (occipital cortex) for wireless recording of the visual evoked potentials (VEP) for regional visual stimuli. This work has been done in conjunction with Mark Ginsberg, owner of Ginsberg Jewelry in Iowa City and Bounnak Thammavong, his design assistant, to develop a molding which can be adjusted to change the tension between the electrodes and the skin and which can be rendered using a 3D printer. This will replace the standard skin electrodes that are adhesive and require skin preparation. This local group has the expertise that we need for developing a 3D CAD model and printing of the molding to which we will mount the electrodes. The 3D printer can print the molding on-site using a flexible and soft polymer resin, whose properties we can specify, including a polymer that can be autoclaved for sterilization. An example of the CAD design of a prototype is shown if Figure 14 in different views.
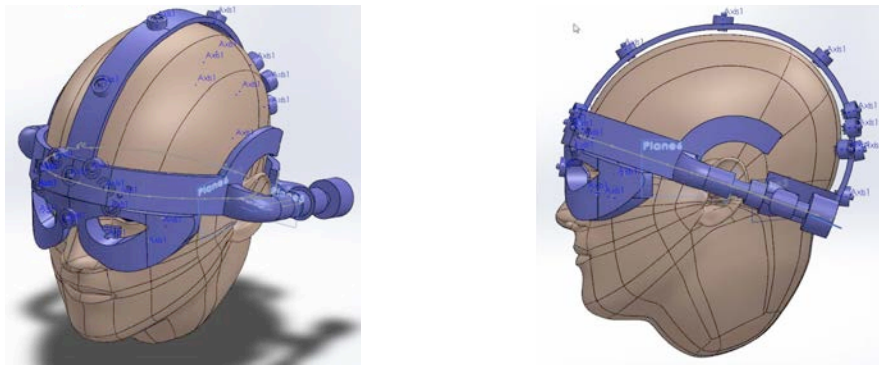
Figure 14. 3D CAD rendering of head/face mount for EMG wireless electrodes that is in the process of being finalized. The end product will be more streamlined, lightweight, and will not be as bulky as shown here. Once finalized, the printing of the head

1e. Multi-camera eye movement monitor (Smart Eye AB) and visual stimulus software platform (months 1-24)

The multi-camera Smart Eye system has been fully operational and was moved into our new clinical testing suite, but had to be moved again during Year 3 to a room in the same hallway in the Ophthalmology Department, due to a noise distraction from the laser in the original room, whose fan must be running at all times. The set up allows the subject to sit 7 feet from the cameras and monitor upon which they are mounted, allowing us to obtain high quality multi-vantage point video recordings of each eye and pupil position during the testing. This allows us to derive point of gaze and pupil response in real time in response to different visual stimulus paradigms. The transfer of all of the equipment into a dedicated human testing room within the confines of the ophthalmology clinic, will allow us to easily test patients and normal subjects.

Visual stimulus presentation and data analysis software platform

In our Annual Report for Year 2, we have reported on our efforts to optimize a combined experimental design-analysis methodology to reliably extract response waveforms for visual stimuli at multiple spatial locations. These waveforms can be in response to both multi-focal visually-evoked potential (VEP) and electroretinogram (ERG) stimuli. We have mentioned that findings from literature make it clear that the traditional evoked response potential (ERP) or transient response (TR) methodology fail to separate individual responses into their corresponding components when stimulus rates are increased in an attempt to shorten testing times and responses inadvertently start to overlap. In an effort to circumvent the time separation limitation of ERPs (TRs), faster stimulation rates can be used with the steady-state response (SSR) method, in which case the complex waveform of the combined responses is analyzed in the frequency domain by measuring spectral peaks. However, overlapping responses and the periodic nature of the SSR stimulus sequence still lead to the loss of critical information provided by waveform morphology and waveform components of individual responses.

