Efficiently Measuring Magnocellular and Parvocellular Function in Human Clinical Studies

Andrew J. Anderson¹, Julie Jiao¹, and Bang V. Bui¹

¹ Department of Optometry and Vision Sciences, The University of Melbourne, Parkville, Australia

Purpose: Pokorny and Smith (J Opt Soc Am A Opt Image Sci Vis. 1997;14:2477–2486) described a laboratory method to behaviorally measure magnocellular and parvocellular pathway sensitivity. We investigated whether their method may be more efficiently applied to clinical populations by reducing adaptation times.

Methods: We measured contrast detection thresholds to a 30-ms increment on a 30 cd/m² background every 2 seconds after a 1-minute preadaptation to either a bright (90 cd/m²) or dim (3 cd/m²) luminance, in four observers. We also measured increment thresholds atop a steady 60 cd/m² luminous pedestal (30 cd/m² above the background) that remained on for 80 seconds, and tracked thresholds for 60 seconds after pedestal offset. We also assessed the minimum number of stimulus presentations required to reliably estimate thresholds using our four alternative forced choice (4-AFC) zippy estimation by sequential testing (ZEST) procedure.

Results: Detection thresholds between the bright and dim preadaptation conditions were identical within seconds after the offset of the preadaptation luminance. Thresholds on the steady luminance pedestal reached stable values within approximately 10 seconds from pedestal onset, and recovered within 2 seconds of pedestal offset. Analysis of the 4-AFC ZEST procedure found little decrease in threshold variability after approximately 14 stimulus presentations.

Conclusions: Preadaptation and stimulus adaptation times may be reduced dramatically from those described by Pokorny and Smith, without altering thresholds.

Translational relevance: Experimental time with clinical populations often is limited. Increasing the efficiency of the method of Pokorny and Smith allows for either shorter test sessions, or for a more extensive range of experimental parameters to be explored in disease.

Introduction

It is well established that visual information travels from the eye by means of parallel pathways, with two principal pathways being the magnocellular (M) and parvocellular (P) pathways.¹ These two pathways subserve different visual functions, with the M pathway being achromatic and showing high temporal frequency tuning, and the P pathway showing chromatic sensitivity as well as high spatial resolution.²,³ The retinal ganglion cells within each pathway,⁴ along with corresponding cells in the lateral geniculate nucleus,⁵ display anatomical and physiological differences,⁴ prompting researchers to investigate whether different pathways may be differentially affected in disease.⁶ Although direct measurement of the functional responses of each pathway is possible electrophysiologically in animal models, such invasive electrophysiological measures cannot be performed in human observers. Conventional electrophysiological measures in humans, such as the electroretinogram or visual evoked potential, do not distinguish between M and P pathway function. Therefore, finding behavioral methods that can distinguish between M and P pathways is important for human research.

There is a substantial overlap in the spatiotemporal tuning of the M and P pathways,³ making it difficult to isolate the response of one pathway using simple contrast detection tasks. Pokorny and Smith⁷ described a method that exploits contrast gain differences in the M and P pathways⁸ to better isolate each pathway behaviorally. Briefly, their method involves a
contrast discrimination task in which an observer is presented with a small group of luminous increments (pedestals), with one differing from the others in luminance (pedestal + ΔL). The observer must discriminate which target is more intense among a group. If the pedestals are briefly flashed (the pulsed-pedestal paradigm), the M pathway response saturates, leaving the nonsaturated P pathway to perform the discrimination task. If, however, the pedestals are steadily presented for a period before the presentation of the ΔL to be discriminated (the steady-pedestal paradigm), local adaptation mechanisms that are preferentially performing the discrimination task. An advantage of the method of Pokorny and Smith is that M and P isolating stimuli are identical at the time of the contrast discrimination judgment, differing only in terms of the eye's adaptation state immediately preceding this judgment. Because of this, the influence of preretinal factors, such as refractive error and media opacities, would be expected to be similar for both stimuli, unlike if M and P functions were attempted to be separated through the use of gratings of differing spatial frequencies or color. Their method has been used widely to measure how M and P functioning is altered in ocular, systemic, and cortical conditions (see review by Pokorny).

The method of Pokorny and Smith method was used originally in comprehensive laboratory investigations on young, healthy observers, and involved comparatively long periods of adaptation. For example, participants preadapted to a uniform display for 2 minutes, and then to the steadily presented pedestal for a further 1 minute, before any measurements commenced. Translations of their method for use in more clinical populations similarly have tended to use long adaptation times, although the selection of these times is somewhat ad hoc: for example, 2-minute preadaptation and 1-minute pedestal adaptation in a study of Leber's optic neuropathy, 2 minutes preadaptation plus 30 seconds before each stimulus condition for a study of amblyopia, and 1 minute pedestal adaptation only (no preadaptation) in studies of glaucoma. Given that test time often is limited with clinical observers, it would be useful to know if preadaptation and pedestal adaptation times might be reduced without significantly altering results. Investigation of the time course of adaptation processes reveals mechanisms operating on timescales of tens of milliseconds to a few seconds for contrast-based tasks, suggesting that there may be substantial scope for reducing adaptation times currently used in current clinical investigations that use the Pokorny and Smith method. Protracted adaptation times in the order of minutes do exist for tasks such as brightness perception, however.

