INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.
2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International

300 N. Zeeb Road
Ann Arbor, MI 48106

8401935

Herskovic, Joseph Elliott

RATING SCALE AND RESPONSE TIME ANALYSES OF VISUAL FLICKER IN DEPRESSED PATIENTS AND NORMAL SUBJECTS

City University of New York PH.D. 1983

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

RATING SCALE AND RESPONSE TIME ANALYSES OF VISUAL FLICKER IN DEPRESSED PATIENTS AND NORMAL SUBJECTS

by

JOSEPH E. HERSKOVIC

A dissertation submitted to the Graduate
Faculty in Psychology in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy, The City University
of New York

1983

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 28, 1983
Date

Mitchell L. Kietzman
Chairman of Examining Committee

August 15 1983
Date

Herbert D. Saltzstein
Executive Officer

Prof. Samuel Sutton
Prof. Sandra Shapiro
Prof. Stanley Novak
Prof. Frederic Levine
Supervisory Committee

The City University of New York

Abstract

RATING SCALE AND RESPONSE TIME ANALYSES OF VISUAL FLICKER IN DEPRESSED PATIENTS AND NORMAL SUBJECTS

by

Joseph E. Herskovic

Adviser: Professor Mitchell L. Kietzman

Using signal detection methods to obtain separate measures of sensory sensitivity and response criterion, 11 major depressive patients, 5 dysthymic patients, and 9 normal control subjects were tested on a visual flicker discrimination task. It was hypothesized that major depressive patients would have the lowest flicker sensitivity values, dysthymic patients would have higher values, and normal control subjects would have the highest values. In addition, it was hypothesized that major depressive patients would respond significantly more conservatively than either the dysthymic patients or the normal subjects.

A second purpose was to compare rating scale and response time procedures. It was hypothesized that these procedures would yield similar conclusions with respect to visual flicker performance and depression.

The subject's task was to discriminate between a flickering light (16 Hz) and a fused light (116 Hz) which were equated for duration, apparent brightness, and probability of occurrence. Subjects were instructed to respond to each trial by saying "flicker" or "fused" and to rate the confidence of their response by saying "positive", "fairly sure", or "guess". In addition, response times were measured from the onset of the stimulus to a finger lift response which subjects were instructed to make as soon as they decided whether the stimulus was flickering or fused. Rating scale receiver operating characteristic (ROC) curves were generated from the confidence ratings. Response time ROC curves were generated in an analogous manner by assuming that faster response times meant greater response confidence.

The results indicate that, in general, major depressive patients respond more conservatively than either dysthymic patients or normal subjects. The results also indicate that these subject groups do *not* significantly differ on flicker sensitivity.

The major conclusions are that previously reported visual flicker differences between depressed patients and normal subjects were probably due to the different decision-making strategies used by the patients and not due to flicker sensitivity differences between the groups and that the rating scale and response time procedures yield similar conclusions with respect to visual flicker sensitivity and response criterion.

ACKNOWLEDGMENTS

First, I wish to thank my sponsor and friend, Professor Mitchell L. Kietzman with whom I spent countless hours discussing this research. I also thank Prof. Kietzman for inspiring me to continue at those times when I thought I would never finish.

Special thanks are due to Dr. Samuel Sutton, my "co-sponsor", who read, commented on, and returned early drafts of this thesis with unbelievable speed and who had much to do with the design and analysis of this study.

I am grateful to Drs. Sandra Shapiro, Stanley Novak, and Frederic Levine for serving on my committee and for their insightful and helpful comments.

The present study would not have been possible without the cooperation of the staff of the Depression Evaluation Service at the New York State Psychiatric Institute (NYSPI) especially Drs. Quitkin, Stewart, McGrath, Liebowitz, Schwartz, and Harrison; Nurses Elaine Tricamo and Lorraine Calev; and Ms. Martha Anderson and Ms. Marie Smith.

I thank the staff of the computer center at the NYSPI especially Henny Wolland, Murray Solomon, and Beverly Overton.

Thanks also to Nurse Beverly Nelson for assisting in the recruiting of normal control subjects and depressive inpatients.

I am indebted to the entire staff of the Department of Psychophysiology at the NYSPI: Drs. Stuart Goldberg, Kieth Cicerone, and Barbara Bienstock-Zachodin. Special thanks to Malvin Janal for his statistical help and to Steve Fairhurst and Amy Waring for their assistance in the preparation of the figures.

Dr. Stuart Steinhauer helped with preparing this thesis for final copy.

Many thanks to Ray Simon and Carl Regnat who designed and built the optical equipment (Mr. Simon) and the artificial pupils (Mr. Regnat).

I especially thank all of my subjects, particularly those who came back for the second session.

Finally, I thank my wife, Julie, for understanding why I needed to work 14 hrs. a day to finish this research.

This thesis is dedicated to the memory of my mother,
Roselyn Herskovic (1925-1977).

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

The Third Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III; American Psychiatric Association, 1980) defines the essential feature of a major depressive episode as "either a dysphoric mood, usually depression, or loss of interest or pleasure in all or almost all usual activities and pastimes" (p. 210). Other symptoms of depression may include appetite disturbance, change in weight, sleep disturbance, psychomotor agitation or retardation, decreased energy, feelings of worthlessness or guilt, difficulty concentrating or thinking, and thoughts of death or suicide, or suicidal attempts.

It is estimated that some 127 million or more people, including between 8 to 20 million Americans, currently suffer from one or another type of depressive disorder (Teuting, 1981). The more that is known about the etiology of this disorder, the better our chances of finding a rational and effective way of treating it.

One research strategy in the study of depression is the behavioral experimental approach. The goal of such a strategy is to identify objectively measured behavioral characteristics which are typical of and specific to the subtypes of depressive disorder. Ultimately the hope is that clinical, epidemiological, and experimental data will converge to help identify people who are suffering, or are likely to suffer, from a specific depressive disorder.

Although the experimental approach has had some success in the study of schizophrenia (Kietzman, Spring, & Zubin, 1980; Wynne, Cromwell, & Matthysse, 1978), thus far it has not been very fruitful in the study of depressive disorders. Miller (1975) reviewed the cognitive, motor, perceptual, and communication deficits in depression, and concluded that there is little evidence that these deficits are unique to depression. Also, the various subtypes of depression tended to exhibit similar deficits. Miller pointed out several problems with the research studies of depression: a lack of theoretical orientation; a failure to adequately define the sample of depressives tested; a failure to make direct comparisons of deficits in different subtypes of depressives; and a confounding of sensory sensitivity and response criterion factors making interpretation of deficits difficult.

Rather than abandon experimental research on depression, Miller (1975) suggested that future researchers address the problems mentioned above by developing theories of psychological deficits in depression, using the latest available research diagnostic methods, testing different subtypes of depressives in the same study, and using signal detection methods to separate sensory sensitivity from response criterion.

Domains of Information Processing

Kietzman, Spring, and Zubin (1980) discussed the relationship between psychopathology and three domains of information processing; sensory, perceptual, and cognitive.

Considered in a general way, sensory processing is defined as processing which occurs approximately within the first second after the presentation of a simple stimulus such as a light flash or an auditory click. Perceptual processing occurs within about the first five seconds after stimulus presentation. The stimuli for perceptual processing can entail a wide variety of stimuli such as geometric forms, dot configurations, or patterns of musical notes. Cognitive processing can involve complex stimuli such as words and its time course ranges from tens of seconds (short-term processing) to hours, days, or even years (long-term processing) after stimulus presentation. The present study concerns sensory processing and depression. It is assumed that faulty sensory processing can affect processing at the later stages and may partially explain the perceptual and cognitive deficits often reported in depressed patients (McAllister, 1981; Miller, 1975).

Sensory Processing and Depression

According to Johnson (1975), depression is a defense mechanism that enables a patient to cope with anxiety. Presumably, anxiety results from a defect in the patient's sensory processing mechanism, rendering the patient hypersensitive to environmental stimuli. (The notion of hypersensitivity to environmental stimuli is similar to Miller's (1964) information overload theory of psychopathology.) In some patients this hypersensitivity may manifest itself in a manic episode. In other patients depression results when, as a defense against anxiety, they invoke involuntary physiological "sleep" mechanisms and reduce their analysis of sensory information. These "sleep" mechanisms do not actually put patients to sleep (except in extreme cases) but do make them lethargic and drowsy and generally lacking in responsiveness. Thus, a sensory processing defect at one extreme (hypersensitivity to stimulation) results in a defensive reaction which takes the defect to the other extreme (hyposensitivity).

Johnson's (1975) theory offers an explanation for the oft-seen relationship between anxiety and depression (e.g., Anisman & Zacharko, 1982; Sartorius, Jablensky, Gulbinat, & Ernberg, 1980), the sleep disturbances characteristic of depression (by using the concept of patient's invoking of sleep mechanisms), manic-depressive cycling (depression is a response to the heightened sensory awareness of mania), and the effectiveness of lithium salts in the treatment of manic-depressive illness (lithium reduces or attenuates the processing of sensory information, see Johnson, 1979a, 1979b, and 1981 for a review of this possible action of lithium). The concept of sleep mechanisms to cope with hypersensitivity and anxiety provides a potential explanation as to why there might be a sensory processing deficit (hyposensitivity) in depressed patients.

There are many important issues that Johnson (1975) did not address. For example, he did not explain why a sensory processing mechanism defect is antecedent to and results in anxiety. What causes the sensory processing deficit in the first place? Also, why does hypersensitivity result in a manic episode in some patients and not in others? Why do these patients develop mania or depression but not one of the various anxiety disorders? It is also not clear that faulty sleep mechanisms and/or insomnia is a major symptom in all depressed patients. Finally, it is possible that the presumed sensory processing deficit in depressed patients is is not central to the depressive disorder but is only a consequence of various biological changes and/or social pressures associated with depression. However, the main value of Johnson's (1975) theory is that it provides testable hypotheses concerning the nature of sensory processing in depressed patients.

Nunn (1980) has formulated another perceptual model of the functional psychoses, (i.e., schizophrenia, hypomania, and depression) based on the findings of the type of information-processing deficits which characterize each of those groups. According to Nunn's model, information reaches higher brain centers in a series of packages. These packages are controlled by attentional processes and make up the flow of conscious experience. Nunn described four types of abnormalities which may occur to these information-carrying packages and how these abnormalities may lead to different types of psychotic disorders.

First, if the packages were split into fractions and some contained unchanging information, then this information would be constantly available to higher brain centers. The result of the intrusion of unchanging information into the flow of consciousness may lead to schizophrenia, a disorder whose symptoms can include intrusions of unwanted material into consciousness (see Frith, 1979, for a discussion of consciousness, information processing, and schizophrenia).

Second, if the amount of information in each package were increased above the normal range, then the higher centers would be bombarded by excessive sensory stimulation. The result of this abnormality would be hypomania. This notion is similar to Johnson's (1975) concept of a characteristic hypersensitivity in manic disorders.

Third, if the amount of information in each package were reduced below the normal range, then the higher centers would be understimulated resulting in a slowness of information processing. Nunn refers to this as "type 1 depression" and is similar to Johnson's (1975) concept of a characteristic hyposensitivity in depressive disorders.

Fourth, if all of the packages were to contain the same information, then the higher centers would be occupied by the same (usually unpleasant) thoughts. Nunn refers to this as "type 2 depression" which is related to the tendency of some depressive patients to constantly brood over feelings of guilt or worthlessness.

Nunn's (1980) model also addresses the issue of intermediate and mixed categories for patients who cannot be easily categorized into only one of the above disorders. For example, "if the unchanging information contained in a package was only slightly less than the normal quantity, a state intermediate between 'schizophrenia' and 'type 2 depression' would result" (Nunn, 1980, p. 76).

Nunn's (1980) model, like Johnson's (1975), leaves unanswered the question of whether a sensory processing deficit (if such a deficit truly exists) is a cause or a consequence of depression. Also, the relationship between Nunn's (1980) perceptual model and other models of depression (e.g., biochemical, cognitive, learning) was not stated. However, Nunn's model, like Johnson's, consists of testable hypotheses. The specific evidence of impaired sensory processing in depressed patients, which forms the basis of the models proposed by both Johnson (1975) and by Nunn (1980), is reviewed next.

Using a forced-choice staircase type of psychophysical procedure, Bruder, Sutton, Babkoff, Gurland, Yozawitz, and Fleiss (1975) reported that affective patients (five manic-depressive depressed, two manic-depressive manics, and one manic-depressive mixed) were less sensitive in detecting the presence of a transient monotic auditory signal delivered to the right ear than were schizophrenic patients (five paranoids, three schizoaffectives, one simple, and one unspecified), or non-patient control subjects. That finding was replicated in a later study (Bruder, Spring, Yozawitz, & Sutton, 1980). In addition, when the left ear was tested, affective patients were still less sensitive than other groups of subjects (the difference was not statistically significant), suggesting a possible lateralized sensory processing deficit in depression (Bruder et al., 1980).

In contrast to those findings, Bruder, Sutton, Berger-Gross, Quitkin, and Davies (1981) reported that bipolar depressives, unipolar depressives, and normal control subjects did not differ in the thresholds for detecting a single click in either ear. Bruder et al. (1981) discussed the differences between that study and the earlier studies in terms of the symptom of speech retardation. Speech retardation is defined as slowed speech, increased pauses before answering, low or monotonous speech, or a markedly decreased amount of speech (poverty of speech). The depressed patients in the earlier studies (1975, 1980) showed high symptom ratings of speech retardation whereas the patients in the later study (1981) displayed less evidence of that symptom. Indeed, speech retardation has been shown to be highly correlated with self-reports of increased unpleasant affect (Teasdale, Fogarty, & Williams, 1980). Thus, for monotic click detection, the symptom of speech retardation may be a more important predictor of performance than a diagnosis of a particular depressive subtype.

Malone and Hemsley (1977) used a signal detection theory analysis of a task involving the detection of a one second tone embedded in a six second burst of white noise. They demonstrated that depressed patients (the subtype was not mentioned) had reduced auditory sensitivity when they were depressed as compared to when they were in remission. Severity of depression was measured by the ratings by the patient's psychiatrist on the Beck Depression Inventory (Beck, Ward, Mendelson, Mack, & Erbaugh, 1966) and the Hamilton rating scale for depression (Hamilton, 1960).

Byrne (1977) reported that, in a vigilance task where the subject had to detect three odd digits in sequence from a background noise of random digits, psychotically depressed patients were able to detect an average of only 7.5 sequences (out of a possible 30) compared with a group of neurotically depressed patients and normal control subjects who were able to detect an average of 21.7 and 26.4 sequences, respectively.

In a visual detection threshold task using a forced-choice staircase procedure, Mannuzza, Spring, Gottlieb, and Kietzman (1980) found that schizophrenic and major depressive patients needed about twice as much stimulus energy as normal control subjects to detect a small white flash of light. The mean visual thresholds for the schizophrenic and depressive groups were not significantly different from each other.

The above studies are generally consistent with Johnson's (1975) and Nunn's (1980) prediction of a sensory processing deficit in depression, although this deficit may not be specific to depression (Mannuzza et al., 1980) nor correlated with diagnostic subtype of depression (Bruder et al., 1981).

Visual Flicker Studies of Sensory Processing

The above studies of sensory processing and depression were concerned with detection threshold which is only one type of sensory processing. Temporal resolution, or how stimulus events are perceived in time, is another type of sensory processing which can be used to study depression. One such measure of temporal resolution is visual flicker (or critical flicker-fusion frequency, CFF). Visual flicker is an objective behavioral measure in which an individual's ability to resolve successively presented pulses of light as discrete events is tested (Kietzman et al., 1980).

