Temporal Integration at Extremely Low Contrast

Gabriel Jacob Diaz

Temporal Integration at Extremely Low Contrast

by

Gabriel J. Diaz

In partial fulfillment of the degree
Bachelor of Arts

Skidmore College
Psychology

Advisor

Flip Phillips

Contents

Temporal Integration at Extremely Low Contrast

The goal of this research is to investigate the relationship between the ability to recognize objects at very low contrasts and the duration for which they are presented. Empirical evidence suggests that duration plays a significant role in our ability to recognize objects in these impoverished conditions. To investigate this phenomenon, low contrast objects were presented to observers for extended durations of time and performance was assessed. The findings suggest that recognition complies with an extended duration power integration model of recognition. More specifically, our research suggests that the mechanism is integrating features over time and using them to hypothesis test against the pool of possible stimuli in order to narrow the pool and increase guessing accuracy.

1 Introduction

Figure 1.1 demonstrates a chart similar to those designed by Pelli and Robson for measuring contrast sensitivity. This chart depicts an array of equally spaced letters of equal size, incrementally decreasing in contrast. It illustrates (and indeed relies on) the relationship between contrast and detection - this is apparent in the fact that, as the contrast of the chart's letters is decreased detection becomes more difficult, and may eventually seem unattainable [1].

It is then, when one attempts to identify the lowest-contrast letters, that an interesting and unintended phenomenon arises. While at first these letters do not appear to be 'there' at all, over an extended period of time (roughly ten to twenty seconds) these low contrast letters gradually fade into appearance. The process evident over this extremely long time course seems to play a critical role in mediating our ability to discriminate at low contrasts.

By comparing behavior with past models of recognition, we may be able to speculate on the mechanism's behavior. The *power integrator* model of temporal recognition suggests that information accumulates and continuously offers increasing levels of insight concerning low contrast information [3]. Alternatively, the observers might be operating in concordance with the all-or-none *probability summation* model [3] [4]. This would suggest that the mechanism does not detect objects until a minimum amount of contrast is present at any point in time.

The power integrator model is a single channel theory of temporal recognition and integration. Single channel theories presume that a single agent mediates the information gathered within a finite interval. This integrating agent acts as both a detector and a repository for the information. Additivity

Figure 1.1 A contrast sensitivity chart similar to those created by Pelli and Robson [2]. Each consecutive group of three letters is at half the contrast of the preceding group. This chart provides the empirical evidence of the mechanism that is responsible for the perception of low contrast objects over extended durations (10s - 20s)

implies that any signal detected contributes toward the eventual percept. The end result is an integration of information over a given duration, resulting in a cumulative detection.

Different values have been suggested for the duration for which the power integrator can integrate information over time. Rashbass [5] proposes an integration that spans ±100ms, while Koenderink and Van Doorn [6] propose a span of ±250ms. Based upon the subjective experience of gradual perception evident while using the contrast sensitivity chart, the long-time course recognition in question lasts considerably longer than a few hundred milliseconds.

If the mechanism is integrating over this time-course, what information is being integrated? The process of temporal accumulation of contrast information may be using a feature integration process. Treisman's feature integration theory implies that the mechanism scans the stimulus for basic units such as areas of curvature, lines, and junctions, and then uses the cumulative percept as a basis for comparison with objects stored in memory [7].

However, if there seems to be no relation between the duration for which the stimulus is presented and the observer's ability to detect the stimulus, the mechanism may be acting in a way that is congruous with the all-or-none theory of probability summation. Unlike the single channel power integrator model, the probability summation model uses multiple channels/agents to evaluate the information. Each channel is attuned to contrast information within a specific range of spatial frequencies. When the combination of contrast information and noise excites a detection channel past a certain point (threshold contrast sensitivity) the observer has detected a stimulus. This highly variable noise provides the uncertainty responsible for the variance within the model [3][4]. Unlike the power integrator model, probability summation does not integrate information over time, but instead implies that the probability that this combination of contrast and variable noise will pass the observer's threshold increases as the duration of exposure increases.