To circumvent the limitations imposed by conventional averaging, researchers in ophthalmology and visual sciences use specially designed stimulus sequences called maximum length sequences (MLS or m-sequences), which allows for the separation of an individual response from other overlapping responses. Although m-sequences are not part of the International Society for Clinical Electrophysiology of Vision (ISCEV) standard for multi-focal VEP studies, it forms the basis of the ISCEV multi-focal ERG standard, and compelled us to investigate the use of the MLS method in more detail. We have therefore decided to compare the m-sequence model to our proposed design matrix-based model in terms of statistical metrics in order to shed more light on the strengths and weaknesses of both models in terms of our scenarios of use. It is in fact possible to use a design matrix to model (capture) the stimulus sequences used in an MLS experiment (i.e., m-sequences), which allows for a direct comparison of the analysis results between the MLS method and the least squares solution of the design matrix-based method on the same recorded or simulated data set.

Researchers have previously proofed the validity and reliability of the MLS method and the existence of other similar sequences such as the Legendre sequences. However, a major limitation of the MLS method is rooted in its limited choice of sequences. Each sequence out of an MLS set is constrained to be approximately orthogonal with respect to the other sequences from the same set, imposing a wide range of stimulation rates with a high degree of stimulus onset jitter on each sequence. On the other hand, a design matrix offers flexibility in terms of the stimulation rate, semi-random onset jitter, and other stimulation sequence parameters that can be set for each individual stimulus location. It is also important to note that the occurrence of missing or corrupted data segments poses a significant problem for the MLS method, while it is trivial to remove time points which are deemed corrupted due to noise, subject inattention, etc. from a design matrix.

In general, a broad-band stimulation rate (e.g., an m-sequence) favors efficient waveform morphology estimation, while a narrow-band stimulation rate allows for a more accurate estimate of the overall response amplitude, and therefore greater detection power. For the case of semi-random onset sequences, waveform estimation and detection power are balanced more evenly. A significant strength of using a design matrix revolves around the ability to alter the balance of the model to focus on efficient estimation of the unknown waveform morphology *versus* focusing the model on the accurate detection of responses in the presence of noise. This trade-off can be adjusted for each response by adding additional constraints to the model in the form of a set of basis functions to describe each response waveform in terms of smoothness, frequency spectrum, etc. For example, it is possible to use a set of sinusoids at relatively low frequency if the expected response waveform is smooth and lack higher frequency components.

From our literature study, we have found a few model evaluation guidelines, which include both statistical and graphical techniques. Quantitative statistics include standard regression, dimensionless, and error index metrics. Standard regression statistics determine the strength of the linear relationship between simulated and measured data, and include Pearson's correlation coefficient (r) and the coefficient of determination (R^2). The correlation coefficient r, ranging from -1.0 to 1.0, indicates the degree of linear

relationship between observed and simulated data. If r = 0.0, no linear relationship exists, while for r = 1.0 or −1.0, a perfect positive or negative linear relationship exists. Similarly, R^2 describes the proportion of the variance in the data explained by the model and ranges from 0.0 to1.0. However, although widely used for model evaluation, these statistics are oversensitive to high extreme values (outliers) and insensitive to additive and proportional differences between model predictions and measured data.

Dimensionless techniques constitute a second class of quantitative statistics, and provide for relative model evaluation assessment. The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance ("noise") compared to the observed data variance ("information") (Nash and Sutcliffe, 1970). The NSE metric indicates how well the plot of observed (i.e., underlying "ground truth") data $versus$ modeled data fits the 1:1 line, and is computed as:

$$ NSE = 1 - \left[\frac{\sum\limits_{n} \left(Y_i^{obs} - Y_i^{model} \right)^2}{\sum\limits_{n} \left(Y_i^{obs} - \overline{Y} \right)^2} \right] $$

where $\overline{Y}$ is the mean of the observed ("ground truth") data, and n is the total number of observations. The NSE metric ranges from −∞ to 1.0, with $NSE = 1.0$ being the optimal value. Values between 0.0 and 1.0 are generally viewed as acceptable levels of performance, whereas values below 0.0 indicate that the mean observed value is a better predictor than the modeled values, and signify unacceptable performance of the model.

A third class of quantitative statistics consists of error indices to quantify the deviation between the modeled and observed underlying data in the units of the data of interest. The trade-off between bias and residual variance has to be considered, as recommended by Boyle et al. (2000). Bias measures the average tendency of the modeled data to be larger or smaller than their observed counterparts (Gupta et al., 1999). Residual variance is the difference between the measured and simulated values, often estimated by the residual mean square error (*MSE*) or root mean square error (*RMSE*). Optimizing *RMSE* during model calibration may give small error variance but at the expense of significant model bias.