The literature on visual flicker is enormous. Bibliographies by Landis (1953) and Ginsberg (1970) each list well over 1,000 references. Visual flicker is interpreted as a useful, noninvasive index of the excitability of the nervous system (Douthwaite & Williams, 1974; Landis, 1954; Misiak, 1967; Pieron, 1965; Simonson & Brozek, 1952) and therefore has been used in a variety of studies of drug and chemical toxicity (e.g., Bobon, Ott, & Holmberg, 1982; Leigh, 1980; Smith & Misiak, 1976), aging (e.g., Falk & Kline, 1978; Misiak, 1967), neuropathology (e.g., Goldman, Lodge, Hammer, Semmes, & Mishkin, 1968; Koerner & Teuber, 1973; Medina, 1957; Parsons & Huse, 1958; Riklin, Levita, & Misiak, 1970; Titcombe & Willison, 1961), and psychopathology (discussed below).

The following studies are some that have used visual flicker as an index of cortical arousal:

Harper (1979) measured visual flicker sensitivity and response criterion as a function of white noise intensity (a cortical arouser) in three male graduate students. He found sawtooth-like changes in visual flicker sensitivity; i.e., the function peaked at 40, 70, and 100 dB (SPL) and troughed at 50 and 90 dB (SPL). Except for one subject who was extremely conservative at the 70 dB condition, different intensity levels of white noise had no effect on response criterion. Harper's (1979) results indicated a complicated interaction between white noise and a visual flicker measure of arousal.

Frith (1967) formulated a visual flicker theory in relation to introversion-extraversion personality characteristics based on an inverted-U model of arousal. Visual flicker performance was measured by a criterion-free four-alternative spatial forced-choice method under noisy (50 dB) and quiet (20 dB) conditions. The percentage of correct responses for the extraverts was 58.56 (noisy condition) and 52.64 (quiet condition); for the introverts it was 53.16 (noisy condition) and 53.12 (quiet condition). Analysis of variance showed a significant main effect for condition and a significant condition by personality interaction, thus confirming the theory's prediction that an increase in noise would improve the visual flicker performance of extraverts more than that of introverts.

Visual flicker has been used as a measure of stimulus persistence. Stimuli are said to persist when sensory processes continue to occur after cessation of the physical stimulus which gave rise to them (Falk & Kline, 1978). If the individual light pulses which constitute the stimulus used in flicker research persist, then a person's flicker threshold is reduced. In a study of stimulus persistence and arousal, Falk and Kline (1978) tested two groups of adult volunteer subjects: a young group (mean age of 19.1 years) and an old group (mean age of 70.1 years). Half of the members of each group were female. Visual flicker was determined using the psychophysical method of adjustment and arousal was measured by visual flicker and by skin conductance level, two measures which have been shown to be significantly correlated (Maley, 1967). Falk and Kline found a complicated interaction between visual flicker, age, sex, and arousal and postulated that the increase in stimulus persistence with age supported an overarousal theory of age related persistence performance.

In a visual flicker study which did not manipulate or measure arousal level, Bross, Harper, and Sicz (1980) reported a decrease and then an increase in flicker performance during 24 hours of auditory deprivation. The initial decrease in performance was explained by a decrease in arousal because, due to the auditory deprivation, less sensory information was being transmitted to the brainstem ascending reticular activating system (an important group of neurons involved in arousal). The subsequent increase in flicker performance was attributed to a denervation supersensitivity of the neural fibers from the auditory system to the reticular activating system which arises when neural pathways are not used.

The finding that fatigue and repetitive tasks reduce visual flicker performance also shows flicker's utility as an indicator of arousal (Grandjean, Baschera, Martin, & Weber, 1977; Weber, Fussler, O'Hanlon, Gierer, & Grandjean, 1980).

In addition to being a noninvasive index of the excitability of the nervous system and a measure of arousal and stimulus persistence, the following studies support the concept that visual flicker performance and anxiety level are negatively correlated:

Perris and d'Elia (1971) found that flicker threshold (descending method of limits) but not fusion threshold (ascending method of limits) was significantly negatively correlated with anxiety (as measured by the

Taylor Manifest Anxiety Scale) in a group of alcoholics. The reason why the correlation between fusion threshold and anxiety (though negative) was not statistically significant is unclear.

Maley (1967) reported a significant negative correlation ($\underline{r}$ (19) = -0.47, $\underline{p} < .05$) between visual flicker threshold and a range-corrected measure of skin conductance in a group of applicants for admission to the psychiatric ward of a general hospital. Flicker thresholds were obtained using an adaptive staircase procedure and skin conductance was measured from a constant voltage electrode placed on the subject's fingers. Skin conductance is a physiological response which is correlated with autonomic nervous system activity and is presumed to be a measure of arousal (e.g., Duffy & Lacey, 1946).

Using a modified method of limits procedure to determine the flicker threshold, Wagoner (1960) reported that outpatients suffering from high anxiety displayed significantly poorer flicker performance (i.e., had lower flicker thresholds) than low anxiety relatives of outpatients.

In a study of the relationship between stress and visual flicker, Stewart, Groff, and Kiker (1971) administered random surprise shocks to a group of college students while simultaneously measuring visual flicker by the psychophysical method of limits. Comparison of pre-shock, shock, and post-shock trials indicated that shock, a stressor, decreased visual flicker performance.

Buhler (1961) determined visual flicker thresholds of 30 patients before and after minor surgery using the ascending method of limits (flicker to steady). The assumption was that surgical patients are more anxious preoperatively than postoperatively. To control for possible practice effects a group of 30 hospital employees was also tested and retested about a week later. Visual flicker performance was found to be superior postoperatively as compared to preoperatively. The control group showed no significant differences in visual flicker from one session to another. Buhler concluded that visual flicker is a sensitive, reliable, and valid index of anxiety level.

It is reasonable to assume that the visual flicker measure is a potentially useful one to help differentiate depressed (or anxious) from nondepressed (or nonanxiuos) people in an objective, reliable way. For example, in a review of the peripheral indices of depressive states (Christie, Little, and Gordon, 1980), stated that there is considerable evidence (though not unanimous) that retarded or endogenous depressive patients have reduced electrodermal activity indicating a hypoaroused state (see also Laux & Streichenwein, 1979). In addition, anxiety and tension appear to be among the most frequently reported symptoms in depressive patients (Sartorius et al., 1980). Johnson (1975) and Anisman and Zacharko (1982) have made anxiety and stress central to their theories of depression. Indeed, the distinction between anxiety states and depressive disorders is often a difficult one to make (Roth & Mountjoy, 1982).

Visual Flicker and Psychopathology

Studies using traditional psychophysical methods.

Results of flicker studies may differ for methodological reasons. Therefore, to better understand the results from flicker studies, it is necessary to describe in detail the methods used and to discuss the advantages and limitations of those methods.

To determine a visual flicker threshold using the descending method of limits, an experimenter would set the frequency well above the anticipated threshold (e.g., 80 Hz) and decrease the frequency in discrete steps until the subjects report seeing a flickering stimulus. To determine a visual flicker threshold using the ascending method of limits, an experimenter would set the frequency well below the anticipated threshold (e.g., 10 Hz) and increase the frequency in discrete steps until the subject reports seeing a fused stimulus. Typically, a number (e.g., 20) of ascending (A) and descending (D) determinations are presented in a counterbalanced order (e.g., ADDA, DAAD) and the visual flicker threshold would be defined as the average threshold of the ascending determinations plus the average threshold of the descending determinations divided by two (but see Perris, 1966, and Perris & d'Elia, 1971, for the advantages of analyzing the ascending and descending threshold determinations separately).

The psychophysical method of adjustment is identical to the ascending and descending method of limits except that the subject, rather than the experimenter, controls the frequency in an analogue fashion until the stimulus changes from flicker to fusion or from fusion to flicker.

The advantages of traditional psychophysical methods are their: consistency over time; easy administration; and the rapidity with which thresholds can be determined. The limitations of these methods are discussed after a review of some visual flicker studies of psychopathology which used traditional psychophysical methods.

1. Schizophrenia. Studies comparing visual flicker thresholds of schizophrenic patients with those of normal subjects or patients with neurological damage have yielded mixed results. Studies have found that schizophrenic patients have higher thresholds (McDonough, 1960), lower thresholds (Johannsen, Friedman, & Liccione, 1964), or are no different (King, 1962) than normal control subjects. One study found that the visual flicker threshold successfully distinguished schizophrenic from paretic patients (Irvine, 1954), another found the visual flicker performance of schizophrenic patients to be significantly better than that of organic patients (McDonough, 1960), while another study found no difference between the visual flicker thresholds of schizophrenic and organic patients (Watson, Thomas, Felling, & Andersen, 1969).

McDonough (1960) used an ascending method of limits to determine the visual flicker thresholds of 20 organic (multiple sclerosis, brain trauma, brain tumor, Huntington's Chorea, Parkinson's Disease, cerebral vascular accident, and intracranial infarction), 20 medical control (postoperative fractures, appendectomies, torn cartilage, hemorrhoids, and sacroilliac strain), 20 process schizophrenic, and 20 reactive schizophrenic patients. The mean visual flicker thresholds were 13.20, 16.38, 16.43, 16.67 Hz for the organic, normal, process schizophrenic, and reactive schizophrenic groups, respectively. Interestingly, the organic group differed significantly from each of the other three groups and both schizophrenic groups were more sensitive than the normal group, although statistically, only the difference between the reactive and normal groups was significant ($p < .05$).

Using the psychophysical method of adjustment, Johannsen et al. (1964) reported that chronically ill schizophrenics (those hospitalized for 11 to 20 years) exhibited significantly poorer flicker performance than a control group of hospitalized, non-psychiatric subjects. Even the significance of that result is tempered by the fact that Johannsen et al. (1964) did a number of comparisons but only found a few significant differences. They tested six groups of schizophrenic patients (classified on the basis of duration of hospitalization) at five stimulus intensity levels and only found significant differences between control subjects and two groups of chronically ill schizophrenics at one intensity level.

Other related studies using the method of limits revealed a negative correlation between flicker performance and severity of mental illness (Barry, 1962) or no difference between schizophrenic patients and normal subjects on visual flicker (King, 1962). In a study comparing visual flicker performance of schizophrenic to organic patients, Watson et al. (1969) found that visual flicker threshold did not differentiate the two groups but was instead a function of length of previous hospitalization regardless of diagnosis.

2. Depression. The research on visual flicker threshold in depression is not as extensive as that of schizophrenia or organic mental disorders. In some studies depressed patients were were combined with schizophrenic patients to form a "psychiatric" patient group (Barry, 1962; Clark, 1966; Clark, Rutschmann, Link, & Brown, 1963). Only two studies have investigated visual flicker in depressed patients (Black, Franklin, de Silva, & Wijewickrama, 1975; Perris, 1966).

Perris (1966) used fusion thresholds (ascending method of limits) and flicker thresholds (descending method of limits) for several reasons: (1) to try to differentiate different types of depressive patients; (2) to differentiate depressive patients from a normal control group; (3) and to see if differences in visual flicker between patient groups are due to relatively permanent characteristics of the subject (trait variables) or to the state of the illness (state variables). Perris tested 102 depressed patients diagnosed as either bipolar (depressed phase), unipolar, or reactive neurotic. In addition, 41 patients were examined during remission. The 30 control subjects were members of the hospital staff. Patients who had electroconvulsive therapy during the three months prior to testing or who were given psychotropic drugs within three days of testing were excluded. The diagnostic divisions in various groups were made on the basis of the patient's medical history.

Perris (1966) reported that the bipolar and unipolar patients, but not the reactive neurotic depressives, had significantly lower visual flicker thresholds, i.e., were less sensitive, than the normal control subjects. Also, during remission, only bipolar patients had a significantly lower flicker threshold than the other groups. During remission, the unipolar and reactive neurotic patients were not significantly different from each other or from the control group. Fusion thresholds did not differentiate any of the groups either in the depressed state or during remission.

Perris (1966) concluded that the flicker threshold differentiates psychotic (unipolar and bipolar) depresives from neurotic depressives, psychotic depressives from normal control subjects, and possibly, trait from state variables. He concluded that bipolar depression is probably more of a trait variable because that group's visual flicker performance did not improve in remission. The unipolar group's performance did improve, indicating that visual flicker threshold was more of a state variable related to the depressed mood in the unipolar group.

Black et al. (1975) determined the fusion thresholds of 76 patients who complained of depression and who were diagnosed as suffering from either depression or schizophrenia and 28 normal volunteers. They found that the normal control subjects were most sensitive (had higher fusion thresholds), followed by the schizophrenic patients, with the depressive patients being the least sensitive. The difference between the three groups was highly significant ($\underline{p} < 0.0005$) but the authors did not describe the statistical test used or the results of any post-hoc analysis. The difference between the schizophrenic and depressive patients was probably not due to differences in medications because they found, compared to a medication-free condition, six healthy normal volunteers were less sensitive to visual flicker when they were given chlorpromazine, which is widely prescribed in the treatment of schizophrenia. Also, four depressed patients demonstrated improved flicker performance after taking amitriptyline, an antidepressant. Thus, any effect of those medications on each group of patients probably would have narrowed, not accentuated, any difference between them on fusion threshold.

The conclusion of the Perris (1966) and Black et al. (1975) studies is that depressed patients have lower visual flicker thresholds than normal subjects. However, it is difficult to interpret those and other visual flicker studies of psychiatric patients for several reasons: in all of those studies, possible differences in pupil size were not controlled; the diagnosis of the patient groups was clinical and usually unsystematic; and the studies used traditional psychophysical methods which did not control, eliminate, or measure the effects of the willingness of the observer to respond "flicker" (i.e., they did not address the problem of possible response criterion differences between psychiatric patients and normal subjects). These problems are now discussed in detail.

Methodological Problems of Prior Flicker Studies

Controlling pupil size. Turner (1975), Smith and Misiak (1976), and Lawerence, McEwen, and Pidgen (1982) have all emphasized that the visual flicker threshold is affected by changes in pupil size. None of the studies mentioned above of visual flicker and psychopathology controlled for possible differences in pupil size among psychiatric patients, organic patients, and normal subjects. This is particularly disturbing given the small absolute flicker threshold differences found in some of those studies and the changes in pupil size which may have occurred due to any medications taken by the patients. An artificial pupil is a simple device which adequately controls for differences in dark-adapted pupil size among groups of subjects. Such a device was used in the present study.

Diagnosis. Most of the investigators who studied visual flicker and psychopathology diagnosed their patients on the basis of chart reviews or unstructured interviews. The criteria for a given diagnosis varied widely. Psychiatric diagnosis has made many advances since those studies were published. Semi-structured interviews such as the Schedule for Affective Disorders and Schizophrenia (SADS: Endicott & Spitzer, 1978) and specific criteria for the definition of mental disorders

(DSM-III: American Psychiatric Association, 1980) have made diagnosis more reliable (Andreasen, 1982). A comprehensive review of current diagnostic techniques and systems of classification used in research on depression is available (Andreasen, 1982). Although there are pitfalls in the diagnosis of patients with affective disorders (Charles, 1982), the system of DSM-III diagnosis based on a SADS interview used in the present study clearly defines the type of patients studied. Having diagnoses that are more reliable makes the present study easier to compare with future studies.

Response criterion and visual flicker. By using traditional psychophysical methods, those studies that reported visual flicker threshold differences between psychiatric patients and normal subjects may have confounded flicker sensitivity with response criterion factors. A subject who is reluctant to respond "flicker" very often will have a low (less sensitive) flicker threshold but this may not reflect the subject's true sensitivity to visual flicker.

Response criterion can be handled in three ways. Classically, criterion was controlled by using only well practiced observers (usually the experimenters served as subjects) who were trained to maintain a constant, usually strict criterion. Under these conditions, the classical psychophysical methods were appropriate because once the criterion was controlled, only the sensitivity remained. Secondly, criterion can be eliminated by using a forced-choice method (e.g., Clark, Brown, & Rutschmann, 1967; Clark et al., 1963). Finally, response criterion can be measured, and that is the value of signal detection theory (Clark, 1966). (There are many ways to measure response criterion (Dusoir, 1975) and in this review response criterion measures other than beta (McNicol, 1972) are discussed, where relevant.) It is possible that the decision making strategies of psychiatric patients are more important determiners of the visual flicker threshold than the patients temporal resolution sensitivity.