Our experiments attempt to disambiguate the behavior of this mechanism over time by investigating the relationship between the exposure time of a low contrast stimuli and the resulting contrast sensitivity. Observer performance is expected to be in concordance with the power integrator model of recognition. If it conforms to the power integrator model of recognition then we will explore the possibility that the observer is integrating features over time. If it does not conform, then we will explore the possibility that it operates in a way similar to a probability summation model of recognition over time.

2 Experiment 1: Finding Threshold

These experiments will determine the behavior of the mechanism responsible for the temporal recognition of objects at a variety of low contrasts with their surroundings. Experiment 1 will examine the mechanism's behavior over a duration of up to 16 seconds in order to find the average observer threshold.

Observers

Participants included the author and two additional laboratory members. All observers had normal or corrected-to-normal vision, and were aware of the purpose of the experiment.

Stimuli

Each stimulus depicts one individual letter at low contrast. The level of contrast was varied between trials.

It is important to understand that letters were used independent of their literary utility, but because they are objects that all observers will have become familiar with in their daily lives. By using letters as familiar objects we can be sure that object recognition accuracy will not be lowered as a result of unfamiliarity with the objects, and that data reflected by the experiment will be unbiased by individual familiarity levels. We can also assume that, because of a universal high level of familiarity, no observer will be any more familiar with the letters than any other observer.

Stimuli were constructed using Sloan letters [8] because they maintain a stable global contrast (the ratio of black to white) across letters, ensuring that

Figure 2.1 Each stimulus featured one of ten Sloan letters [8] at low contrast. Sloan letters were used because they maintain a constant ratio of black to white pixels across characters.

no one character contains any more contrast information than any other (see **Figure 2.1**).

The printing process necessary to create paper stimuli similar to Pelli and Robson's contrast-sensitivity chart [2] is too expensive and time-consuming in comparison with the alternative method of using computer generated low contrast stimuli. There is a slight problem, however: even the best modern cathode ray tube computer monitors cannot display the tiny fluctuations in contrast necessary for this experiment. Pelli and Farell [9] developed a method of simulating this low contrast.

Individual frames of uniform distributions of random noise, which yield a constant base level of luminance, were darkened in the areas that would depict the letter in the resulting stimuli. This method produces a consistent level of contrast when presented as a movie and integrated over brief periods of time (see **Figure 2.2**).

The general idea behind the use of noise is that it introduces a form of malleable intrinsic uncertainty to the neuronal process of recognition. When it is integrated over time it yields a consistent level of luminance. Consistent with low contrast stimuli, this is done without activating edge detectors on the retina, a feature afforded through the ambiguity offered by the noise. One result of this is an inability for observers to identify these low levels of contrast from the information in any individual frame. Only when the frames are played in sequence for a brief duration are the observers are able to detect the contrast along an edge (see **Figure 2.2**). When these stimuli are used in conjunction with the moderately low contrasts that can be achieved by a CRT monitor,