Percent bias ($PBIAS$) is computed as:

$$ PBIAS = \frac{\sum\limits_{n} \left(Y_i^{obs} - Y_i^{model} \right) * 100}{\sum\limits_{n} Y_i^{obs}} $$

The optimal value of $PBIAS$ is 0.0, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias.

RMSE-observations standard deviation ratio (RSR) is one of the commonly used error index statistics in hydrologic modeling (e.g., Singh et al., 2004). The RSR metric standardizes $RMSE$ values using the observations standard deviation, and is calculated as the ratio of the $RMSE$ and standard deviation of the observed data:

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_n \left(Y_i^{obs} - Y_i^{model}\right)^2}}{\sqrt{\sum_n \left(Y_i^{obs} - \overline{Y}\right)^2}}$$

The range of RSR extends from the optimal value of 0.0, which indicates zero RMSE or residual variation and therefore perfect model fit, to a large positive value. The lower the $RMSE$ value, the lower the RSR value, and the better the model performance.

We have implemented a software-based test bed to evaluate the impact of various parameters on the efficiency of different models for extracting the waveform morphology and activation level of multiple responses. The test bed allows us to specify the number of response areas, individual response waveforms, individual stimulus sequences, experiment duration, the nature of basis functions to be used with the design matrix method, as well as additive noise profiles. Based on these parameters, simulated data sets are created and analyzed with the MLS method and the design matrix method afterwards. Model evaluation statistics can then be computed to compare the efficiency of each model to extract the underlying "ground truth" responses in the presence of noise, as well as the effect of other model assumptions on the overall outcome. We plan to use the test bed to fine tune the stimulus sequences for the VEP portion of the project in order to further optimize the bias *versus* error variance trade-off, while trying to make each test protocol as short as possible and maximize the robustness of the methodology against the impact of corrupted data and subject inattention, etc.

For the figures following, we have set up our test bed to evaluate the effect of different levels of noise, as well as different basis function sets on the MLS and design matrix methods for 61 individual stimulation locations (and therefore 61 different responses). The "ground truth" response waveforms were chosen as a windowed sinusoid with a central frequency proportional to the index number of the response (i.e., response #1 had the lowest frequency and response #61 had the highest frequency), while the amplitude of the waveform hills were increased by 2.5x with respect to the valleys to create a non-zero mean value. The sampling rate was set at 1,000Hz, and the minimum inter stimulus interval (ISI, or base period in terms of the MLS method) was 50ms. Two stimulus onset sequence protocols were used: the first set was based on standard m-sequences, and the second set was created by allowing the ISI to vary randomly in a constrained interval of 50 to 150ms. The parameter selection of the latter ensured that both protocols were similar in terms of the mean number of stimulus presentations for each stimulation location. Additive noise values were modeled as normally-distributed with zero mean.

Based on our simulation results, the design matrix method offers us a more efficient way to analyze multi-response data compared to the MLS method, while the m-sequence stimulus onset protocol tends to produce lower bias at the expense of higher error variance when compared to the constrained (pseudo-) random stimulus onset protocol in case of the design matrix analysis method.
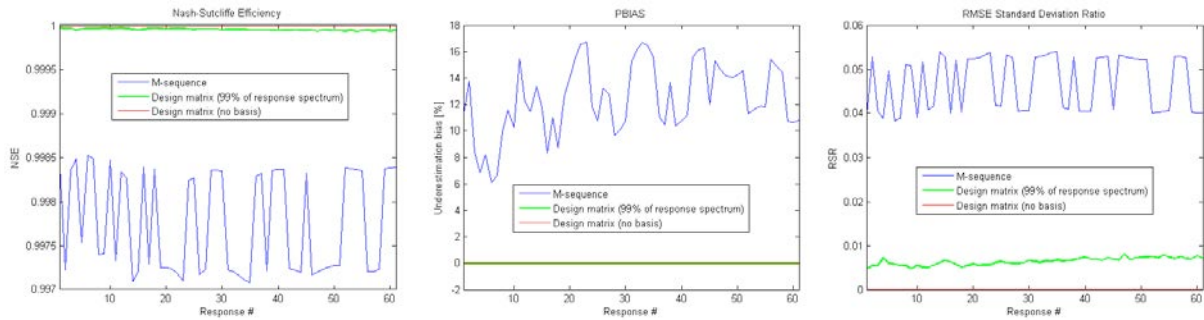