Clark et al. (1963) compared flicker thresholds obtained by the method of limits to thresholds obtained by spatial and temporal forced-choice methods in 48 psychiatric patients, 40 of whom were diagnosed as schizophrenic (the diagnosis of the other eight patients was not stated). In the spatial forced-choice method, subjects were instructed to choose which member of a set of four illuminated test patches was the flickering one. In the temporal forced-choice method, the four stimuli (three steady, one intermittent) appeared sequentially in time in a single spatial position and the subject's task was to identify which of the four temporal intervals contained the flickering pulse. The forced-choice method was found to be superior to the method of limits because the former yielded a much higher Day 1-Day 2 correlation (.96) than did the latter (.76). Also, the forced-choice method resulted in more sensitive flicker thresholds, indicating that it is not affected by such nonsensory factors as the subject having a stricter criterion (less willing to respond "flicker"), which has been observed in certain psychiatric patients (Clark, 1966). The spatial and temporal forced-choice methods afforded similar results, but the former was judged superior because four times as many observations could be collected in the same amount of time.

In a subsequent study, Clark et al. (1967) compared the flicker thresholds obtained by three psychophysical methods: spatial forced-choice; constant stimuli; and limits. They tested 26 psychiatric patients (22 schizophrenic patients and 4 neurotic patients) and 26 normal control subjects. To counterbalance for sequence effects, half of the observers were tested on the method of limits first and a concurrent forced-choice--constant stimulus method second (Order 1 group) and half in the opposite order (Order 2 group). They concluded that flicker threshold differences between psychiatric patients and normal control subjects could be attributed to nonsensory, rather than to sensory, differences in temporal resolution. Further, they reported that the spatial forced-choice method was superior to the other two methods because that method yielded significantly higher (more sensitive) flicker thresholds than the other methods under all but one experimental condition. The one exception was that, for the normal subjects, the Order 1 group method of limits threshold was slightly, but not significantly, higher than the forced-choice threshold. The results of the Clark et al. (1967) study confirmed the earlier finding that the forced-choice method is not as affected by nonsensory factors. Most importantly, Clark et al. (1967) reported that patients had significantly less sensitive flicker thresholds than normal subjects for the method of limits, but for the spatial forced-choice task, the thresholds of the two groups did not differ.

The Clark et al. (1967; 1963) studies indicated that the apparently less sensitive flicker thresholds of psychiatric patients might actually be due to nonsensory factors which are not controlled for when using traditional psychophysical methods. The forced-choice methods used in those studies eliminated the nonsensory factor of response criterion.

In an effort to measure their response criterion to flicker, Clark (1966) determined the thresholds of 16 psychiatric patients (mostly schizophrenic) with two different sets of instruction. In the facilitating instructions condition, the patients were told that most people see the light as flickering almost all the time. Inhibiting instructions consisted of telling the patients that most people do not report the light as flickering unless they are fairly certain. Compared to the inhibiting instructions, the facilitating instructions led to a greater proportion of "flicker" responses to both a physically intermittent light and to a continuous light of the same subjective brightness. Analysis of those data by the method of constant stimuli suggested a change in the flicker threshold as a function of the type of instruction. However, analysis of the same data by signal detection methods showed that the apparent change in flicker threshold was a function of change in subjects response criterion and not in flicker sensitivity. Unfortunately, Clark did not use a normal control group to assess if the patients were actually more conservative than nonpatients under both instruction conditions.

Further evidence that psychiatric patients and normal subjects may differ in response criterion was provided by a signal detection study by Clark and Mehl

(Note 1) who found that while psychiatric patients had lower sensitivity to thermal pain than did normal subjects, the patients were also less willing to report the stimulus as being painful (i.e., they were more conservative). Although Clark and Mehl (Note 1) reported that the patient group consisted of 31 schizophrenics, 24 affectives, and 9 others (not specified), separate analysis of pain sensitivity and response criterion was not done for each diagnostic subtype.

Response criterion and schizophrenia. The concept of differences in response criteria between schizophrenic patients and normal subjects is a complex one. For example, Rappaport, Hopkins, Silverman, and Hall (1972) concluded that paranoid, but not nonparanoid, schizophrenic patients employed a more conservative decision-making criteria than did normal subjects on an auditory signal detection task. The task involved the detection of a tone in a background of continuous white noise. Six different tone intensities were employed to produce six signal-to-noise conditions. Response criterion was defined as the sum of the probability of responding "yes" to an auditory stimulus (hits) plus the probability of responding "yes" when no stimulus was actually presented (false alarms; Rappaport et al., 1972). It was true that the paranoid schizophrenic group responded "yes" less often than either of the other two groups. However, it was also true that the paranoid schizophrenic group was less sensitive than either of the other groups (as measured by d' and the probability of detection). Both the paranoid and nonparanoid patients committed more false alarms than did the normal subjects. Those data suggest that response criterion and sensory sensitivity were not independent in the Rappaport et al. (1972) study and in fact, based on the false alarm data, it could be argued that paranoid and nonparanoid schizophrenic patients are less conservative than normal subjects. This was true of at least four of the six signal-to noise conditions. Therefore, it is concluded here that response criterion could not be accurately measured in the Rappaport et al. (1972) study and that their conclusion that paranoid schizophrenic patients were more conservative than normal subjects is not supported by their data. A more reasonable interpretation of the data of Rappaport et al. (1972) is that schizophrenic patients may employ different decision-making strategies than normal subjects. This interpretation is very tentative given the nonindependence of sensory sensitivity and response criterion as measured in that study.

Cegalis and Deptula (1981), who investigated detection of a target stimulus presented in the visual periphery at two display angles, reported that paranoid schizophrenic patients responded more conservatively than normal subjects. They used a nonparametric analogue of beta and found response criterion differences only at the more difficult display angle.

Contrary to the findings of the Cegalis and Deptula (1981) study, Gruzelier and Venables (1974) used a measure of visual temporal resolution and found schizophrenic patients (both paranoids and non-paranoids) generally had more lenient two-flash response criteria (as measured by an analogue of beta) than did normal subjects.

Price and Eriksen (1966) found that while paranoid schizophrenic patients were about equally accurate with normal subjects in judging whether a comparison stimulus was larger or smaller than a standard, they differed from the normal subjects in having much greater subjective confidence in their judgments. The paranoid patients also had greater confidence in their judgments than did non-paranoid patients. The non-paranoid patients and the normal subjects did not differ in judgment confidence. In that study, patients and subjects had to rate response confidence on a three point scale: positive, fairly sure, or guess. Judgment confidence was defined in terms of number of positive responses irrespectve of whether the trial contained a signal (comparison stimulus was larger than the standard stimulus) or noise (comparison stimulus was smaller than the standard stimulus). Thus, paranoid patients made significantly more positive responses than the other two groups.

The differences in the findings among the studies which investigated response criteria in schizophrenia could be due to differences in the experimental task, in the measure of response criterion, the diagnosis, or in other factors. The important point is that response criterion differences can exist between schizophrenic patients and normal subjects and these differences can affect visual flicker thresholds as determined by classical psychophysical methods.

Unfortunately, no studies of visual flicker in depression have measured response criterion. It is probable that some of the patients in the Clark (1966) and Clark et al. (1967; 1963) studies were depressed, but the patients were not subdivided into distinct diagnostic categories for data analysis. Studies which addressed the question of response criterion and depression are reviewed below.

Response criterion and depression. In a study of analgesia to painful stimuli in affective illness, Davis, Buchsbaum, and Bunney (1979) asked subjects to rate a shock administered to the left forearm as either noticeable, distinct, unpleasant, or very unpleasant. Response criterion was defined as the milliamperage level required for the subject to judge the stimulus as noxious. Those investigators interpreted their measures of response criterion as being nonparametric analogues of beta. In that study, 30 bipolar depressive, 36 unipolar depressive, 10 bipolar manic, and 48 normal control subjects were tested. Davis et al. (1979) reported that the depressed patients as a group rated significantly fewer stimuli as "unpleasant" and "very unpleasant" than did the normal subjects. This indicates that the depressed patients adopted a more conservative criterion for reporting pain. Further analysis indicated that this response criterion difference was primarily due to the unipolar male depressive patients. Bipolar depressive male patients and bipolar or unipolar depressive female patients were not significantly different from normal subjects.

Miller and Lewis (1977) tested recognition memory of simple geometric shapes in elderly depressive patients, elderly demented patients, and elderly control subjects.

In the depressed group, no attempt was made to classify patients into the various subtypes of depression. With respect to memory ablility, the depressed patients did not differ from the normal subjects on d' (the signal detection measure of sensory sensitivity) whereas the demented patients demonstrated significantly poorer memory ablility than the other groups. With respect to response criterion, the depressed patients were significantly more conservative (as measured by beta) than demented patients and normal subjects. The comparison between the demented patients and the normal subjects was not significant. In addition, the depressed patients made many fewer false alarms than the other groups, confirming an earlier study comparing depressed to demented patients on a series of verbal learning and memory tasks (Whitehead, 1973).

In the auditory vigilance study described above, Byrne (1977) reported that the mean false alarm rate for neurotic depressive patients was 31.6; for psychotic depressive patients the rate was 6.8; and for normal subjects it was 2.1. Byrne did not calculate any signal detection measures of sensitivity or response criterion, thus the suggestion that the psychotic depressives adopted a more conservative response criterion (as indicated by a low false alarm rate) than the neurotic depressives is made with caution. However, the low mean false alarm and mean detection rates (7.5 compared with 21.7 for the neurotic depressive patients and 26.4 for the normal subjects) shown by the psychotic depressives indicate their reluctance to report the occurrence of a signal. This reluctance could be due to a conservative response criterion or to a low sensitivity for signal detection, or both.

The psychophysical methods used by Perris (1966) and Black et al. (1975) in their studies of visual flicker and depression did not separate flicker sensitivity from response criterion factors. Given the evidence that depressed patients respond more conservatively than do normal subjects on a variety of psychological tasks, it is possible that the differences in visual flicker threshold between the depressed patients and normal subjects reported in previous studies were due differences in response criterion. To test this possibility, the present study was designed to separate visual flicker sensitivity from response criterion factors by obtaining signal detection measures of both flicker sensitivity (d'e) and response criterion (beta).

The Relationship Between Rating Scale and Response Time Signal Detection Procedures

Usually, the rating scale procedure of signal detection requires subjects to decide whether a given trial contained a stimulus (detection experiments) or contained a certain type of stimulus (discrimination experiments) and to rate the confidence of their decision on some scale (usually ranging from totally certain to totally uncertain). Rating scale receiver operating characteristic (ROC) curves can be generated by utilizing each response confidence category as a data point on a plot of responses to "signal" trials versus responses to "noise" trials. Signal detection measures of sensory sensitivity and response criterion can be derived directly from an ROC curve.

One problem with a rating scale procedure is that even under the best of circumstances subjects often fail to use all of the response categories, making subsequent data analysis difficult (McNicol, 1972). Another problem is that subjects must remember all the response categories and their meaning, a task which may be difficult for certain types of populations (e.g., children, mentally retarded patients). A third problem is that, at least in some situations, subjects may be required to give a meaningful verbal response, something that socially withdrawn or mute patients may be reluctant to do. An alternative measure is response time. This procedure requires subjects to respond as soon as they decide whether or not a certain stimulus was presented. The response is a simple motor movement, such as a finger lift or pressing a telegraph key. This measure assumes that the faster the response time, the greater the response confidence (Emmerich, Gray, Watson, & Tanis, 1972; Katz, 1970; Kerkhof, van der Schaaf, & Korving, 1980; Moss, Myers, & Filmore, 1970; Norman & Wickelgren, 1969). Compared to a rating scale procedure, data analysis in the response time procedure is usually facilitated because the latter procedure often yields more response categories (see <u>Signal Detection Analysis</u> in <u>Method</u> chapter).

Moss et al. (1970) generated rating scale and response time ROC curves in four undergraduate male subjects. The task required the observer to judge whether two temporally sequenced tones (interstimulus intervals were equal to 0.5, 2.0, or 8.0 sec) were the "same" or "different". Sensory sensitivity as measured by d'e (a parametric signal detection measure analagous to d') was similar for the two procedures across interstimulus interval conditions. In a similar study in which normal subjects were required to respond "same" or "different" to two lines which contained a gap (two levels of difficulty were used), Katz (1970) found generally similar results for the rating scale and response time procedures. The measure of sensitivity used in the Katz study was the distance of the ROC curve from the positive diagonal.

In an auditory detection study, Emmerich et al. (1972) compared ROC curves based on binary (yes-no), rating scale, and response time procedures and found that the the rating scale and response time procedures gave similar information and that both procedures yielded more information than the binary procedure. The measure of sensitivity used in that study was the area under the ROC curve (the nonparametric P(_A_) measure).

In a study of metacontrast phenomenon, Weintraub and Fidell (1979) compared the rating scale and response time procedures. They reported that the response time procedure was less sensitive than the rating scale procedure (i.e., yielded lower _z_ values, a measure related to d' with the population variance factored in) for two subjects, was more sensitive for one subject, and essentially equal for another subject. In a follow-up experiment the response time procedure was more sensitive for two subjects and less sensitive for three subjects. Sensitivity in the follow-up experiment was measure by P(_A_). However, because Weintraub and Fidell did not compare the rating scale and response time procedures over experimental conditions or over different groups of subjects, the two procedures could not be compared in a relative way.

Direct comparisons between the rating scale and response time procedures were made by Moss et al.(1970); by Katz (1970); and by Norman and Wickelgren (1969), who studied verbal short-term memory and used the distance of the ROC curve from the positive diagnonal to compare the two procedures on the variable of memory strength. These three studies reported that the rating scale procedure generally yielded higher sensitivity values than the response time procedure. However, the reduced sensitivity scores of the response time procedure did not alter any of their major conclusions.

The above studies strongly suggest that the rating scale and response time procedures lead to similar conclusions in a number of experimental tasks using a number of different measures of sensory sensitivity. Unfortunately, none of those studies directly compared the rating scale to the response time procedures with respect to measures of response criterion. Emmerich et al. (1972) manipulated criterion by instructing the subject to adopt a strict, medium, or lax criterion but they did not calculate beta (or its equivalent) directly from the rating scale or response time ROC curves. Moss et al. (1970) measured response criterion as the proportion of "different" responses and Katz (1970) measured it as the number of "most confident" responses. Both of those measures can be calculated without generating any type of an ROC curve. Norman and Wickelgren (1969) and Weintraub and Fidell (1979) did not measure response criterion. The similarity between the rating scale and response time measures of response criterion remains an unresolved issue.

Rationale and Hypotheses

The present study was designed to test Johnson's (1975) hypothesis of a sensory processing deficit in depression by replicating the visual flicker studies of Perris (1966) and Black et al. (1975). This study incorporated several refinements. It used artificial pupils to achieve optimal stimulus control. Further, signal detection theory methods were employed to obtain separate measures of flicker sensitivity and response criterion. It was hypothesized that major depressive patients would have the lowest sensory sensitivity (d'e) values, dysthymic (neurotic depressive) patients would have higher d'e values, and normal control subjects would have the highest d'e values to visual flicker. In addition, it was hypothesized that the major depressive patients would respond significantly more conservatively than either the dysthymic patients or the normal control subjects.

A second purpose of the present study was to compare the rating scale and response time procedures using two groups of psychiatric patients and a normal control group. It was hypothesized that these procedures would yield similar conclusions with respect to visual flicker performance and depression.

CHAPTER II

Method

Subjects

Pilot subjects. Seven colleagues and friends of the experimenter assisted in determining the stimulus level to be used in the main experiment. The pilot subjects ranged in age from 20 to 30 years (M = 25.4, SD = 3.7) and, with one exception, all were graduate students in psychology. These subjects were paid $5.00 per hour for their participation.