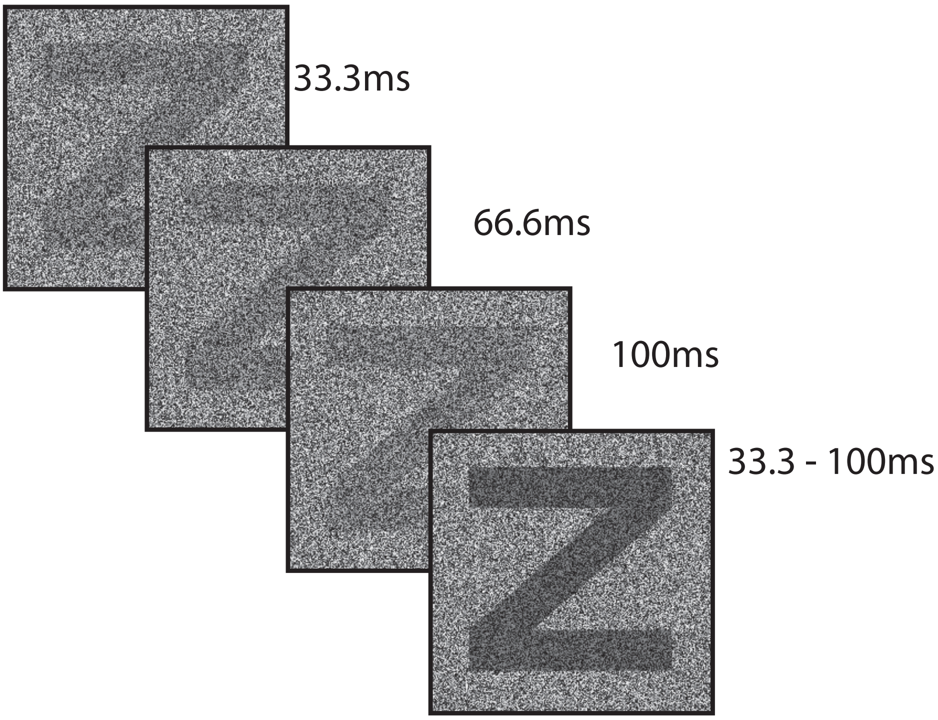

Figure 2.2 An demonstration of the perceptual process underlying the temporal integration of low contrast information to produce the letter percept. Individual frames of uniformly distributed noise displayed at 30 frames per second provide a base level of luminance. This base level of luminance was lowered in areas where there was overlap with the letter depicted in the resulting stimuli.

the resulting percept is equivalent to paper stimuli similar to those found on a contrast sensitivity chart (see **Figure 1.1**).

The dynamic stimuli are movies comprised of individual frames played at 30 frames per second. Each frame is a combination of some base level of randomized pixel-scale resolution noise and some amount of signal (the embedded letter). When the individual frames are shown rapidly they produce a constant level of contrast.

The stimuli are created in Mathematica version 5.0 on a Macintosh G5 (see **Figure 2.2**). Random numbers from -1 to 1 are assigned to a grid of 256 by 256 cells of pixel-scale resolution. A continuous grayscale grid is mapped onto each number from -1 to 1, -1 being black, 0 being gray, and 1 being white. This produces a 256 cell grid comprised of random and quantifiable grayscale values. When presented in sequence as dynamic noise, a constant level of luminance is maintained.

This grayscale grid is then darkened where the letter will be depicted in the final stimulus to produce some constant level of contrast with the surrounding noise. The cells overlapping the area onto which the letter will be embedded are systematically darkened by a constant contrast value decided upon before the trial.

New Cell Value = Noise Luminance+((Contrast Value)(Current Cell Value)).

The resulting level of contrast is the proportional difference between the luminance values of the darkened letter area and the background. The stimuli used exhibited contrasts ratios of .005, .00354, .00250, .00177, .00125, and .00088.

Procedure

A single trial of the experiment presented the observer with the stimulus for a given duration, after which point they were asked to identify which letter they were just presented with in a 10 alternative forced choice task (10AFC). Responses were recorded for later analysis. Each block held a constant duration of exposure of 2, 4, 8, or 16 seconds while an adaptive staircase was used to find threshold contrast.

Apparatus

Participants viewed the stimuli binocularly from a chinrest fixed at a 57 cm from the monitor such that the stimuli of a 256 pixels filled 7.2 degrees of visual angle.

The frames for the stimuli were created and using Mathematica 5.0 on a Macintosh G5, assembled into sequenced movies using Quicktime 6.5, and presented with Mathematica 5.0 on a Viewsonic PT795 monitor. The monitor was calibrated using a Colortron calibration device. A chin rest was utilized to steady the observer's head and maintain a constant viewing distance.

Results

Observer threshold contrast showed a gradual negative slope over the range of exposure times used, suggesting a temporal integration characteristic of the power integrator model of temporal recognition (see **Figure 2.3**). The continual decline of the shallow slope implies room for further improvement over a duration well beyond 16 seconds. This is a possibility that we explore in Experiment 2 by removing the time constraints in a similar 10 AFC task. The mean contrast threshold across all observers after 16 seconds of exposure time was 0.0262. This value will be the highest level of contrast presented in Experiment 2.