Figure 15. Noiseless case: With no additive noise, the design matrix analysis method outperforms the MLS analysis method in terms of bias (*PBIAS* metric), as well as error variance (*NSE* and *RSR* metrics). As expected, the use of a constrained basis for the design matrix method leads to increased bias and error variance. The same m-sequence stimulus onset protocol was used for both MLS and design matrix analysis methods.
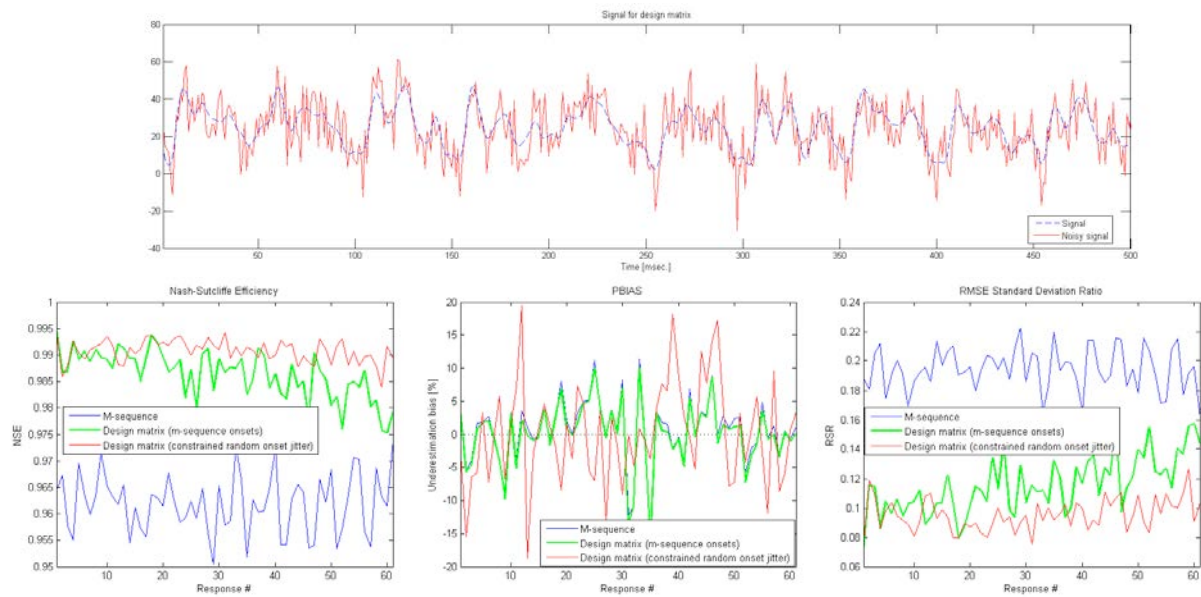
Figure 16. Simulation results with the standard deviation of the additive noise set to 10 units. The top graph shows the combined waveform of 61 responses, activated by the constrained (pseudo-) random stimulus onset protocol, with and without noise over a representative 0.5s interval. The 3 bottom graphs demonstrate the improved analysis efficiency of the design matrix method (green line) in comparison with the MLS method (blue line) for the case of the m-sequence stimulus onset protocol. The m-sequence stimulus onset protocol (red line) tends to produce lower bias at the expense of higher error variance when compared to the constrained (pseudo-) random stimulus onset protocol in case of the design matrix analysis method.