Recruitment and evaluation of depressed patients. Patients were recruited through the Depression Evaluation Service (DES) at the New York State Psychiatric Institute, an outpatient clinic which specializes in the diagnosis and treatment of depression in a research setting. The service advertizes on radio and in newspapers throughout the New York Metropolitan area. When a prospective client calls, he or she is questioned by a staff member on such items as age, referral source, reason for calling, and general medical health. If the person reports feeling depressed and does not have any major health problems, then an intake interview is scheduled with one of five staff psychiatrists.

During this interview the client is administered the Schedule for Affective Disorders and Schizophrenia (SADS: Endicott & Spitzer, 1978), the Hamilton Rating Scale for Depression (Hamilton, 1960), and given a diagnosis based on criteria from the DSM-III (American Psychiatric Association, 1980).

Each case is then discussed in detail at a meeting of the DES which includes psychiatrists, psychologists, social workers, and psychiatric nurses. Any questions concerning the diagnosis are discussed and, if necessary, the taped intake interview is rated by another service psychiatrist.

For the present study, once it had been determined that the client met the criteria for inclusion i.e., a DSM-III diagnosis of major depression, dysthymic disorder, or atypical depression, then he or she was asked by their physician to sign a consent form allowing this investigator to contact them to discuss participation in this study. The consent form was sent to this writer who remained unaware of the diagnosis or medication regime until after all of the subjects had been tested.

Note that each patient in the present study was diagnosed by only one psychiatrist. Therefore, only those patients about whom the psychiatrist was positive of a diagnosis were studied. Unfortunately, there is no measure of interrater reliability of diagnosis.

Originally, the study design called for patients to be tested twice: once before receiving any antidepressant medications and once after they had been on medication for at least four weeks. However, many patients refused to return for the second session. Therefore, the primary analyses for this study includes only those data collected in first session.

A further complication in the study was that, due to an administrative misunderstanding, 9 patients (out of 24) were taking antidepressants when tested for the first time (6 on phenelzine and 3 on imipramine). However, for a number of reasons, it is contended that this factor probably did not confound the results.

First, although the data for some patients who were tested before and after taking antidepressants showed interesting trends which possibly could be attributed to the medication (see <u>Results</u> and <u>Discussion</u> chapters), those patients were taking antidepressants for at least four weeks prior to testing. In addition, those patients who were retested were taking high ("clinical") dosages of either imipramine or phenelzine. In contrast, prior to the first testing session, most patients had not taken medication long enough or in large enough dosages to clinically benefit from it (see Hollister, 1980, for a review of drug treatment in affective disorders). The mean number of days on medication was 13 (range 7-29 days) and the mean dosages were 40 mg/day of phenelzine (range 30-90 mg/day) and 216.7 mg/day of imipramine (range 150-250 mg/day). All patients were clinically depressed at the time of the first testing (assessed by chart review).

Second, two studies (Holmberg, 1981; Karp & Pollack, 1963) have reported no significant effect of imipramine on critical flicker-fusion threshold (CFF). Holmberg (1981) tested healthy volunteer subjects at low (25 and 50 mg), acute dosages of imipramine. Compared to a pre-dose baseline condition, Holmberg (1981) reported no significant changes in CFF threshold up to six hours after imipramine administration. Karp and Pollack (1963) tested a group of hospitalized psychiatric patients (schizophrenics, depressives, and neurotics) after they had been taking 300 mg/day of imipramine for five weeks. They reported a statistically nonsignificant decline in CFF as compared with placebo. Both Holmberg (1981) and Karp and Pollack (1963) measured CFF thresholds using the psychophysical method of limits. The effect of phenelzine on CFF is unknown, although in the present study two patients improved in visual flicker discrimination sensitivity and two patients declined on that measure after having taken phenelzine for at least one month and having taken large (75-90 mg/day) amounts for one week prior to testing.

Third, an equal percentage of major depressive and dysthymic patients were tested on medication (39% and 40%, respectively).

Fourth, of those patients who performed above chance (see below), there was no significant difference between patients on and off medication on flicker sensitivity as measured by rating scale d'e ($\underline{t}$ (13) = 0.71, $\underline{p} > .05$) and response criterion as measured by log beta 3 ($\underline{t}$ (12) = 1.24 $\underline{p} > .05$). Therefore, the medication variable is ignored for all subsequent analyses.

Fifty-seven patients signed the consent form between June, 1981 and July, 1982. Of these, only 24 agreed to participate in the study. Eighteen of these patients were diagnosed as having major depression, 5 had dysthymic disorder, and 1 had atypical depression. With three exceptions (major depressive inpatients nos. 101, 103, and 109) all were outpatients. Since there was only one atypical depressive, and he performed below chance on the flicker discrimination task (see below), his data are not included in any analyses.

Due to hospital policy, patients were not paid for their participation.[1]

Recruitment and evaluation of normal control subjects. Normal control subjects were recruited by a variety of means. Subjects who had participated in other experiments at the New York State Psychiatric Institute (NYSPI) were contacted, hospital personnel (e.g., social workers, nurses) were notified, and friends of the experimenter were asked.

All prospective control subjects were administered a preliminary screening evaluation by the experimenter. This schedule, developed by Dr. Krooss-Glover (Note 2) consists of selected items of the SADS (Endicott & Spitzer, 1978) and is designed to eliminate from the study subjects with the following past or present problems: poor hearing or vision; brain damage or mental retardation; drug or alcohol addiction; depression or other mental disorder; and a family history of mental illness.

Three (of 18) prospective normal control subjects were not tested on the basis of the interview findings: one was clearly dependent on illicit drugs, one had received psychotherapy for depression in the past, and one was taking antidepressants at the time of the interview.

All normal control subjects were paid $5.00 per hour for their participation.

Elimination of subjects based on flicker discrimination performance. The data indicated that some subjects were unable to discriminate the 16 Hz test stimulus from the 116 Hz stimulus. Therefore, those subjects whose sensitivity index was at or below chance level (see Signal Detection Analysis) were eliminated from most analyses because their poor performance could have been due to a number of variables including low flicker sensitivity, fatigue, inattention, boredom, or lack of motivation.

Seven major depressive patients (39% of the major depressive sample) and 6 normal control subjects (40% of the normal sample) were eliminated because their flicker discrimination was at a chance level of performance. None of the dysthymic patients performed at a chance level.

A series of *t*-tests revealed that, with only one exception, there was no significant difference in age, education, or visual detection threshold between those subjects who did and did not perform at chance levels on the flicker discrimination task. The only exception was that the normals who performed at chance level had a significantly higher visual detection threshold than the normals who performed above chance (t (13) = 2.7, p < .05).

Subject samples used in most subsequent analyses. Demographic characteristics and log visual thresholds of all three subject groups (total N = 25) are shown in Table 1. A chi-square test (for the sex variable) and t-tests (for all other variables) showed no significant differences among the groups in sex, age, education, or log visual threshold (see Procedure section below). Ten subjects had normal vision (3 major depressives, 3 dysthymics, and 4 normal controls); two subjects wore contact lenses during testing (1 major depressive and 1 normal control); one normal control subject (no. 117) needed to wear his glasses during testing (because the target was too blurred without them); and the remaining 12 subjects were tested with their glasses off. With one exception (normal control subject #207 who was cross-eyed), all subjects who needed either glasses or contact lenses were myopic.

Table 1

Demographic Characteristics and Log Visual Thresholds of the Three Subject Groups

	Group		
Variable	Major Depressive (N=11)	Dysthymic (N=5)	Normal (N=9)
Age (years)			
Range	26-57	23-47	22-45
Mean	39.3	36.4	30.9
SD	10.97	9.29	7.47
Education (years)			
Range	8-20	12-18	13-18
Mean	14.9	14.8	16.1
SD	3.36	2.39	1.96
Sex (n)			
Male	3	4	6
Female	8	1	3
Log Visual Threshold (Trolands)			
Range	0.81-1.74	0.94-1.45	0.85-1.33
Mean	1.18	1.19	1.11
SD	0.254	0.183	0.147

Apparatus

Optical system. A single-channel optical system (a light source, collimating lens, filters, and a decollimating lens) was used to transilluminate the white-appearing circular target viewed by the observer.

The light source was a glow modulator gas-discharge tube (Sylvania R1131C) which was operated at a constant current of 23 milliamperes and activated with timing and gating circuits. It was irradiated by an argon ultraviolet lamp (General Electric AR-4) to provide short and stable ionization times of less than 10 microsec (see Matin, 1964). This resulted in the production of rectangular light pulses with very short rise (approximately 20 microsec) and decay (approximately 10 microsec) times (see Kietzman & Gillam, 1972).

Light from the glow modulator tube, after collimation, was passed through a six-filter (neutral density Tiffen metallic) intensity programmer and additional gelatin filters (neutral density Kodak Wratten No. 96) mounted in an adjacent fixed filter holder, before being decollimated to transilluminate the target. The switch-controlled, solenoid-driven intensity programmer positioned any of the sixty-four possible filter combinations into the collimated light path. The intensity programmer provided luminance control over a 1.4 log unit range; filter combinations were selected to vary luminance in either .10 (for the threshold procedure) or .88 (for the flicker procedure) log unit steps. Additional luminance control was accomplished by manually changing filters in the fixed filter holders. Low density watch glasses were also placed in the fixed filter holders for daily calibration. All of these filters and combinations were calibrated in our laboratory (Appendix).

Two photomultiplier tubes (RCA 1P21), whose inputs were filtered (Kodak Wratten No. 106) to approximate the C.I.E. photopic-luminous- efficiency curve, monitored relative luminance before and after filtering. Photomultiplier 1 received its input directly from the glow modulator tube while photomultiplier 2 received reflected light from a pellicle beamsplitter (National Photocolor Corp.) situated on the other side of the filters from the glow modulator tube. An oscilloscope (Tektronix 532) display of the photomultiplier outputs enabled the experimenter to monitor intensity and duration characteristics of the stimuli on a trial to trial basis.

After passing through the filters and pellicle, the light was decollimated to transilluminate an opal glass target which was viewed from the other side by the subject. The stimulus was a white appearing, circular target, subtending a visual angle of 22 minutes. By displacing the opal diffusing glass just beyond the focal length of the decollimating lens, the stimulus appeared homogeneous with luminance differences around the circumference, as assessed with a Pritchard photometer (Photo Research Corp.), of less than 10 percent.

The subject was seated in a light-tight booth and binocularly viewed the stimulus at a distance of 52 cm. Stable head position was maintained with the aid of an adjustable chin and forehead rest. Four dim red fixation lines which surrounded the target in an incomplete cross pattern assisted the subject in maintaining foveal fixation. To control for the effects of pupil size on flicker discrimination the stimulus and fixation lights were viewed through two adjustable 2.5 mm artificial pupils mounted to the forehead rest. Subjects were encouraged to position their eyes as close to the artificial pupils as possible.

<u>Daily calibration</u>. A Pritchard photometer (Photo Research Corp.) was modified to convert luminous flux, in the plane of the stimulus target, into a voltage output on an oscilloscope display. Before each testing day, a

.41 log unit neutral density filter and low density watch glasses were placed in the fixed filter holder until the Pritchard photometer output was lowered to .8 volts, yielding a luminance of approximately 1.499 millilamberts. The .41 log unit neutral density filter was then removed so that the average unfiltered retinal illuminance was 1.777 log trolands (Intensity 2). The calibration procedure was repeated after each testing day. Changes were found to be negligible over this period, the amount of filtering needed to reduce the photometer output to .8 volts either did not change, or when it did, the change was never more than .04 log units.

Timers. Trial events, including the presentation of a warning click, the timing of the preparatory interval, and the presentation of the stimuli, were controlled by a nine-channel multivibrator timer (Logical Instruments Co.) with an indeterminancy of 1 part in 10,000 and by a Digital Time Generator Model TG-760 (Analog and Digital Instruments Co.) with an accuracy of plus or minus .005 msec. The timer's stimulus channels delivered logic pulses to a gating device that activated and deactivated the glow modulator tube for the preset durations and intervals.

A response key to measure response and reaction times was located to the subject's right. The time from the onset of the stimulus to the subject's finger lift response was read off a model 5304A timer/counter (Hewlett-Packard) to an accuracy of .0001 msec.

Click and tone generation. The amplitude of the 500 msec warning click and 100 msec warning tones (used in the threshold procedure only) was controlled by a model DC-IMC attenuator (Hewlett-Packard) connected to a model LA 750 speaker (Lafayette Instruments) in the subject's booth. The frequency of the tones was controlled by a model 200 CD wide range oscillator (Hewlett-Packard) set at 275 Hz. The interval between the offset of one warning tone and the onset of another such tone was 900 msec.

The duration of the feedback tone (used in the threshold procedure only) lasted as long as the correct response button was pressed. The frequency of the feedback tone was approximately 1000 Hz. All auditory stimuli were well above the hearing threshold and well below a painful intensity level.

Mood scales. Eleven self-rating scales were given to each subject before and after each experimental session. Each scale was on a separate sheet of paper which contained five 100 mm lines (cf. Zealley & Aikens, 1969). Subjects were instructed to place a check mark somewhere on the top line (before session) or on the second line (after session) to reflect how they felt at the moment. The scale ranged from "not at all" (on extreme left) to "very" (on extreme right). The rated adjectives were: drowsy, energetic, anxious, annoyed, depressed, hungry, sexually aroused, physically uncomfortable, restless, tired, and alert.

Procedure: Pilot Study

Since the main experiment was a signal detection discrimination study, a frequency level of the stimulus that yielded a sufficient number of false alarms had to be determined. This was accomplished using the Up-Down-Transformed-Response method described by Levitt and Treisman (1969). This is a staircase method whereby the stimulus level (in this case the number of cycles per second) was determined by the subject's response. For example, if the subject responded "flicker" on two consecutive stimulus presentations, then the frequency of the stimulus was increased by 2 Hz. If the response was "fused" on two consecutive presentations, then the stimulus frequency was decreased by 2 Hz. By following the instructions on charts prepared by Levitt and Treisman, the experimenter ran a sufficient number of trials to obtain at least three "peaks" (high points on the charts) and three "troughs" (low points on the charts) which were averaged to yield a frequency rate at which the subject responded "flicker" 70.7% of the time and a frequency at which a "flicker" response was given 29.3% of the time. By varying the frequency in this manner and later by varying the luminance using a signal detection procedure, it was determined that a "signal" stimulus of 16 Hz and a "noise" stimulus of 116 Hz matched for apparent brightness (Talbot level)[2] at a retinal illuminance of 1.777 log trolands, generally yielded flicker sensitivity scores (d'e) in the 1.0 to 1.5 range.

Procedure: Main Experiment

After the subject agreed to be tested and met the criteria outlined in the Subject section, she or he went through the following sequence of procedures: informed consent; initial mood scales; visual detection threshold; signal detection flicker discrimination; simple reaction time (for a small subgroup of subjects); and final mood scales. Each procedure is described next in more detail.

Informed consent. Control subjects and patients were told that the purpose of the task was to measure visual performance to flashes of light. After being shown the apparatus, the subjects and patients were informed that they could terminate their participation in the study at any time. Patients were also told that discontinuation would not affect their treatment in any way. Written consent was obtained and all participants were given a copy of the consent form to keep.

Mood scales. Subjects were asked to fill out the 100 mm line mood scales. It was stressed that these scales refer to how the subject felt at the moment and examples of how to use the scale were given.

Visual threshold procedure. Subjects were then read the following instructions:

> This is a procedure to find out how well you can see a brief flash of light. When I say "get ready" you are to sit up and position your head like this (subjects were shown how to use the chin and forehead rest and the artificial pupil). When you clearly see the four red lines say "ready". If you cannot see them clearly, then I will help you. Now you will hear a click, which is only a warning

signal, followed a couple of seconds later by three tones. Along with one and only one of the tones, you will see a light. If the light was presented with the first tone, press the button closest to you. If the light was presented with the second tone, press the middle button. If the light was presented with the third tone, press the button furthest away from you. Sometimes you will not see any light, when this happens guess your best as to when the light was presented by pressing button number 1, number 2, or number 3. Be sure to respond on every trial and wait until the end of the third tone to make your choice. If you are correct, you will hear a high pitched sound (demonstrate). If you are incorrect, you will hear nothing. Any questions?