Most threshold determining procedures inherently assume that threshold remains constant over time, and so are not suited to measuring adaptation effects that might change rapidly. We used a modified method, the “method of a thousand ZESTs,” that can estimate thresholds with a temporal resolution of a couple of seconds. Using this technique, we explored the role of the preadaptation period in the Pokorny and Smith protocol, as well as adaptation to, and recovery from, a steady pedestal. By investigating this, we aimed to determine what preadaptation and stimulus-adaptation times might be used in clinical studies to allow data collection in as short a time as possible.

The efficiency of yes/no and two alternative forced choice (2-AFC) ZEST procedures have been explored, with yes/no being markedly more efficient due to having a much smaller false alarm rate (typically a few percent, in contrast to the 50% false alarm rate in 2-AFC experiments). The four alternative forced choice (4-AFC) procedure used in the method of Pokorny and Smith has a false alarm rate of 25% that is intermediate to yes/no and 2-AFC procedures, suggesting a suitable number of presentations for a 4-AFC would be intermediate between the 8 presentations recommended for yes/no and the 30 recommended for 2-AFC. In addition to exploring adaptation times, we also assessed another important question in terms of test efficiency: how many stimulus presentations are required to return reliable threshold estimates in the Pokorny and Smith method?

Methods

Stimuli and Experimental Protocols

We presented stimuli on a calibrated computer monitor system (Visage graphics card; Cambridge Research Systems, Kent, UK and Diamond Pro 2070SB display; Mitsubishi Electric Visual Systems, Tokyo, Japan). Participants viewed the monitor at a 1 m distance in a dimly illuminated room, with their head supported by a chin rest. Viewing was binocular to avoid systematic changes in threshold over time that can result from opaque monocular patching, while important for the current study design, such
small changes are typically of no consequence when applying the method monocularly in ocular disease.\textsuperscript{10}

Participants performed three interleaved experimental conditions, all of which were matched for run length and the number of stimuli presented (Fig. 1). Each condition ran for 140 seconds, preceded by 60 seconds of preadaptation to a blank screen. The task was a 4-AFC, with stimuli being 1.0° square, 30 ms luminous increments centered 0.81° radially from the fixation point in one of four quadrants, and presented every 2 seconds after the preadaptation period. Participants responded by means of a button press, with a subsequent auditory tone indicating the correctness of the response. Stimuli were assumed to be below threshold if no response was given after 1500 ms,\textsuperscript{19} and so a randomized response was assigned (fewer than 1% of trials for any participant).

Although we test foveal vision in the current study, the Pokorny and Smith\textsuperscript{7} method may be used at eccentric locations in the visual field.\textsuperscript{10}

Conditions 1 and 2 were designed to investigate the influence of preadaptation on thresholds. Participants first adapted to either a bright (90 cd/m\textsuperscript{2}, Condition 1) or dim (3 cd/m\textsuperscript{2}, Condition 2) blank screen, after which increment thresholds were measured on a 30 cd/m\textsuperscript{2} background over a period of 140 seconds. Therefore, the final 20 seconds of each run in Conditions 1 and 2 were measured more than 2 minutes after the onset of the 30 cd/m\textsuperscript{2} background, the time used by Pokorny and Smith\textsuperscript{7} to adapt participants to the background. This final 20-second period served as a control to which earlier thresholds could be compared: if thresholds stabilized to this control level before 2 minutes, this would provide evidence that 2 minutes was not required for full adaptation to be obtained. This final 20-second period also served as a control for Condition 3.

Condition 3 was designed to investigate the adaptation time to, and recovery from, a steady luminous pedestal. Participants first adapted to a 30 cd/m\textsuperscript{2} blank screen, after which four steady 60 cd/m\textsuperscript{2} luminous pedestals appeared. Increment thresholds upon these pedestals were measured every 2 seconds over an 80-second period. The pedestals then disappeared, and increment thresholds measured every 2 seconds until 140 seconds. The final 20 seconds that the pedestals remained on (excluding the trial at 80 seconds, Fig. 1) was treated as a control where adaptation to the pedestal was judged to be complete, based on the 1-minute adaptation criterion used by...
Pokorny and Smith. The time after the pedestal appeared for thresholds to reach this control then could be determined, which would then give the minimum time required to adapt to the pedestal. Similarly, we measured the time required for thresholds to recover after the pedestal was turned off by determining how long it took for thresholds to return to the control values determined in Conditions 1 and 2.