24
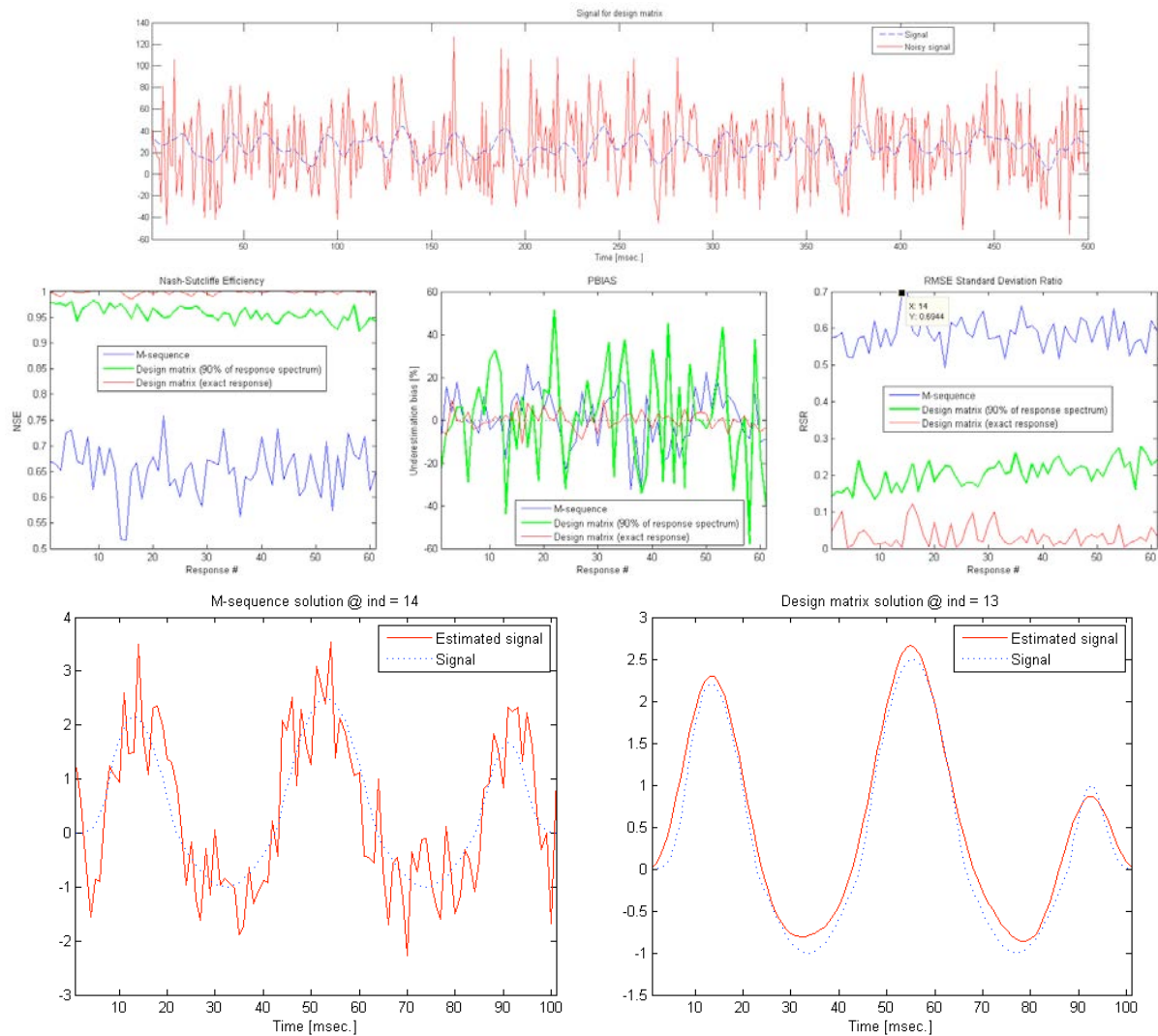
Figure 17. Simulation results with the standard deviation of the additive noise set to 30 units. The top graph shows the combined waveform of 61 responses, activated by the constrained (pseudo-) random stimulus onset protocol, with and without noise over a representative 0.5s interval. The 3 graphs in the middle demonstrate the lower error variance but higher bias tendencies of the design matrix method (green line) with the constrained (pseudo-) random stimulus onset protocol in comparison with the MLS method (blue line). Also shown is the special case for which accurate waveform morphology information is available, in which case the design matrix method (red line) can efficiently extract the activation amplitude of the responses in the presence of overwhelming noise. The 2 bottom graphs show the "ground truth" *versus* modeled response waveforms for the MLS analysis method (left) and the design matrix analysis method (right) for representative responses with larger deviations from the "ground truth" data.