Subjects then dark adapted for five minutes and their absolute thresholds (the luminance level yielding 67% correct responses) were estimated using a three-interval forced-choice procedure. This procedure is an adaptation of the Block Up and Down Two-Interval Forced-Choice procedure introduced by Campbell (1963) and has been used extensively in our Department (cf. Bruder et al. 1975). The test-retest reliability of this procedure is greater than 0.9 with normal subjects and greater than 0.8 with psychiatric patients (Bruder, Note 3).

The detection threshold procedure consisted of a warning click followed two seconds later by a series of three tones (observation intervals). A light pulse (2 msec) was randomly presented in one of the observation intervals. The subject indicated which interval contained the signal by pressing one of three buttons. Feedback was given after every trial via the presence or absence of the feedback tone.

Inspection of the data revealed that none of the patients or control subjects had a marked position bias which might have affected the detection threshold. That is, none of the observers responded disproportionately often to any one of the three buttons.

The intensity level used in any given trial was determined by first setting the intensity well above threshold and decreasing it in .2 log unit steps until an incorrect response was made. That intensity was then used for the next block of three trials. The stepping rules for estimating the 67% correct point were:

0 or 1 correct--increase intensity by .2 log units
2 correct-------stay at the same intensity
3 correct-------decrease intensity by .2 log units.

When the same intensity level was obtained three times, the step size of was varied in .1 log unit steps until a final threshold estimate was reached. Threshold was defined as the intensity level that was revisted three times in the .1 log unit step size condition.

Subjects were given a sufficient number of practice trials to ensure understanding of the task.

Flicker discrimination procedure. Subjects were then read the following instructions:

> This is a procedure to find out how well you can tell the difference between a light that appears to be flickering and a light that eppears to be fused or steady. Again, when I say "get ready" you are to sit up and position your head as you did before. When you clearly see the four red lines, press this button (the reaction time button) with the index finger of your right hand. If you cannot see them clearly, I will help you. Now you will hear a click, which is only a warning signal, followed by a light. You will always see the light. Your job is to decide if the light appears to flicker or if it appears as a fused or continuous light. As soon as you make your decision, flicker or fused, lift your finger from the button, tell me what you decided,

> flicker or fused, and tell me how confident you are in your decision by saying "positive", "fairly sure", or "guess". Remember, you will not always be able to tell whether or not the light was flickering or fused so please use all categories, that is, positive, fairly sure, or guess. When you are ready for the next light, that is when you clearly see the four red lines, simply press the button with your index finger and we will begin again. The light should appear to flicker about half the time and it should appear fused or continuous about half the time. Any questions?

Subjects dark adapted for five minutes and visual flicker discrimination sensitivity and response criterion were determined using the following procedure:

After the subjects pressed the response time button, a 500 msec warning click was followed 2 sec later by a 500 msec stimulus train of either 16 Hz (signal) or 116 Hz (noise). For the signal condition, nine (2 msec) pulses with an interpulse interval of 60.5 msec were delivered. For the noise condition, 59 (2 msec) pulses with an interpulse interval of 6.6 msec were delivered. The stimuli were equated for brightness (Talbot level) using neutral density filters. The subjects then lifted their finger, responded "flicker" or "fused", gave their rating response, and prepared for the next trial. The experimenter noted the verbal response and response time and delivered the next stimulus approximately 15 sec later.

A session consisted of nine blocks of 30 trials. The first block of 30 trials were practice and not counted in the data analysis. Blocks 2 through 5 (120 trials) were run at a retinal illuminance of 1.777 log trolands (Intensity 2). Blocks 6 through 9 (120 trials) were run either at 1.892 log trolands (Intensity 3) or 1.676 log trolands (Intensity 1) depending on the subject's performance at Intensity 2. If the subject scored at least 20 out of 30 correct in each block on blocks 2 through 5, then blocks 6 through 9 were run at Intensity 1. If the subject scored less than 20 out of 30 correct in each block on blocks 2 through 5, then blocks 6 through 9 were run at Intensity 3. Six of nine major depressive patients, two of five dysthymic patients, and three of nine normal subjects were tested at Intensity 1. Two major depressive patients (nos. 101 and 102) were tested only at Intensity 2 for eight blocks (210 analyzable trials).

The probability of any one trial being the signal was .5 with the restriction that an equal number of signal and noise trials be presented within a block. Thus, each Receiver Operating Characteristic (ROC) curve (see Signal Detection Analysis section) was based on 60 signal and 60 noise trials except for major depressve patients 101 and 102, whose Intensity 2, Session 1 ROC curves were based on 105 signal and 105 noise trials.

Each block lasted about seven minutes and subjects were given a two minute rest period between blocks. Prior to each block, subjects were presented with at least two examples of each of the signal and noise stimuli and told which was which. More presentations were given if requested.

Simple reaction time. To assess the effect of speed of response on the response time variable, two major depressive patients, three dysthymic patients, and seven normal control subjects were given a simple reaction time task after the flicker discrimination procedure. Subjects were instructed to lift their right index finger as quickly as possible when they saw a light. It was emphasized that they did not have to make any decisions or give any verbal responses and that speed was the important variable.

Subjects were given one block of 30 trials at Intensity 2 and one block of 30 trials at whatever intensity they received in the flicker discrimination procedure. Half the subjects received Intensity 2 first. The stimuli were identical to the stimuli presented in the flicker discrimination task and the same random ordering of signal and noise applied.

Finally, subjects were again given the self report scales to assess any mood changes between the beginning and the end of the session; they were then asked for an opinion of their performance and of the tasks; and an appointment was made for them to return in about four weeks. For most subjects, the entire session lasted about 2 hours. It should be considered that most subjects found the flicker discrimination task difficult, long, and boring. This may help explain why a number of subjects performed at chance levels of flicker sensitivity.

Signal Detection Analysis

Rating scale data. Two ROC curves per subject were generated for each session, one for each intensity level tested. One curve was for Intensity 2 and the other curve was for either Intensity 1 or 3. Unless otherwise indicated, all subsequent discussions apply only to Intensity 2 of Session 1.

All ROC curves were generated according to the method described by McNicol (1972, especially Chapter 5). A rating scale ROC curve is a plot of cumulative responses to signal (16 Hz) trials (in _z_-scores) versus cumulative responses to noise (116 Hz) trials (in _z_-scores). The six response categories (from "flicker positive" to "fused positive") yielded ROC curves with a maximum of five data points. Categories with empty cells were combined with adjacent categories; in this case ROC curves had less than 5 data points. With one exception, all ROC curves were based on at least two data points. The one exception was a major depressive patient (#120) who never said "flicker" to the 116 Hz stimulus (i.e., never "missed"). Therefore, the path of her ROC curve was estimated from one data point by assigning a value of one to the standard deviations of underlying distibutions of both the 16 Hz and the 116 Hz trials. The mean number of data points for the different groups of subjects was 3.6 (_SD_ = 1.2) for the major depressive patients; 4.2 (_SD_ = 0.8) for the dysthymic patients; and 4.1 (_SD_ = 1.6) for the normal control subjects.

The points on the three ROC curves for each of the three subject groups (_Results_ chapter, Figures 2 and 3) were calculated by taking the mean of the _z_-scores of all subjects for each response category, separately for the

16 Hz and the correspondinW 116 Hz stimulus (McNicol, 1972).

The path of the ROC curve was estimated using the mutual regression method (Grice, 1966), a technique which bypasses the inadequacies of the method of least squares, the method frequently used for signal detection data (see discussion by McNicol, 1972). The mutual regression method gives the best fitting straight line through a set of rating scale derived data points and is given by the equation:

$$Y = sy/sx\ X + My - sy/sx\ Mx \qquad (1)$$

where:

Y = *z*-score of 16 Hz trial

sy = standard deviation of 16 Hz trials

sx = standard deviation of 116 Hz trials

sy/sx = slope of the ROC curve

X = *z*-score of 116 Hz trial

My = mean of the 16 Hz trials

Mx = mean of the 116 Hz trials

Using this formula, it was determined that a straight line reasonably fit each individual set of data points. However, a statistical test of the slopes of the ROC curves of all 39 subjects (including analysis of those subjects who performed at chance levels) indicated that the average slope significantly deviated from unity

($\underline{M}$ = 0.81, $\underline{t}$ (38) = 5.84, $\underline{p} < .05$). A statistical test of the slopes of the ROC curves of just those 25 subjects who performed above a chance level also indicated that the average slope significantly deviated from unity ($\underline{M}$ = 0.76, $\underline{t}$ (24) = 5.99, $\underline{p} < .05$). To use d', the most preferred parametric measure of sensory sensitivity in signal detection theory, it is necessary to assume that the underlying distributions of signal and noise are Gaussian and of equal variance. However, since the present data met the former but not the latter criterion (slope deviations from 1.0 violate the equal variance assumption), an alternate parametric measure of flicker sensitivity, d'e, was used instead.

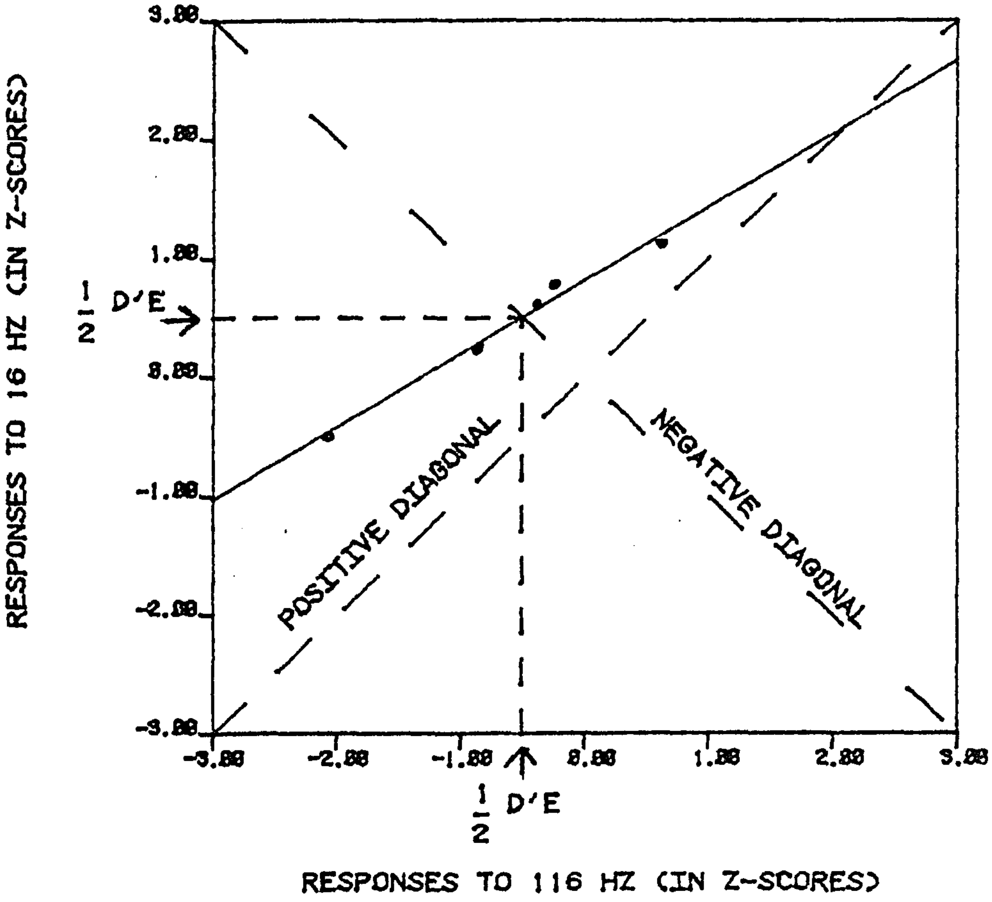
The d'e measure was defined as twice the value of the $\underline{z}$-score to the 16 Hz or 116 Hz stimulus, ignoring sign, at the point where the ROC curve intersects the negative diagonal. Figure 1 shows the rating scale ROC curve for normal control subject #210 illustrating a d'e of 1.0. Thus, even though d'e is a parametric measure, it is not as affected by the slope of the ROC curve as is the value of d' (Markowitz & Swets, 1967; McNicol, 1972). It can be calculated by the equation:

$$d'e = -2b / (1 + a) \qquad (2)$$

where:

b = Y-intercept of ROC curve (My - sy/sx Mx,

Figure 1. Rating scale ROC curve for normal control subject #210 illustrating a d'e of 1.0.


RESPONSES TO 16 HZ (IN Z-SCORES)
3.00
2.00
1.00
$\frac{1}{2}$ D'E
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
$\frac{1}{2}$ D'E
POSITIVE DIAGONAL
NEGATIVE DIAGONAL
RESPONSES TO 116 HZ (IN Z-SCORES)

from Equation 1)

a = slope of ROC curve (sy/sx, from Equation 1)

To determine if the subjects were responding at a better than chance level, two by two contingency tables were generated for each set of data. The entries of the cells were the frequency with which the subject responded "flicker" to the 16 Hz stimulus (hits); "flicker" to the 116 Hz stimulus (false alarms); "fused" to the 116 Hz stimulus (correct rejections); and "fused" to the 16 Hz stimulus (misses). If the subjects response distribution was random (as determined by a nonsignificant chi-square), then it was assumed that the subject was not discriminating the 16 Hz from the 116 Hz stimulus and that subject's data were not analyzed further. Based upon this procedure, the minimally acceptable d'e (i.e., a d'e greater than chance level) was found to be 0.29.

Two measures of response criterion, log beta and the false alarm rate, were calculated. McNicol (1972) pointed out that distributions of beta scores are often skewed. Therefore, to help normalize skewed data and thereby facilitate statistical analysis, the beta scores were transformed to log beta scores for analysis.

Log beta, the parametric measure of response criterion, was defined as the natural logarithm of beta. Beta is the ratio of the height of the underlying signal distribution to the height of the underlying noise distribution at a given criterion point. Beta was calculated for each of the points (maximum of five) on the ROC curve, before adjacent empty-celled categories were combined, using the equation:

$$\text{Beta} = Ys/Yn \qquad (3)$$

where:

Ys = ordinate of the 16 Hz distribution for a given response category

Yn = ordinate of the 116 Hz distribution for the corresponding response category.

The logarithm of beta at the first data point on the ROC curve is referred to as log beta 1, at the second point as log beta 2, etc.

In the present study, since, on any given trial there was an equal probabilty (0.5) that the stimulus would be 16 Hz or 116 Hz, a log beta value of zero indicates that the observer was unbiased; a positive log beta value indicates that the observer adopted a strict or cautious criterion and was biased towards making fused responses; and a negative log beta value indicates that the observer adopted a lax or risky criterion and was biased towards making flicker responses.

The false alarm rate was calculated by obtaining the probability of the subject's responding "flicker" to the

116 Hz stimulus and converting this probability to a *z*-score.

Response time data. ROC curves for response times were calculated using the method described by Moss et al. (1970) and others (Emmerich et al., 1972; Katz, 1970, Norman & Wickelgren, 1969). This method depends on the assumption that faster response times means greater response confidence. Table 5 of the *Results* chapter indicates that this is a reasonable assumption for all of the subject groups.

Geometric mean response times (the antilogarithm of the sum of the logarithms divided by the number of response times) were used in Table 5 to reduce the skewness of response time distributions.