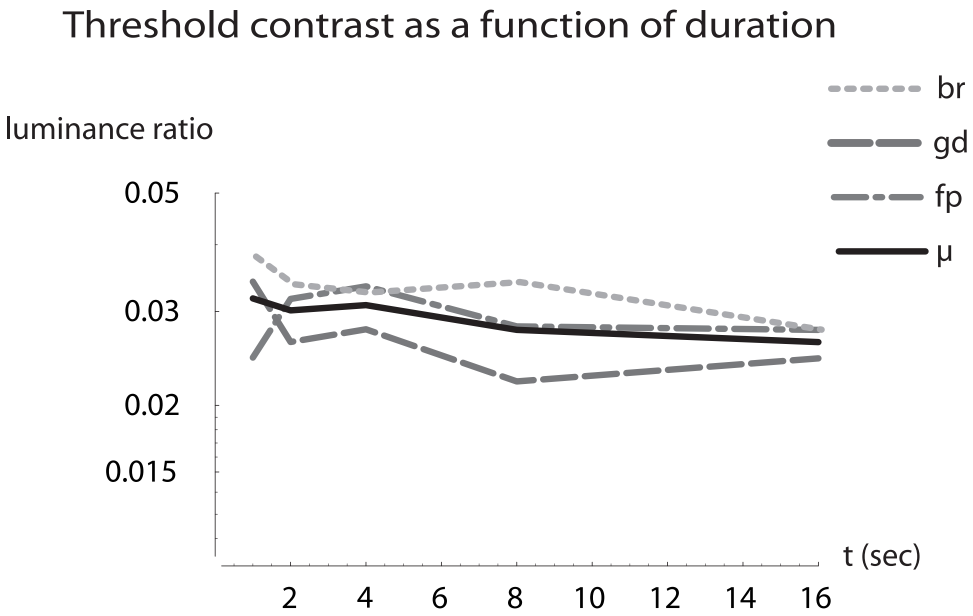

Figure 2.3 Experiment 1 Results. Observer thresholds decreased as the duration of exposure increased in a 10AFC task. This implies a relationship between duration and exposure time which is characteristic of the power integrator model of recognition.

3 EXPERIMENT 2: EXPLORING TEMPORAL BEHAVIOR

Experiment 1 examined observer contrast sensitivity over a 16 second duration using an adaptive method. Experiment 2 used the threshold exhibited in Experiment 1 after 16 seconds of exposure as the maximum contrast ratio presented to observers in a 10 alternative forced choice task. Unlike in Experiment 1, Experiment 2 was a free-viewing task (there were no time constraints) and no adaptive method was used.

Stimuli

The movies were identical to those used in Experiment 1, except only those depicting the lowest contrast ratios (.0025, .00177, .00125, and .0088) were used.

Observers

Participants included the same three observers from Experiment 1 and one additional laboratory member. All observers had normal or corrected-to-normal vision and were aware of the purpose of the experiment.

Procedure

The interaction between contrast and duration was stressed. Observers were instructed only to key in a response when they were confident in their choice, or to guess after at least 30s had passed without the presence of any perceived contrast information. Observers ran four blocks, each consisting of five trials at each of the four contrasts which were displayed in a random order. The letter presented, response, and time of response were collected for later analysis.

Apparatus

Experiment 2 utilized the same apparatus as Experiment 1.

Results

Observer accuracy was impressive. In Experiment 1, observer performance at a luminance ratio of 0.02 was well below threshold after 16 seconds of exposure time. In Experiment 2, performance at this level in a free viewing condition showed above 85% response accuracy.

Mean detectability even at the lowest contrast (.0008) was 21.2%. Despite the fact that observers reported not being able to detect a letter in the stimulus, the data suggests that, even at under extremely deprived viewing conditions, observers were able to utilize the low contrast information (see **Figure 3.1**).

Experiment 2 provides further evidence of the interaction between duration and contrast sensitivity. While Experiment 1 showed a threshold luminance ratio of .0262 after 16 seconds of exposure time, Experiment 2 showed contrasts below .015 to be above observer threshold in a free-viewing 10AFC task. This increase in contrast sensitivity over time suggests that this long time-course mechanism conforms to the power integrator model of recognition.