As stated previously, because of difficulties in identifying adequate space and location to accommodate our subjects, the human testing has been delayed. Since we recently were
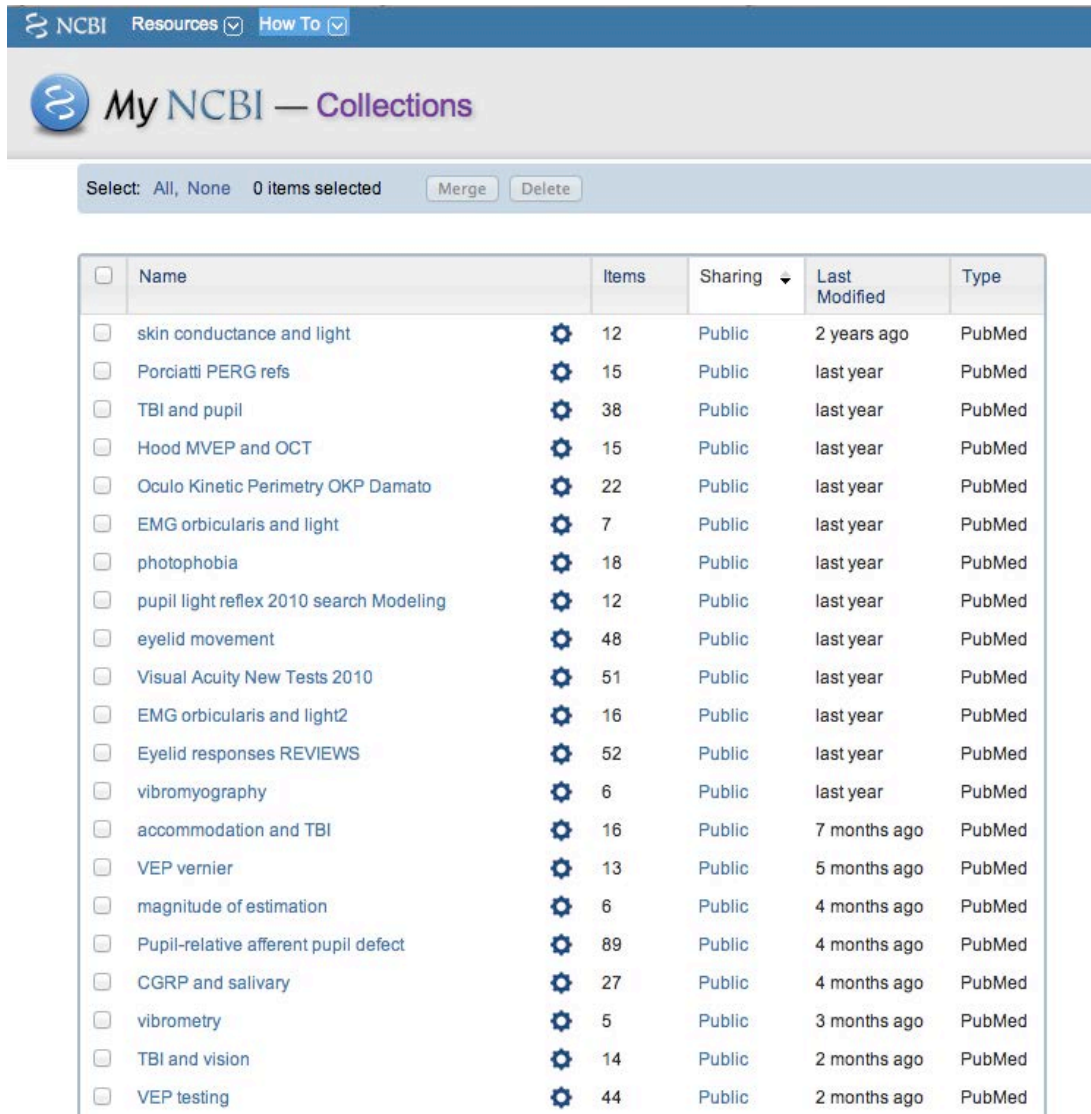
able to reconfigure human testing space within the ophthalmology clinic, we can now start recruiting subjects to begin testing. We have identified normal human subjects and those with TBI from our database patient and subject files and we are now in the process of scheduling their testing. Figure 18 shows the new testing suite and set-up platform for the objective testing of visual responses (pupil responses, eye tracking of visual optotypes, and visual evoked potentials).