Also to reduce the skewness of response time distributions, all response times were transformed into their base ten logarithmic equivalents and partitioned into ten equal log unit steps. The slowest response time was 3.6990 log msec, the fastest time was 2.3200 log msec, and the range was 1.3790 log msec. The step size was the range divided by the number of steps (10) or .1379 log units. Thus, a response time between 2.3200 and 2.4579 log msec was equivalent to the most confident response, a time between 2.4580 and 2.5958 was a less confident response, and a time between 3.5612 and 3.6990 log msec was the least confident response. This technique yielded ROC curves with a maximum of 19 data points. Empty cells were combined with adjacent categories. All subjects had ROC curves based on at least three data points. The mean number of data points was 5.2 (*SD* = 2.3) for the major depressive patients, 7.2 (*SD* = 2.4) for the dysthymic patients, and 4.6 (*SD* = 2.5) for the normal control subjects.

The dependent variables, d'e and log beta, were calculated in the manner described in the previous section.

Statistical Analyses

Separate one-way analysis of variance tests were performed on the dependent variables of d'e, log beta, false alarm rate, and self-rating scores. Homogeneity of error variance was tested using Cochran's procedure (Winer, 1971), and if significant ($p < .05$) heterogeneity of error variance was found, the data were logarithmically transformed (except for log beta). If that transformation did not significantly reduce error variance, then nonparametric statistics were used (Kruskal-Wallis procedure; Siegel, 1956). Post-hoc analyses were done using the Student-Newman-Keuls or Mann-Whitney *U* procedures for parametric and nonparametric data, respectively.

All correlations reported in this study refer to Pearson product-moment correlation coefficients.

All statistical hypotheses were two-tailed and a *p* value of less than 5% was interpreted as statistically significant.

CHAPTER III

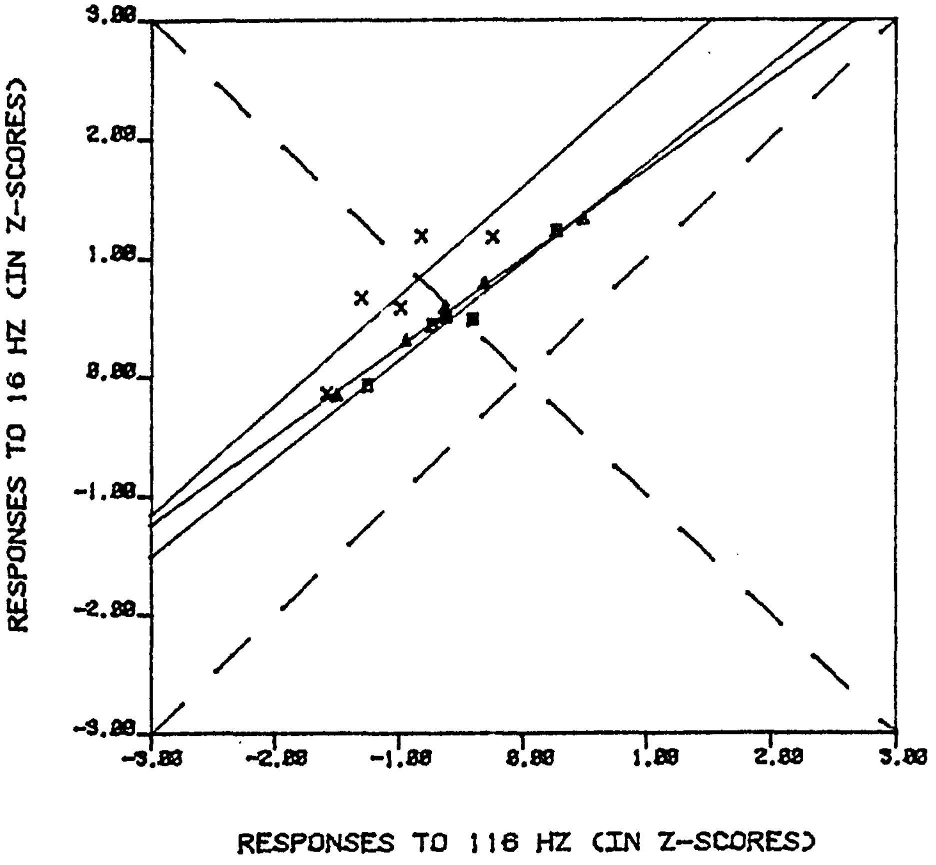
Results

Rating Scale Analysis

Flicker sensitivity. Figure 2 shows the rating scale ROC curves based on mean z-scores at each data point for the three groups of subjects. All average rating scale ROC curves were based on five data points. The slopes of the curves were 1.07, 0.84, and 0.75 for the major depressive patients, dysthymic patients, and normal control subjects, respectively. The slopes indicate that the underlying variances of the signal and noise distributions are essentially equal for the major depressive patients. For the dysthymic patients and the normal control subjects the underlying variance of the signal distribution is greater than the variance of the noise distribution.

The rating scale ROC curves of the dysthymic patients and normal control subjects overlap; the curve of the major depressive patients is displaced furthest from the positive diagonal (Figure 2), suggesting that the major depressives were able to discriminate the 16 Hz stimulus from the 116 Hz stimulus better than both of the

Figure 2. Rating scale ROC curves based on mean *z*-scores at each data point for the major depressive patients (x's), dysthymic patients (squares), and normal control subjects (triangles).


3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
RESPONSES TO 16 HZ (IN Z-SCORES)
RESPONSES TO 116 HZ (IN Z-SCORES)

other two groups. Table 2 shows that the mean rating scale d'e values were 1.78 (*SD* = 1.03), 1.07 (*SD* = 0.23) and 1.16 (*SD* = 0.67) for the major depressive patients, dysthymic patients, and normal control subjects, respectively. However, this difference in d'e (logarithmically transformed) was not statistically significant (F (2,22) = 1.16, $p > .05$), thus the original hypothesis that normal subjects would have higher sensory sensitivity values to visual flicker than depressed patients was rejected.

Response criterion. Figure 2 also shows that each data point on the ROC curve of the major depressive patients is closer to the y-axis than the corresponding data point on the curves of the other two groups, indicating that the major depressive patients adopted the most conservative criterion at each point. Table 3 shows that the major depressive group has the highest log beta values and that all log beta values for the dysthymic and normal control groups are similar.

One way analysis of variance test of log beta 1, log beta 2, log beta 4, and log beta 5 yielded no significant F ratios. The mean log beta 3 value was significantly different (F (2,21) = 3.99, $p < .05$) but Cochran's test for homogeniety of error variance was also significant (C

Table 2

Mean Rating Scale and Response Time d'e Values for the Three Subject Groups*

Group	Rating Scale d'e	Response Time d'e
Major Depressive		
Mean	1.78	1.71
<u>SD</u>	1.03	1.03
<u>n</u>	11	11
Dysthymic		
Mean	1.07	0.96
<u>SD</u>	0.23	0.29
<u>n</u>	5	5
Normal Control		
Mean	1.16	1.01
<u>SD</u>	0.67	0.61
<u>n</u>	9	9

---

*Correlation coefficients between rating scale and response time d'e were r (9) = 0.96, r (3) = 0.96 and r (7) = 0.96, for the major depressive patients, dysthymic patients, and normal control subjects, respectively. All correlations were significant at the 5% level of confidence.

Table 3

Mean Log Beta Values for the Rating Scale ROC Curve for the Three Subject Groups

	Log Beta				
	1	2	3*	4	5
Group					
Major Depressive					
Mean	1.31	0.77	0.54	0.28	-0.57
SD	0.66	0.69	0.63	0.72	0.86
n	8	10	10	10	10
Dysthymic					
Mean	0.88	0.20	0.03	-0.10	-0.67
SD	0.73	0.37	0.25	0.26	0.47
n	5	5	5	5	5
Normal Control					
Mean	1.04	0.34	-0.02	-0.25	-0.77
SD	0.64	0.27	0.27	0.42	0.63
n	7	9	9	9	9

*Major depressive patients significantly differed from normal control subjects only, U (9, 10) = 15, $p < .05$.

= 0.74, $p < .05$); therefore, the nonparametric, Kruskal-Wallis one way analysis of variance test was employed and found to be significant (H (2) = 6.64, $p < .05$). Subsequent Mann-Whitney U tests indicated that the major depressive patients were significantly more conservative than the normal control subjects for the log beta 3 measure (U (9,10) = 15, $p < .05$). None of the other comparisons was significant.

It is important to remember that a rating scale ROC curve is based on cumulative, not independent, frequencies, which explains the fact that the log beta values get progressively smaller from log beta 1 to log beta 5. Therefore, unless one group of subjects is unusually extreme in their rating responses, group differences are not likely to be obtained at the extreme end of the curve (point 5). In fact, the smallest differences among the groups was for the log beta 5 measure (see Table 3). It is interesting that the only statistically significant difference was at log beta 3, the middle point of the five point curve and the point which represents the cumulated frequencies of all "flicker" responses.[3]

The mean log beta 3 values for the dysthymic patients and the normal subjects were very close to zero which indicates that these groups generally adopted an unbiased response criterion. The mean log beta 3 value for the major depressive patients was greater than zero which indicates that this group generally adopted a conservative criterion, i.e., had a bias towards fused responses.

Another measure of response criterion is the false alarm rate. In the present study a false alarm was defined as any "flicker" response to the 116 Hz stimulus. A small number of false alarms indicated an unwillingness to respond "flicker" very often and was interpreted as a conservative criterion. Table 4 shows that the mean false alarm rates (in *z*-scores) for the three groups of subjects were 1.20 (*SD* = 0.61), 0.54 (*SD* = 0.31), and 0.62 (*SD* = 0.33) for the major depressive patients, dysthymic patients, and normal control subjects, respectively. The larger the *z*-score, the smaller the number of false alarms. The difference between groups in mean false alarm rates was statistically significant ($F(2,22) = 5.12$, $p < .05$). Subsequent Student-Newman-Keuls procedures indicated that the major depressive patients gave significantly fewer "flicker" responses (were more conservative) than the other groups ($p < .05$). None of the other pairwise comparisons was significant.

Table 4

Mean False Alarm Rates (in *z*-scores)* For The Three Subject Groups

Group	*n*	Mean	*SD*	*F*	*p*
Major Dep.	11	1.20	0.61		
Dysthymic	5	0.54	0.31	5.12**	< 0.05
Normal	9	0.62	0.33		

*Note: The larger the *z*-score, the smaller the probability of a false alarm.

**Student-Newman-Keuls procedure indicated that major depressive patients (Major Dep.) significantly differed from both other groups ($p < .05$).

Response Time Analysis

The geometric mean response times at the different levels of response confidence for the signal (16 Hz) and noise (116 Hz) stimulus conditions are shown in Table 5. Notice that, in all cases, response times increased as response confidence decreased, a finding consistent with that of other studies (Emmerich et al., 1972; Katz, 1970; Kerkhof et al., 1980; Moss et al., 1970; Norman & Wickelgren, 1969). Also, within any one response confidence category, the response times among the groups did not differ (none of the ten one way *F*-tests were significant at the .05 confidence level), indicating that the response time-response confidence relationship is independent of diagnosis.

Each geometric mean in Table 5 was based on a different number of trials. For example, more subjects responded "flicker positive" than "fused guess" to the 16 Hz stimulus. Inspection of the response time-response confidence data of each subject revealed that, in general, response times increased as response confidence decreased for both signal (16 Hz) and noise (116 Hz) conditions. There were no indications in the individual or group data to suggest any systematic differences among the subject groups based on responses times.

Table 5

Geometric Mean Response Times (in msec) at Different Levels of Response Confidence for Signal (16 Hz) and Noise (116 Hz) Conditions

Signal (16 Hz) Condition

	Response Confidence					
Group	Flicker Positive	Flicker Fairly Sure	Flicker Guess	Fused Guess	Fused Fairly Sure	Fused Positive
Major Depressive (*N*=11)	874	986	1219	1137	1098	911
(*n*)	(11)	(11)	(8)	(4)	(10)	(10)
Dysthymic (*N*=5)	927	1042	1340	1363	1161	1091
(*n*)	(5)	(5)	(3)	(4)	(5)	(5)
Normal Control (*N*=9)	895	973	1178	1191	999	962
(*n*)	(9)	(9)	(9)	(7)	(9)	(9)

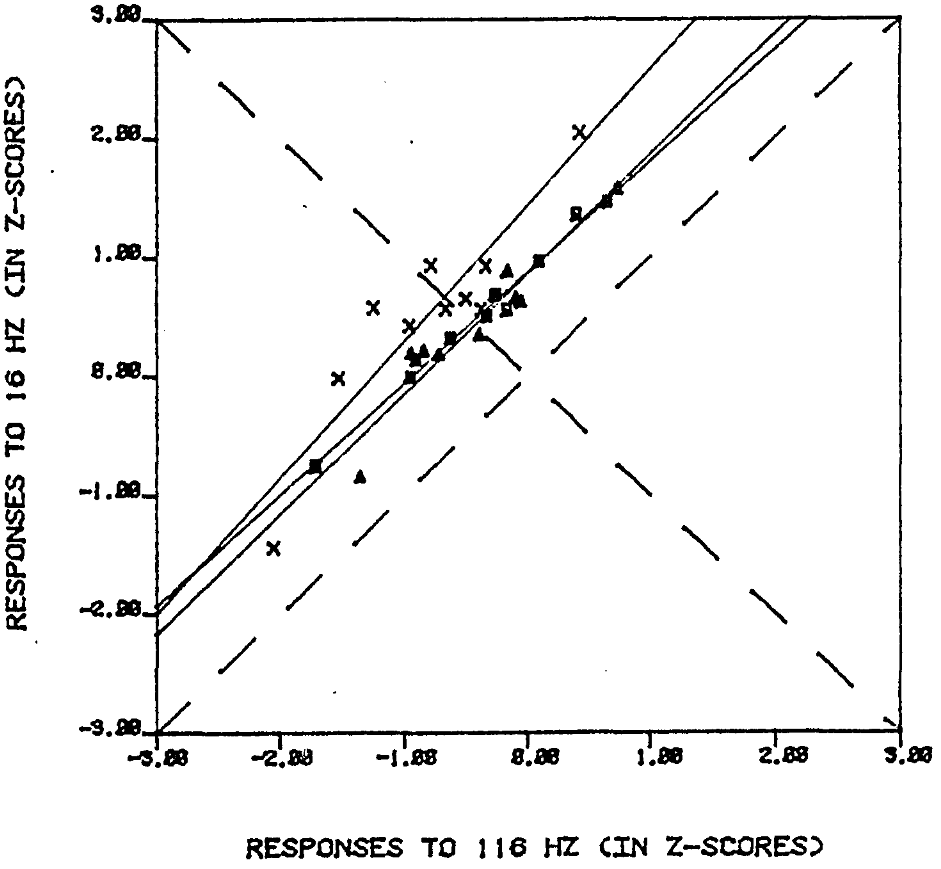
Noise (116 Hz) Condition

	Response Confidence					
Group	Flicker Positive	Flicker Fairly Sure	Flicker Guess	Fused Guess	Fused Fairly Sure	Fused Positive
Major Depressive (*N*=11)	947	1057	1193	1338	1031	896
(*n*)	(8)	(11)	(8)	(6)	(11)	(11)
Dysthymic (*N*=5)	973	1076	1403	1369	1114	952
(*n*)	(5)	(5)	(4)	(2)	(5)	(5)
Normal Control (*N*=9)	896	973	1070	1070	1008	946
(*n*)	(7)	(9)	(9)	(9)	(9)	(9)

Flicker sensitivity. Response time ROC curves based on mean z-scores at each data point for the three subject groups are shown in Figure 3. The curves of the major depressive and dysthymic patients are based on ten data points and the curve of the normal control subjects is based on nine points. The slopes of the curves were 1.27, 0.94, and 1.06 for the major depressive patients, dysthymic patients, and normal control subjects, respectively. These slopes are somewhat steeper than the slopes of the rating scale ROC curves indicating that the relationship between the underlying variances of the noise and signal plus noise distributions are somewhat different for the two types of curves. The data points for these curves are more scattered than the rating scale curves because they are based on fewer z-scores per point and are therefore more variable. However, as is shown subsequently, the conclusions with respect to d'e and log beta are virtually identical for data based on rating scale and response time ROC curves.