To allow us to measure thresholds at closely spaced intervals, we used a method of a thousand ZESTs. In this method, participant performed 20 runs of an experimental condition with a separate ZEST threshold algorithm operating between, rather than within, runs at each time-point. All stimuli in the first run were presented at a predetermined level, therefore, with each subsequent run gradually converging on the participant’s threshold based on the participant’s responses in earlier runs. The average intensity of the initially presented stimuli (i.e., run 1) was 1 cd/m², with a uniformly distributed randomization ±0.5 log units applied to avoid long stretches of stimuli all appearing either below or above threshold. A further randomization of ±0.1 log units was applied to the optimum stimulus intensity returned from the ZEST procedure: while this reduces efficiency slightly, it makes the task more comfortable by avoiding every stimulus being placed at threshold, particular in latter runs were our ZEST procedures has largely converged. Each ZEST used a flat prior probability density function of 4 log unit range, centered on the initial threshold guess and encoded with a resolution of 0.01 log units. The slope of the assumed psychometric function was 3.5, with false alarm and lapsing probabilities of 0.25 and 0.01, respectively. Each ZEST was set to converge at the 84% correct level.

Participants

Four observers (aged 27, 28, 27, and 40; observers A–D) participated in the experiment, with spectacle corrections worn as required. Observers C and D also were authors. The study complied with the tenets of the Declaration of Helsinki, and was approved by the authors’ institutional ethics committee, with participants giving informed consent before participating.

Results

Figure 2 shows the results for Conditions 1 (preadaptation = 3 cd/m², open symbols) and 2
(preadaptation = 90 cd/m$^2$, closed symbols), with horizontal dashed lines giving ±2 SD limits around the mean threshold when adaptation is presumed to be complete (from 120 seconds onwards). Despite a 1.5 log unit variation in the preadaptation luminance, thresholds for both conditions fell within these limits for the bulk of measurements after 4 seconds into the run. Participant C showed the greatest number of thresholds falling outside these limits (14 for conditions 1 and 2 combined), with a tendency for threshold early in the run to be towards the upper boundary of the ±2 SD limits. The 95% confidence interval for the exponent of an exponential decay fitted to these data encompasses zero (i.e., a horizontally flat line); however, (GraphPad Prism 4; GraphPad Software, Inc., San Diego, CA).

Figure 3 shows the results for Condition 3 (steady pedestal of 60 cd/m$^2$ for 80 seconds). Leftmost dashed lines give ±2 SD limits around the mean for thresholds after 60 seconds adaptation to the pedestal. For all participants, thresholds are within these limits from a few seconds after the onset of the pedestal (time = 0). For all participants there is a prominent (0.5–1.4 log unit) elevation of thresholds coinciding with the pedestal offset (t = 80 seconds), after which thresholds immediately fall within the limits denoting complete adaptation to the background, as determined in Conditions 1 and 2 (rightmost dashed lines).

Figure 4 shows the standard deviation of threshold estimates as a function of the number of presentations in the 4-AFC procedure. All participants showed an initially rapid decrease in standard deviation that progressively slowed, which could be approximated by decaying exponential functions (solid lines) whose parameters varied little between participants (see Fig. 4 legend). Simulations for the original ZEST procedure found a standard deviation of log thresholds of approximately 0.11 for the recommended 8-trial yes/no procedure, and our participants reached this level of variability after an average of approximately 14 trials.

**Discussion**

We found little evidence that extended adaptation periods to either the background or stimulus pedestals are required to implement the Pokorny and Smith method for assessing the function of the M and P
pathways. Increment thresholds reached stable levels for most of our participants within seconds, irrespective of whether they had previously adapted to a bright or dim uniform field (Fig. 2). In a clinical environment it may be difficult to tightly control an eye’s state of adaptation immediately before testing: participants may have come immediately from waiting within a lit hallway or have been waiting in the darkened test room for many minutes. Our results suggested that such adaptation differences are not likely to be of significance to measured thresholds, and that testing may commence without extended preadaptation to the screen luminance. Similarly, we found that an extended adaptation to the steady pedestals used to assess M function is not required (Fig. 3). This is of particular importance where investigators measure thresholds on a series of steady pedestals of differing magnitudes to determine the contrast gain characteristics of the M system (e.g., a series of seven pedestals used to investigate glaucoma10). By allowing an extended time for participants to adapt to each pedestal, waiting times can quickly accumulate. Equally, the length of the interstimulus interval is directly related to how quickly thresholds can be measured. The rapid recovery demonstrated after a pedestal (within 2 seconds, Fig. 3) suggests that the 3 seconds interstimulus interval sometimes used8,9 may be unnecessarily long, particularly given that interstimulus intervals in clinical perimetric tests are around half this time (e.g., < 1.5 seconds).22 By using an interstimulus interval of 2 seconds, the duration of a test run is decreased by approximately a third.