Figure 18. New clinical testing space recently acquired within the confines of the Ophthalmology Clinic. The test subject sits comfortably in front of a large screen which can present a large range of light stimuli, varying in color, intensity and size which are used to elicit objective responses such as EMG of the orbicularis and procerus muscles, skin conductance and heart rate.

1e. Literature study (months 1-12)

Completed in year 1 and reported in Year 1 annual report :

Figure 13. Example of literature search on topics in PubMed that was performed.

The literature search is an ongoing process, however, as we continually are updating our information with recently published articles that are relevant to our proposed research.

Task 2. In normal eyes, collect, correlate, and define normative range of values for both objective and standardized tests (months 13-18, now anticipated to occur in months 37-46):
2a. Pupil light reflexes (revised; months 37-46)
2b. Evoked potentials (revised; months 37-46)
2c. Eye movements to targets changing in resolution (revised; months 37-46)
2d. Standard light threshold perimetry (revised; months 37-46)

In Year 2 we submitted the forms and protocol to our local IRB3 Human Use Committee and finally did receive the IRB-approval, and this was followed by final approval by Brigit Ciccarello after submission to ORP HRPO after their second-level review and approval was received. This occurred in the last quarter of Year 2, which took longer than anticipated, putting us a year behind schedule in the planned human testing of normal subjects and patients with anterior pathway disorders (Task 3, below). During Year 3, while waiting to secure a clinical testing suite in the Ophthalmology Department at the University of Iowa, we took advantage of this time to further refine our software and hardware capabilities for testing the pupil light reflex, visual evoked potentials from the eye and visual cortex, and eye movement tracking to vanishing optotypes. We have subsequently submitted a two year, no cost extension of the grant starting in October 2013 in order to complete all of the planned testing (see SOW below).

Task 3. In eyes with damage to the retina or optic nerve, collect, correlate, and define normative range of values for both objective and standardized tests (revised; months 47-50):
3a. Pupil light reflexes (revised; months 47-50)
3b. Evoked potentials (revised; months 47-50)
3c. Eye movements to targets changing in resolution (revised; months 47-50)
3d. Standard light threshold perimetry (revised; months 47-50)

Task 4. In eyes with damage to primary visual cortex (V1), collect, correlate, and define normative range of values for both objective and standardized tests (revised; months 51-54):
4a. Pupil light reflexes (revised; months 51-54)
4b. Evoked potentials (revised; months 51-54)
4c. Eye movements to targets changing in resolution (revised; months 51-54)
4d. Standard light threshold perimetry (revised; months 51-54)

Task 5. Optimize hardware systems and vision testing protocols to reduce testing time, maximize the signal/noise, and minimize cognitive demands placed on the patient for proposed objective tests (revised; months 37-58):
5a. Develop robust discriminators of visual pathway dysfunction and location by testing normal participants and those with well-defined visual field defects from damage to specific sites in their visual pathway (e.g. retina, optic nerve, post-geniculate homonymous damage) (revised; months 37-58)
5b. Optimize signal to noise by informed selection of visual stimuli and innovative analysis approaches (revised; months 37-58)

KEY RESEARCH ACCOMPLISHMENTS (SUMMARY)

- Development of a completely objective test protocol and associated data analysis routines to derive metrics that correlate with clinical diagnoses of afferent and efferent pathway pathologies based on the pupil light reflex. This test protocol consists of 8 white light impulses presented sequentially to each eye across a range

of intensities (1-400 lux), and is less than 22 seconds in duration. The Pearson correlation coefficient between the continuous RAPD metric derived from this test protocol and clinical measurements taken from the same subjects is 0.86.