The response time ROC curves of the dysthymic patients and normal control subjects overlap and the curve of the major depressive patients is displaced the furthest from the positive diagonal as in the rating scale ROC curves (Figure 2), suggesting that the major depressive patients were the most sensitive to visual

Figure 3. Response time ROC curves based on mean _z_-scores at each data point for the major depressive patients (x's), dysthymic patients (squares), and normal control subjects (triangles).


3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
RESPONSES TO 16 HZ (IN Z-SCORES)
RESPONSES TO 116 HZ (IN Z-SCORES)

flicker. Table 2 shows that the response time d'es were 1.71 (_SD_ = 1.03), 0.96 (_SD_ = 0.29), and 1.01 (_SD_ = 0.61) for the major depressive patients, dysthymic patients, and normal control subjects, respectively. As in the rating scale analysis, these d'e (logarithmically transformed) differences were not statisticaly significant (F (2,22) = 2.29, $p > .05$). Therefore, the original hypothesis that depressed patients would have lower sensory sensitivity values to visual flicker than normal subjects was rejected. The correlation coefficients between rating scale and response time d'es were statistically significant for each subject group indicating a strong relationship between the two methods in calculating that measure of sensory sensitivity. For the major depressive patients, r (9) = 0.96; dysthymic patients, r (3) = 0.96; and normal control subjects, r (7) = 0.96. All correlations were significant at the 5% level of confidence.

Response criterion. Mean log beta values for the response time ROC curve for the three subject groups are shown in Table 6. These data are based on ROC curves with 19 data points. Very few subjects responded either very quickly or very slowly which yielded empty cells at the extreme ends of the curves. There were not enough

Table 6

Mean Log Beta Values for the Response Time ROC Curve For The Three Subject Groups

	Log Beta*							Log Beta					
Group	4	5**	6	7***	8***	9**	10***	11***	12***	13**	14****	15	16
Major Depressive													
Mean	0.88	1.05	0.57	0.59	0.55	0.54	0.54	0.54	0.52	0.49	0.29	-0.17	-1.78
SD	1.10	0.39	0.65	0.59	0.62	0.63	0.63	0.63	0.65	0.65	0.52	0.42	1.03
n	5	8	11	10	10	10	10	10	10	10	10	10	4
Dysthymic													
Mean	0.97	0.48	0.42	0.12	0.05	0.03	0.03	0.02	-0.01	-0.09	-0.23	-0.65	-0.58
SD	0.27	0.41	0.51	0.22	0.24	0.25	0.25	0.26	0.28	0.30	0.34	0.58	0.47
n	3	4	5	5	5	5	5	5	5	5	5	5	2
Normal Control													
Mean	0.57	0.39	0.19	0.05	0.01	-0.01	-0.02	-0.02	-0.03	-0.04	-0.08	-0.31	-1.07
SD	0.50	0.35	0.42	0.31	0.28	0.27	0.27	0.27	0.27	0.27	0.25	0.43	0.87
n	3	8	9	9	9	9	9	9	9	9	9	8	3

---

*There was an insufficient number of subjects for any calculations at data points 1, 2, 3, 17, 18, and 19.

**Major depressive patients significantly differed from both other groups ($p < .05$).

***Major depressive patients significantly differed from normal control subjects only ($p < .05$).

****$F(2,21) = 3.453$, $p = .0505$. Student-Newman-Keuls procedure indicated that no two groups significantly differed at the .05 level.

data to statistically analyze six log beta values (nos. 1, 2, 3, 17, 18, and 19); thus calculations were performed on the remaining 13 log beta values.

The one way analysis of variance for log beta 5 was significant (F (2,17) = 6.64, p < .05). Subsequent Student-Newman-Keuls procedures indicated that major depressive patients were significantly more conservative than both other groups (p < .05). Due to significant heterogeneity of error variance, log beta 9 and log beta 13 values were analyzed using nonparametric tests. The Kruskal-Wallis one way analysis of variance was significant for both log beta values (for log beta 9, H (2) = 7.74, p < .05; for log beta 13, H (2) = 6.24, p < .05). Subsequent Mann-Whitney U tests indicated that major depressive patients were significantly more conservative than both other groups for log beta 9 and log beta 13 (p < .05 for each log beta value).

For five log beta values (nos. 7, 8, 10, 11, and 12) the major depressive patients did not differ from the dysthymic patients but were significantly more conservative than the normal control subjects. Significant heterogeniety of error variance indicated the use of nonparametric tests; therefore Kruskal-Wallis tests were used and all were significant at the 5% level of confidence. For log beta 7, H (2) = 7.64; for log beta 8, *H* (2) = 6.49; for log beta 10 and 11, *H* (2) = 6.64; and for log beta 12, *H* (2) = 6.34. For each log beta value subsequent Mann-Whitney *U* tests were signifcant at the 5% level of confidence.

A one-way analysis of variance approached significance for the log beta 14 value: (*F* (2,21) = 3.45, *p* = .0505) but subsequent Student-Newman-Keuls procedures indicated that no two groups significantly differed at the .05 level. Other analyses of variance did not reach statistical significance at points 4, 6, 15, and 16.

As was the true of the rating scale data, the response time data indicated that the major depressive patients were the most conservative of the three groups and this was most manifest at those points in the middle of the ROC curve. The dysthymic patients and the normal control subjects behaved similarly for all log beta values.

Reaction Time Analysis

Reaction times were in the range of 200 to 500 msec and response times were generally greater than 500 msec for the "positive" response condition and in the 1000 to 2000 msec range for some subjects for the other response conditions. Patients (two major depressive and three dysthymic combined) were slightly slower than normal control subjects (n = 7) and reaction times were faster in the signal than in the noise condition. Geometric mean reaction times for patients were 284 and 328 msec for signal and noise stimuli, respectively, and for normal control subjects the corresponding times were 279 and 312 msec. The antilogarithm of the standard deviations was 1.3 in all cases. Based on these data, it seems doubtful that speed of response per se had any significant effect on the results.

Self-rating Analysis

Of all the self-rated variables, only depression, alertness, and drowsiness significantly differentiated the three subject groups. Table 7 shows the mean self-ratings (in mm) for these variables.

Not surprisingly, the major depressive group rated themselves as being the most depressed before (M = 44.8) and after (M = 45.3) psychophysical testing, dysthymic patients were next (M = 21.0, before and M = 14.6, after), and normal control subjects were least depressed (M = 2.8, before and after). Both patient groups were significantly more depressed than the normal control group (F (2,22) = 13.99, p < .05, before testing; F

Table 7

Significant Mean Self-Ratings (in mm) for the Three Subject Groups

	Self-Rating Variable			
	Depressed Before Testing***	Depressed After Testing****	Alert After Testing**	Drowsy Difference*
Group				
Major Depressive (*N*=11)				
Mean	44.8	45.3	45.1	15.3
SD	24.2	25.9	21.1	33.4
Dysthymic (*N*=5)				
Mean	21.0	14.6	28.6	58.0
SD	15.8	15.2	14.0	33.3
Normal Control (*N*=9)				
Mean	2.8	2.8	65.4	16.2
SD	3.8	3.5	20.6	20.2

---

*Drowsy difference = Drowsy after testing minus Drowsy before testing. *F* (2,22) = 4.17, *p* < .05. Student-Newman-Keuls procedure indicated that dysthymic patients significantly differed from both other groups (*p* < .05).

***F* (2,22) = 5.96, *p* < .05. Student-Newman-Keuls procedure indicated that normal control subjects significantly differed from both other groups (*p* < .05).

****F* (2,22) = 13.99, *p* < .05. Student-Newman-Keuls procedure indicated that normal control subjects significantly differed from both other groups (*p* < .05).

*****F* (2,22) = 15.80, *p* < .05. Student-Newman-Keuls procedure indicated that normal control subjects significantly differed from both other groups (*p* < .05, data were logarithmically transformed).

(2,22) = 15.80, $p < .05$, after testing; Student-Newman-Keuls tests were significant at the .05 level). The major depressive patients were not significantly more depressed than the dysthymic patients either before or after testing.

Normal control subjects rated themselves as significantly more alert after psychophysical testing than either patient group (F (2.22) = 5.96, $p < .05$; the Student-Newman-Keuls test was significant at the .05 level). The patient groups did not significantly differ from each other.

The effect of psychophysical testing on ratings of drowsiness was most pronounced in the dysthymic patients whose mean drowsiness difference (drowsy rating after testing minus drowsy rating before testing) was 58.0 compared to 15.3 for the major depressive patients and 16.2 for the normal control subjects (F (2,22) = 4.17, $p < .05$; the Student-Newman-Keuls test was significant at the .05 level).

Self-rating on the depression scale was not correlated with the dependent variables of d'e and log beta 3 (both derived from the rating scale analysis).[3] The correlation between depression rating and d'e was r (16) = -0.12, for the major depressive patients including those who performed at or below chance, and r (3) = 0.16, for the dysthymic patients. The correlation between depression rating and log beta 3 was *r* (15) = 0.04, and *r* (3) = 0.19 for those groups (all *p* values were greater than .05). Major depressive patients, including those who performed at or below chance were divided into two groups, above and below the median depression rating of 48. These two groups did not differ on d'e (*t* (16) = 0.85) or log beta 3 (*t* (15) = 0.55). Further, the five patients with the highest self-rated depression score were compared to the five patients with the lowest self-rated depression score on d'e and log beta 3. Neither *t*-test reached the 5% level of significance (*t* (8) = 0.29, for d'e and *t* (8) = 0.52 for log beta 3).

Test-retest Analysis

The following analyses include those subjects who scored at chance levels.

Eight (of 18) major depressive patients, 2 (of 5) dysthymic patients, and 11 (of 15) normal control subjects returned for the second session. The low return rate of the patients may have been because they found the task long and boring and were not paid to participate. The daily dosages (in milligrams) of antidepressant medication received one and two weeks prior to the second testing session for the patient groups are shown in Table 8. Generally, a patient was started on 30 mg of phenelzine or 100 mg of imipramine and the dosage was increased every week by 15 mg (phenelzine) or 50 mg (imipramine) per day.

Table 9 shows the rating scale log beta 3 and d'e values for Session 1 and Session 2 for the major depressive patients. Six (of eight) major depressive patients showed better flicker discrimination performance and seven of them were less conservative in the Session 2. The mean (rating scale) d'e score for Session 1 was 0.80 (*SD* = 1.07); for Session 2 it was 1.35 (*SD* = 0.97). However, this difference was not statistically significant as determined by a *t*-test for related measures (*t* (7) = 1.80, *p* > .05). The mean log beta 3 for Session 1 was 0.28 (*SD* = 0.43) and for Session 2 it was 0.09 (*SD* = 0.67). This difference also was not statistically significant (*t* (7) = 1.90, *p* > .05).

Both dysthymic patients exhibited superior flicker discrimination in Session 1 and one was less conservative in Session 2.

Five of the 11 normal control subjects improved their flicker discrimination and 4 of them were less conservative in Session 2. The mean d'e for Session 1 was 0.73 (*SD* = 0.93) and for Session 2 it was 0.75 (*SD* =

Table 8

Daily Dosages (in Milligrams) of Antidepressant Medication Received One and Two Weeks Prior to Second Testing Session

Patient No.	Medication	Two Weeks Prior to Testing	One Week Prior to Testing
	Major Depressive Patients		
101	Phenelzine	60	75
113*	Phenelzine	75	90
115	Phenelzine	75	90
121**	Phenelzine	75	90
106*	Imipramine	200	250
107	Imipramine	150	200
112***	Imipramine	250	300
116	Imipramine	250	300
	Dysthymic Patients		
118	Imipramine	250	300
124	Imipramine	250	300

---

*Scored at chance level for Session 1.

**Scored at chance level for Session 2.

***Scored at chance level for both sessions.

Table 9

Rating Scale Log Beta 3 and d'e Values for Session 1 and Session 2 for the Major Depressive Patients

Patients Taking Phenelzine Prior to Session 2

	Log Beta 3		d'e	
	Session		Session	
	1	2	1	2
Patient No.				
101	1.14	-0.41	2.79	2.35
113	0.06	1.59	-0.25	1.99
115	-0.08	-0.27	1.41	1.64
121	0.01	0.00	0.29	0.06

Patients Taking Imipramine Prior to Session 2

106	-0.01	-0.39	0.18	0.94
107	0.57	-0.26	1.58	2.08
112	0.02	0.01	-0.35	-0.20
116	0.53	0.47	0.75	2.02

1.13). This difference was not statistically significant (*t* (10) = 0.09, *p* > .05). The mean log beta 3 for Session 1 was -0.04 (*SD* = 0.24) and for Session 2 it was 0.04 (*SD* = 0.54). This difference also was not significant (*t* (10) = 0.42, *p* > .05).

Stimulus Intensity Analysis

All 25 subjects were tested at the Intensity 2 level of the stimulus. Twelve subjects (3 major depressive patients, 3 dysthymic patients, and 6 normal control subjects) were tested at a slightly higher intensity (Intensity 3) and 11 subjects (6 major depressive patients, 2 dysthymic patients, and 3 normal control subjects) were tested at a slightly lower intensity (Intensity 1).

Increasing stimulus intensity improved the flicker discrimination performance of 1 major depressive patient, 1 dysthymic patient, and 4 normal control subjects. Flicker discrimination performance decreased in all 11 subjects tested at Intensity 1.

It is difficult to understand why only half of those subjects tested at the slightly higher Intensity 3 level improved in flicker discrimination performance. One explanation is that the subjects began to fatigue towards the end of the session. Intensities 1 and 3 were always given in the second half of the session, after the threshold procedure and the visual flicker procedure at Intenstity 2. Twenty out of the 25 subjects tested rated themselves as being more drowsy after, as compared to before, the experimental session. Therefore, due to the possible confounding factor of fatigue during the second half of the session, the effect of changing intensities on flicker discrimination performance cannot be assessed in the present study.

CHAPTER IV

Discussion

The present study was designed to extend the Perris (1966) and the Black et al. (1975) studies of visual flicker and depression by using signal detection methodology to separate flicker sensitivity from response bias. The results indicate that, in general, major depressive patients have higher log beta values and lower false alarm rates (i.e., are more conservative) than either dysthymic patients or normal control subjects. The results also indicate that these groups do not differ on d'e, a criterion-free measure of flicker sensitivity.

Both Perris (1966) and Black et al. (1975) reported that psychotically depressed patients had lower visual flicker thresholds than normal subjects. In addition, Perris reported no significant flicker performance difference between reactive neurotic depressive patients and normal subjects while Black et al. reported psychotically depressed patients performed more poorly than schizophrenic patients. Perris interpreted his findings as due to either an attentional deficit or to personality differences among his subject groups, i.e., the psychotic depressed group were probably more anxious than the other groups. Black et al. suggested that visual flicker may be a useful diagnostic aid to distinguish depression from schizophrenia in psychotic patients who complain of a depressed mood.

The high log beta values and low false alarm rates shown by the major depressive patients in the present study suggest that the psychotically depressed patients in previous studies were probably less willing to respond "flicker" than the other groups which led to less sensitive visual flicker thresholds. The present study also demonstrated no differences between dysthymic patients and normal control subjects on the measures of d'e, log beta, or false alarm rate. These findings are consistent with those of the Perris study which found that the neurotic reactive group (who were probably similar in symptomatology to the dysthymic group in the present study) did not differ from the normal subjects on visual flicker performance.

The findings of the present study are consistent with the previous studies which reported no differences in visual flicker sensitivity between psychiatric patients and normal control subjects when response criterion was controlled (Clark et al., 1967; 1963). This study of depressed patients also adds the visual flicker measure to the other tasks in which it has been reported that depressed patients tend to adopt a more conservative response criterion than do normal subjects. These previous studies were of pain discrimination (Davis et al., 1979) and recognition memory (Miller & Lewis, 1977).