Looking times during trials using the same extremely low contrast stimuli suggest that the duration of presentation plays a large role in their ability to identify stimuli (see **Table 3.1**). The mean length of trials that produced correct responses (31.12 s) was significantly less than the mean length of trials that produced incorrect responses (45.77 s). With an average difference of 14.64 s, (p = .00002), observers clearly felt the need to attend to the lower contrast stimuli for longer durations of time. It is highly unlikely that observers were able to differentiate between these extremely low contrasts. This

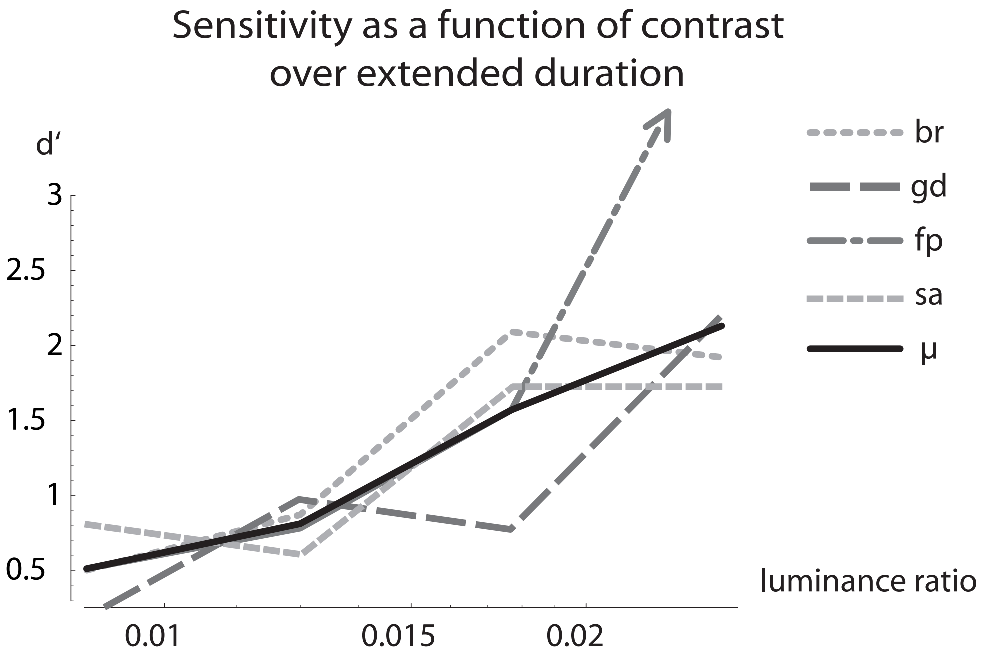

Figure 3.1 Stimulus detectability over a range of contrasts. Experiment 1 exhibited a threshold of 0.0262 after 16 seconds of exposure time. Observers were over 85% accurate at the same luminance ratio in a free-viewing version of the same 10AFC task.

preference to look at stimuli that are difficult to recognize for longer periods of time could be explained as a need to resolve incomplete feature integration.

Correct	Mean	43791.5	36818.1	29900.5	26320.6
	Variance	28366.2	18490.0	24916.0	25776.3
Incorrect	Mean	54854.9	43235.5	29900.5	25320.6
	Variance	39514.4	27261.3	30092.8	29992.2

Table 3.1 In a 10AFC task at a luminance ratio of .0008, observers chose to look at trials longer when they led to an incorrect response.

4 General Discussion

Motivated by empirical exposure to what may possibly be a new mechanism for the temporal recognition of low contrast objects, this research is an exploration into the interaction between the duration of exposure to low contrast stimuli and their detectability. In doing so we sought to establish the mechanism as cohesive with either the power integrator model or the probability summation model. Our original hypothesis was that observers would illicit behavior similar to a power integrator model over extremely long time courses. Further, we believe that they are integrating features over time and using them to hypothesis test against the pool of possible stimuli in order to narrow the pool and increase guessing accuracy.

A series of tasks was used to investigate the relationship between exposure duration and contrast. Experiment 1 explored the relationship between time and detectability over the course of 16 seconds. Experiment 2 used the mean observer threshold from Experiment 1 as the ceiling contrast ratio in an exploration of the relationship between the detectability of extremely low contrast stimuli and the duration of presentation time in a free-viewing task.