We found a prominent elevation of thresholds at the moment of pedestal offset (Fig. 3). This is consistent with the increased masking effect that luminous pedestals show at onset and offset, both for increment23 and flickering24 stimuli. That we failed to find similar increased masking effect at pedestal onset (with the possible exception of the initial datum for Participant D) likely reflects that the pedestal onset 2 seconds before our first measurement. The decay in the initial masking effect has been modelled previously using a decaying exponential24 with a half-life in the order of 100 ms for a pedestal of similar contrast to that investigated here. As such, one would expect thresholds to have largely stabilized over a 2-second period. The half-life increases as the pedestal increases, although is likely still to be less than a couple of

Figure 4. Standard deviation of threshold estimates as a function of the number of presentations. Trials from 120 seconds onwards were pooled across Conditions 1 and 2. Solid lines give best fitting decaying exponentials of the form $y = A.b^x + C$, where $A$ gives the magnitude of the decay, $b$ gives the rate of the decay, and $C$ gives the final asymptote. Parameters values were $A = 0.68, 0.73, 0.64,$ and $0.72$ (Participants A–D), $b = 0.85, 0.78, 0.77,$ and $0.78$, and $C = 0.05, 0.10, 0.11,$ and $0.08$.

Anderson et al.
hundred milliseconds for the typical combination of pedestals used on computer displays.\textsuperscript{8,9,11} The time-course of M and P sensitivity changes to light onsets and offsets also have been estimated, at least for relatively small pedestals: again, recovery was found to be in the order of hundreds of milliseconds.\textsuperscript{25}

Our study was performed on healthy observers, and so the presence of disease may act to alter adaptation timecourses from those found above. However, such adaptation changes would need to be marked to affect thresholds given how rapidly healthy observers adapt to backgrounds and steady pedestals. In diseases where profound adaptation changes are either known or suspected, pilot data comparing longer adaptation periods to the shorter periods recommended in the current paper could confirm whether shorter adaptation periods are appropriate: such testing could be confined to those conditions where adaptation times are likely to be longest (e.g., the most intense pedestal to be used in the study). Studies using staircase methods to determine thresholds would likely require even greater changes in adaptation time before measurements are affected, as thresholds estimates are based on response reversals that typically happen sometime into an experimental run. Indeed, it is not uncommon for the first reversals in a run to be excluded from the final threshold estimate,\textsuperscript{11} making such methods effectively immune to any small, early changes as a result of ongoing adaptation.

Regarding the ZEST procedure and test time, we found that approximately 14 trials were sufficient to reliably measure thresholds in our 4-AFC procedure (Fig. 4). This value is intermediate between the eight trials previously recommended for yes/no procedures\textsuperscript{16} and the 30 suggested for 2-AFC procedures,\textsuperscript{18} as predicted from the intermediate false alarm rate (25\%) associated with the 4-AFC task. The variability in our threshold estimates is comparable to that from longer staircase procedures, such as in Figure 4 of the study of Pokorny and Smith\textsuperscript{7} ($2 \times$ standard error of the mean limits of approximately 0.1 to 0.2 log units, equating to a standard deviation of approximately 0.1 to 0.2 log units given their averaging of four threshold estimates of approximately 40 to 50 trials each). It should be noted that our method’s good performance was obtained despite the use of a small amount of deliberately introduced jitter in stimulus intensity, which slightly decreases efficiency but improves participant comfort.\textsuperscript{20} Combined with an interstimulus interval of 2 seconds and no fixed preadaptation or stimulus adaptation times, our result would suggest that thresholds might be appropriately measured in approximately half a minute. Assuming that in an actual testing environment thresholds could be measured only once every 60 to 90 seconds once breaks and stimulus set-up are incorporated, this still would allow for 30 to 45 thresholds to be estimated over a 45-minute period. In contrast, Pokorny\textsuperscript{6} has suggested that only 10 to 15 thresholds might realistically be measured in the same period using the conventional version of his method.

In summary, we showed that preadaptation and stimulus adaptation times may be dramatically reduced from those described by Pokorny and Smith,\textsuperscript{7} without altering thresholds. Further efficiencies also may be achieved by reducing the interstimulus interval, and using efficient threshold procedures, such as a 14 presentation 4-AFC procedure. Experimental time with clinical populations often is limited compared to laboratory-based investigations, and so increasing the efficiency of the Pokorny and Smith method allows the effects of disease on M and P functions to be either measured in a shorter test session, or to be explored over a more extensive range of experimental parameters.

**Acknowledgments**

Supported by Australian Research Council Future Fellowship FT120100407 (AJA) and FT130100338 (BVB).

Disclosures: A. J. Anderson, None; J. Jiao, None; B. V. Bui, None

**References**