- Implementation and testing of a multi-functional user interface to allow an expert user to quickly and efficiently perform manual artifact detection and removal, using both video and pupil tracings, in cleaning pupil light reflex (PLR) data. Clean data are needed to populate our PLR database, and to provide a sufficient number of instances of paired artifact and cleaned data segments to train future automated artifact detection/removal algorithms to minimize manual operations.
- Comparison of our in-house design matrix method for visually-evoked potential (VEP) stimuli with the well-known maximum length sequences (MLS or m-sequences) method based on simulations and various statistical metrics. We have shown that the design matrix method offers us a more efficient way to analyze multi-response data compared to the MLS method, while the m-sequence stimulus onset protocol tends to produce lower bias at the expense of higher error variance when compared to the constrained (pseudo-) random stimulus onset protocol used as part of the design matrix analysis method.

REPORTABLE OUTCOMES

Since the first two years of this research was restricted to implementation of a hardware and software testing and analysis platform, we have not yet tested human subjects (years 3 and 4). Human approval was granted in the last quarter of year 2. Therefore, we do not yet have results of testing to report in the literature. However, the following two manuscripts were published on relevant work to this project for using the pupil light reflex to chromatic stimuli to diagnose and characterize photoreceptor disease of the retina:

1: Kawasaki A, Crippa SV, **Kardon R**, Leon L, Hamel C. Characterization of Pupil Responses to Blue and Red Light Stimuli in Autosomal Dominant Retinitis Pigmentosa due to NR2E3 Mutation. Invest Ophthalmol Vis Sci. 2012 Aug 15;53(9):5562-9. Print 2012 Sep. PubMed PMID: 22807301.

2: Kawasaki A, Munier FL, Leon L, **Kardon RH**. Pupillometric quantification of residual rod and cone activity in leber congenital amaurosis. Arch Ophthalmol. 2012 Jun;130(6):798-800. PubMed PMID: 22801849.

CONCLUSIONS

The research work that we are carrying out has important implications for the greater public good, in addition to its military relevance. Visual impairment from traumatic brain injury can occur in military personnel exposed to direct trauma to the brain or indirectly from blast injury. Similar damage to the visual system can also occur in the civilian population from TBI resulting from motor vehicle accidents and also from head injury due to contact sports at both the school and professional level. Traumatic causes of visual damage can also be additive after repeated episodes of head injury. Patients with visual pathway damage are often unaware of the problem and their associated cognitive

impairment may mask the underlying vision impairment and also prevent detection with standard tests of visual function, which require good cognitive performance and focused attention during the test. In addition, other forms of cognitive impairment in the general population such as attention deficit disorder, depression, and dementia prevent the accurate assessment of visual function. Patients with undiagnosed visual dysfunction and superimposed cognitive impairment may pose a danger to themselves and to others when tasks such as driving and other tasks, which demand good visual performance, cannot be safely carried out.

For this research, our main goal is to use objective reflexes of the visual system to diagnose vision deficits and ensure effective monitoring of their treatment, when indicated. Such tests will allow accurate testing of the visual system with almost no demands on cognitive function during testing. This will be possible because the constriction of the pupils in response to light, the electrical recording of light evoked potentials (voltage) from the skin overlying the vision centers of the brain and the monitoring of purposeful eye movements to track moving targets are all objective, natural reflexes of the visual system. We are taking advantage of these reflexes by implementing an integrated system to quantify them using a specially designed suite of rapidly performed tests requiring little patient cooperation. Once validated in our proposed study, these tests can be used in cognitively intact or cognitively impaired individuals to assess visual function, leading to rehabilitation and treatment when appropriate.

The availability of the objective tests of vision being developed and implemented will greatly improve eye care by providing faster, lower cost testing that can be performed in remote settings. This will provide easier access of the public to accurate assessment of their visual function and will also reduce the cost associated with current testing and transportation to sites of testing. Such tests will also provide a new tool for assessing innovative treatments being developed to save or restore vision.

REFERENCES

None.

APPENDICES

None.

SUPPORTING DATA

All figures including in body of report