The results from Experiment 1 show a slight increase in sensitivity as the exposure duration is increased (see **Figure 2.3**). The results from Experiment 2 provide further evidence of a relationship between exposure time and threshold contrast (see **Table 3.1**). In a free-viewing task, observer accuracy increased to over 85% accuracy when given unlimited exposure to stimuli that were below threshold in Experiment 1.

When viewing stimuli of the lowest contrasts in Experiment 2, observers spent a significantly longer amount of time looking at trials that produced incorrect responses than looking at trials that produced correct responses.

This preference is apparently mediated by the stimuli's contrast ratio, despite the fact that no observer reported having any insight into the contrast level presented, even on an ordinal level. This suggests that the mechanism does not conform to the multiple channel, single threshold probability summation model but to the single channel model of power integration.

If, as the data implies, the mechanism does indeed rely on power integration over time, what is being integrated? Could it be that the channel is integrating letter features in order to hypothesis test and narrow down the pool of possible alternatives?

In order to examine this possibility we performed a multidimensional scaling (MDS) of the stimuli (see **Figure 4.1**). The position of the letters is arbitrary, however the distance is symbolic of their similarity.

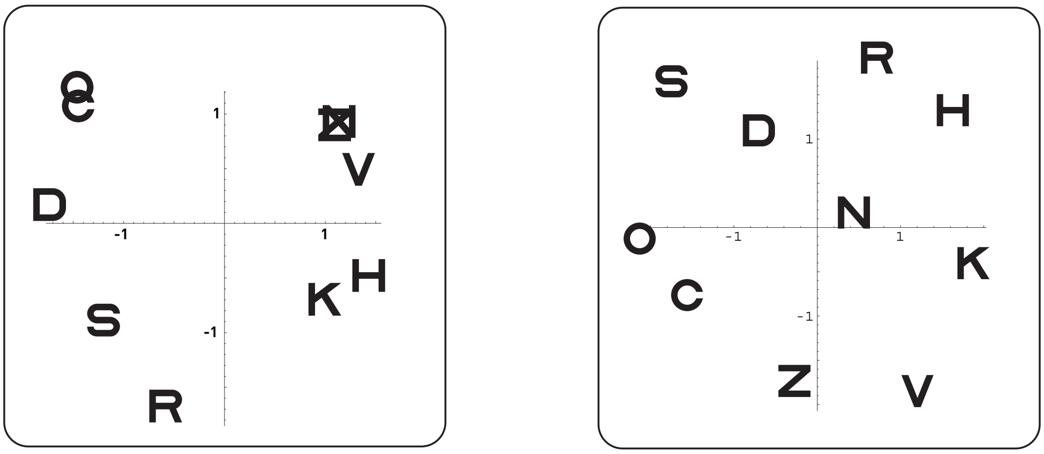

Figure 4.1 Multidimensional scalings of the stimuli created from naïve participants (left), and from the paired errors made at contrast levels of .0025, .00177, and .00125 (right). The proximity of similar letters suggests feature integration.

The MDS constructed from paired errors at luminance ratios .0025, .00177, and .00125 offers insight into what kind of errors were made at these extremely low contrasts. There are some trends suggestive of incomplete feature integration.

For instance, **O** and **C** are in extreme proximity, suggesting that one was often mistaken for the other when they were displayed at low contrasts. This makes sense, as their extreme curvature is unique from the other letters. The straight vertical line on the left side of letters such as **R**, **H**, **N**, **D**, and **K** also seems to have been a very salient feature.

Another MDS was constructed from data that was collected from participants who were unfamiliar with our research (see **Figure 4.1**). There are only a few disagreements between the two scalings, with the exception of most pairings that included the **Z** or the **N**. Naïve participants rated the two letters as highly similar, which is not a surprise considering the fact that they are rotated versions of one another. On the other hand, they were rarely mistaken for one another in Experiment 2. Given the context of the task, this is also not a surprise. The two letters had a very small amount of overlapping area, and shared very few salient features independent of their rotated states.

The similarities between the two scalings and the trends apparent in the scaling of paired errors made in Experiment 2 both suggest that errors made in Experiment 2 reflect incomplete feature recognition.

Our research suggests that a single channel temporal feature integrator mediates object recognition. This is in agreement with research by Solomon and Pelli [10] that finds letter identification to be mediated by a single channel. Subsequent research suggests that the channel used is completely decided upon by the stroke frequency of the letter set [11]. Stroke frequency was defined as “the average number of lines crossed by a slice through a letter, divided by the letter width.” Could the stroke frequency be mediating the channel used to detect the letters in our stimuli? A test of this would be to repeat the experiments using a letterset that differs from the Sloan letters in stroke frequency by some constant ratio. This ratio of difference should be reflected by the data, shifting thresholds up or down by some fraction of that constant.

Past research has suggested the ability for observers to show an adaptive increase in contrast sensitivity over time. Greenlee and Heitger [12] produced psychophysical evidence in support of an adaptive increase in contrast sensitivity. Observers were adapted to a high contrast image before discriminating between either low or high contrast test images. There was an improvement in the discrimination of high contrast images, while there was a detriment in the discrimination of low contrast image signals. This improvement in performance supports Blakemore *et al's* [13] predictions that high contrast signals adapt observers for discrimination, and that this adaptation is probably due to neural fatigue of the high contrast detector mechanism and a resultant hypersensitivity to low contrasts [14].

It is unlikely that the increased sensitivity present when using a contrast sensitivity chart is a result of fatigued detectors. First of all, there is no prior exposure to stimuli of extremely high-contrast that might fatigue the observer's contrast receptors.

Still, if one were to repeat the experiment, it is important to keep in mind some considerations this that previous research raises that were not incorporated into the design of this research. It is possible that the gradual increase in sensitivity took place between trials as well as within them. In Experiment 1, as the length of the total block increased so did the opportunity for this adaptation to bias the data. This may be another possible explanation for the graph's gentle slope, though it is unlikely.

There may still be some similar kind of neural adaptation similar to that found by Greenlee and Heitger [12]. A brief empirical test supports the presence of adaptive increases to contrast sensitivity. If the fading-in phenomenon works similarly to a fatigued system for high-contrast detection we should instantaneously be able to recognize all the letters above the contrast that we are currently adapted to. This is apparent when an observer, without explicit knowledge of the letters used in the chart, attempts to recognize the letters starting from the lowest contrast first, giving each letter enough attention that they might adapt before moving on. While the experience is subjective, this seems to be the case, and may be another valid path for future research.

References

1 Campbell, F. and Robson, J. (1968). Application of fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197:551-556.

2 Pelli, D., Robson, J. and Wilkins, A. (1987). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 2(3):187-199.

3 Watson, A. B. (1979). Probability summation over time. *Vision Research*, 19:515-522.

4 Quick, R. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, 16:65-67.

5 Rashbass, C. (1970). The visibility of transient changes of luminance. *The Journal of Physiology*, 210:165-186.

6 Koenderink, J. and Van Doorn, A. J. (1978). Detectability of power fluctuations of temporal visual noise. *Vision Research*, 18:191-195.

7 Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychologist*, 47:862-875.

8 Sloan, L. L. (1959). New test charts for the measurement of visual acuity at far and near distances. *American Journal of Opthalmololgy*, pages 807-813.

9 Pelli, D. and Farell, B. (1999). Why use noise?. *J. Opt Soc. Am.*, 369:647-653.

10 Solomon, J. A. and Pelli, D. (1994). The visual filter mediating letter recognition. *Nature*, 369(6479):395-396.

11 Majaj, N. J., Pelli, D. G., Kurshan, P. and Palomares, M. (2002). The role of spatial frequency channels in letter identificiation. *Vision Research*, 42:1165-1184.

12 Greenlee, M. and Heitger, F. (1988). Functional role of contrast adaptation. *Vision Research*, 28:791-797.

13 Blakemore, C., Muncey, J. and Ridley, R. (1971). Perceptual fading of a stabilized cortical image. *Nature*, 233:204-205.

14 Blakemore, C., Muncey, J. and Ridley, R. (1973). Stimulus specificity in the human visual system. *Vision Research*, 13:1915-1931.