In a related auditory vigilance task, Byrne (1977) reported that psychotic depressive patients were less willing to respond to the presence of a signal than either neurotic depressive patients or normal control subjects. One interpretation of that finding is that psychotic depressive patients were more conservative than either of the other groups, a conclusion which is consistent with that of the present study. However, Byrne (1977) reported that the psychotic depressive patients were the least sensitive group (i.e., psychotic depressive patients had fewer hits than the other subjects). The present study did not find a sensitivity difference. Apparently, in the Byrne study the measurement of response criterion was not independent of the measurement of sensory sensitivity. Thus, the two studies cannot be readily compared for several reasons: because of the different tasks used; the different data analysis techniques (Byrne only reported the hit and false alarm rates. He did not report any signal detection theory measures such as d' or beta); and differences in patient populations (e.g., some of the patients in the Byrne study probably had overt psychotic symptoms such as delusions or hallucinations, whereas none of the patients in the present study displayed such symptoms).

Johnson (1975) has postulated that depressed people have a sensory processing deficit. He also suggested that future experiments should attempt to determine if the sensory processing systems of depressed patients are deficient or if the deficit has more to do with output or response factors. For example, do patients correctly register environmental events but cannot respond to them? The present study suggests that the main source of the differences in performance of the major depressive patients is not in their sensory processing nor in their motor output but rather in their extremely conservative response criterion. Thus, the decision-making stategy is one important factor which differentiates the major depressive patients from the dysthymic patients and normal subjects.

Differences in the decision-making strategy between depressed patients and normal subjects have also been reported by Glass, Uhlenhuth, Hartel, Matuzas, and Fischman (1981) who tested 31 normal subjects, 25 major depressive, 3 minor depressive, and 4 intermittent depressive patients on a number of tasks including a modification of the Sternberg (1966) memory scan procedure. In this version of the Sternberg procedure subjects were presented with a series of digits for three seconds and then given a test digit one second later. Subjects lifted their right index finger to indicate that the test digit was a member of the preceding series or their left index finger to indicate that it was not. Glass et al. (1981) reported no differences between the groups in number of errors but the depressed patients appeared to decide more slowly (as indexed by response time) than did normal subjects if the test digit was a member of the search set. The results of the Glass et al. study suggest that the depressed patients maintained accuracy comparable with the control subjects at the expense of slower performance. Glass et al. postulated that the performance of the depressed patients may have been an indicator of a slowing in the decision-making process of those patients. It is possible that the slower performance of the depressed patients was due to a more conservative, deliberate decision-making style.

Note that the present study did not find response time differences between depressed patients and normal subjects whereas the Glass et al. (1981) study did find such differences. This may be due to the fact that the two studies are very different procedurally. In the present study, the subject had to decide whether or not a given stimulus was flickering or fused. This may or may not involve a comparison of the presented stimulus with an internal representation of what the flickering light looks like. In Glass et al. (1981) study, the subject had to maintain the members of the search set in memory in order to determine if the probe digit was a member of that set. Also, the Glass et al. study was specifically designed to measure speed of information processing whereas in the present study the response time data were logarithmically transformed and used to generate ROC curves. Thus, two very different procedures lead to the conclusion that major depressive patients and normal subjects display differences in their decision-making strategies.

In the present study, all subjects were informed that the stimulus should appear to flicker about half the time and it should appear fused about half the time. Both the normal control subjects and the dysthymic patients seemed to use this information because they generally gave an equal number of flicker and fused responses. In contrast, the major depressive patients did not seem to use this information because they generally gave more fused than flicker responses.

Note that under the conditions of the present study there was no reason for the subjects to assume that the 16 Hz stimulus was the "signal" and the 116 Hz stimulus was the "noise". A conservative response criterion here means a bias towards giving "fused" responses. This bias towards "fused" responses is evident even if the data are analyzed as if 116 Hz stimulus is considered the "signal" and the 16 Hz stimulus is considered the "noise". It is this bias towards "fused" responses which may account for the low thresholds of the depressed patients in the visual flicker study of Black et al. (1975).

It is not clear how a bias towards "fused" responses relates to the findings of the Perris (1966) study. Perris reported that the flicker threshold obtained by a descending method of limits differentiated psychotically depressed patients from normal subjects. If it is assumed that depressed patients prefer "fused" to "flicker" responses, then their lower flicker thresholds may be partially explained by the perseveration of "fused" responses (the error of habituation). However, on the ascending series (from flicker to fused) it might be expected that the depressed patients would change their responses from "flicker" to "fused" sooner than normal subjects (the error of anticipation), thus lowering their fusion thresholds. In fact, in the Perris study, the unipolar and reactive depressive groups did have lower fusion thresholds, albeit nonsignificant ones, than the normal subjects. Interestingly, with one exception (the bipolar disorder--depressed group), all subject groups in the Perris study had lower fusion than flicker thresholds, including the normal group (whose mean flicker threshold was 39.9 Hz and whose mean fusion threshold was 37.1 Hz). This suggests that the fusion threshold procedure is not sensitive enough to detect performance differences between depressed patients and normal subjects.

The question remains as to why some psychiatric patients may wish to avoid the error of reporting something which is not there (i.e., their reluctance to give false alarms). Clark et al. (1963) offered a possible explanation when they suggested that the cost of reporting an hallucination in a hospital setting can be very high. This may explain a conservative response criterion in experiments in which the patient must detect a faint stimulus from no stimulus at all, but it does not explain why some psychiatric patients are biased towards "fused" responses in a situation in which a stimulus (either flickering or fused) is presented on every trial. Perhaps a fruitful way of investigating response criterion in depressed patients would be to generate ROC curves by varying the probabilty of the occurrence of a flickering or fused stimulus. Another method might involve generating ROC curves by giving different sets of instructions to the patient (e.g., instruct some patients that the flickering light is the signal and instruct other subjects that the fused light is the signal). Although Clark (1966) investigated visual flicker performance under two instruction conditions (a facilitating condition and an inhibiting condition), conclusions from his study are limited because the psychiatric patient group was diagnosed in a general way (as mostly schizophrenic) and no normal control group was tested.

A signal detection study of the auditory behavior of the elderly found that aged subjects were more cautious than younger ones (Rees & Botwinick, 1971). Another auditory detection study found a significant correlation ($\underline{r}$ (13) = +0.786, $\underline{p} < .001$) between response bias and age in two depressed groups; one with obsessional symptoms and one without such symptoms (Milner, Beech, & Walker, 1971). Since the major depressive patients in the present study were older than the normal subjects (though not significantly so), it is possible that age was a confounding factor in the results reported here. However, since none of the correlations between log beta

3 and age were statistically significant ($p > .05$), it is unlikely that age played a significant role in the results of the present study. The correlations between log beta 3 and age were $r\ (8) = -0.003$, $r\ (3) = -0.002$, and $r\ (7) = -0.037$ for the major depressive patients, dysthymic patients, and normal subjects, respectively. The age range of the older group in the Rees and Botwinick (1971) study was 65 to 77 years and in the present study the oldest patient was 57 years old. The age range of the subjects in the Milner et al. (1971) study was not reported.

It is possible that a factor such as the severity of a patient's illness and not the subtype of their depressive disorder can help to explain why the major depressive patients were significantly more conservative than the dysthymic patients (as measured by the false alarm rate) and normal subjects (as measured by the false alarm rate and log beta 3).

The relationship between severity of depression and performance on visual flicker has not been previously investigated. The results of studies in schizophrenic and organic patients suggest that severity of illness may be related to visual flicker performance. Barry (1962) found a significant negative correlation between severity of illness (as measured by the resident's mental status report at the time of admission) and visual flicker performance. Watson et al. (1969) found a significant interaction between diagnosis (schizophrenia or chronic brain syndrome) and chronicity (length of previous hospitalizations) meaning that visual flicker performance was a sensitive indicator of length of previous hospitalizations regardless of diagnosis. However, because those studies used the method of limits to measure visual flicker performance, it cannot be determined whether the relationship between visual flicker performance and severity of illness was due to flicker sensitivity or to response criterion.

In the present study severity of illness was not related to visual flicker performance because the correlations between self-rating on the depression scale and the variables of d'e and log beta 3 were not significant ($p > .05$). Further, the correlations between scores on the Hamilton Depression Rating and d'e and log beta 3 also were not significant ($p > .05$). For the major depressive patients the (Hamilton scale) correlations were; r (15) = -0.33 (d'e) and r (14) = -0.01 (log beta 3). For the dysthymic patients they were; r (3) = 0.25 (d'e) and r (3) = 0.16 (log beta 3).

Finally, a possible confounding factor in the present study relates to the log beta measure itself. In order to compare the log beta measure among different subject groups it is assumed that the underlying variance of the noise and signal plus noise distributions are approximately equal across groups. Since the slope of the ROC curve is directly related to the underlying variances (McNicol, 1972), a one way analysis of variance was performed on that measure. The slope of each individual's rating scale ROC curve was calculated. The mean slopes of those curves were 0.74 (*SD* = 0.23), 0.88 (*SD* = 0.07), and 0.72 (*SD* = 0.20) for the major depressive patients, dysthymic patients, and normal subjects, respectively. The analysis of variance was not significant (*F* (2, 22) = 1.11, *p* > .05) indicating that the variances of the underlying noise and signal plus noise distributions were comparable.

To summarize, it is likely that the findings concerning response criterion differences between the major depressive patients and the other groups was due to a different decision-making strategy adopted by that group and not due to age, severity of illness, or statistical artifacts associated with the log beta 3 measure.

In a signal detection study of visual sustained attention, Nuechterlein, Parasuraman, and Jiang (1983) stated that:

> "... certain clinical disorders may be characterized by altered response criteria rather than by sensitivity differences. These distinctions may provide theoretically more useful separations for clinical and psychopharmacological purposes than the more global findings of 'attention deficits' characteristic of much work on sustained attention among clinical populations" (p. 329).

The present study appears to illustrate the validity of that statement for differences in response criterion between major depressive patients and normal subjects using a visual flicker discrimination task.

If affective illness is considered as a continuum ranging from depression to mania, then maybe a cautious and conservative attitude is not inconsistent with the dysphoric mood, feelings of guilt, and generally pessimistic outlook characteristic of depressive disorders. By adopting a cautious attitude, depressed people may be able to avoid taking the risks often necessary for success in experimental or social situations. Conversely, at the other end of the spectrum of affective illness, manic patients characteristically behave in an irresponsible, grandiose manner; their mood is expansive; and their outlook is usually overly optimistic. During a manic phase, these patients often exhibit poor judgment by going on buying sprees or becoming sexually indiscrete. It is hypothesized that manic patients would adopt a liberal response criterion on psychological tasks (i.e., be biased in the opposite direction from depressed patients, cf. McAllister, 1980). In fact, Perris (1966) reported that both the flicker and fusion thresholds of the bipolar--manic patients were higher (though not statistically significantly so) than that of the normal subjects. That finding might be due to a bias towards "flicker" responses exhibited by the manic patients.

When stimulus events occur with equal probability, the optimal decision-making strategy is to adopt an unbiased response criterion. This tendency is illustrated in the present study by both the normal subjects and the dysthymic patients whose log beta 3 values were very close to zero, which indicates an unbiased response criterion. These groups are approximately midway between the overly cautious attitude of the major depressive group and the presumed liberal attitude of manic patients.

The deficits that most depressed patients exhibit on various cognitive, motor, sensory, and perceptual tasks (Miller, 1975) can be partially explained by the more cautious attitude adopted by this group. Any interpretations concerning a psychological deficit in depression must consider the differences in the decision-making strategy between depressed and normal people and the differences among people with various subtypes of depressive illness.

The fact that in the present study flicker processing differences among the groups was not found deserves comment. One possibility is that the flicker task as used in the present study was not sensitive enough to detect sensory sensitivity differences between groups. There are several reasons why this possibility seems unlikely.

First, the nature of the stimulus waveform used in the present study was one that is known to provide the most sensitive flicker thresholds. Maruyama (1976) studied the relationship between stimulus waveform and visual flicker in normal subjects. Rectangular waveforms yielded more sensitive flicker thresholds than either sinusoidal, triangular, or saw-tooth waveforms. Thus, the stimulating conditions of the present study can be considered to have been maximal to elicit differences in flicker sensitivity.

A second factor which argues that the procedures of the present study were sensitive enough to detect patient-normal flicker differences is the choice of d'e as the measure of visual flicker processing. The parametric d'e measure which makes certain assumptions concerning the relationship between the shapes of the underlying noise and signal plus noise distributions has been shown to be a reliable, valid measure of sensory sensitivity (McNicol, 1972). Those assumptions were met by the data in the present study; thus the use of the d'e measure was justified. Any patient-normal differences should have been detected by this measure.

A third factor is the systematic diagnosis of the depressed patients. The patients in the present study were diagnosed according to DSM-III criterion (American Psychiatric Association, 1980) based on a clinical interview using the SADS (Endicott & Spitzer, 1978). The DSM-III is a symptomatic approach to classification which also includes elements of family history and possible etiological factors (Andreasen, 1982; Roth & Barnes, 1981). Although far from perfect (Charles, 1982), the DSM-III represents the state-of-the-art in psychiatric classification. Also, the diagnosis was made by research psychiatrists for research purposes and therefore they tended to reject those patients who did not meet strict, well-defined criteria for major depressive or dysthymic disorder. This resulted in more homogeneous patient groups (especially the dysthymic group, see below).

In summary, patient-normal sensory processing differences, if they existed, should have been accentuated by the fact that the present study used a stimulus which delivered rectangular light stimuli, used a parametric measure of flicker sensitivity, had homogeneous patient groups, and screened the normal control group for signs of psychopathology.

The finding that the major depressive group had the highest mean d'e values for both the rating scale and response time procedures (Table 2) also deserves some comment. As Table 2 indicates, the major depressive patients were more variable than the other groups. In fact, the subject with the highest d'e score (3.93, rating scale) and the subject with the lowest d'e score (0.29, rating scale) were both major depressive patients. If the one major depressive patient who never responded "fused" to the 16 Hz stimulus (i.e., never missed) is not included in the calculations, the mean d'e score for that group drops from 1.78 (SD = 1.03) to 1.57 (SD = 0.79) for the rating scale procedure and from 1.71 (SD = 1.03) to 1.48 (SD = .75) for the response time procedure. Therefore the fact that the major depressive group had the highest mean d'e values was probably due to large variability within that group.

Interestingly, with respect to variability, the standard deviations of the dysthymic patient group were almost always lower than those of the major depressive patient group and were approximately equal to or lower than those of the normal control group on the measures of log visual threshold (Table 1); rating scale and response time d'e (Table 2); rating scale and response time log beta (Tables 3 and 6); false alarm rate (Table 4); self-ratings (Table 7); and the slope of the rating scale ROC curve. This suggests that the dysthymic patients, compared especially to the major depressive group, consituted a more homogeneous diagnostic group worthy of further investigation.

In the present study, there was no difference on the log visual threshold measure between the major depressive patients and the normal control subjects (Table 1). However, Mannuzza et al. (1980) reported that schizophrenic and major depressive patients needed about twice as much stimulus energy as normal control subjects to detect a brief visual stimulus. The optical bench, stimulus conditions, and psychophysical procedures were very similar between the Mannuzza et al. (1980) and the present study. (Although Mannuzza et al. did not use an artificial pupil, this probably did not account for the large threshold differences among the groups found in that study). The main difference between the two studies is that Mannuzza et al. reported that, of 20 factors of psychopathology, visual threshold was significantly ($p < .05$) correlated with only 1 factor; auditory hallucinations. The patients in the present study displayed neither auditory hallucinations nor any other psychotic symptoms. Thus, the presence of psychotic symptoms in the patients of the Mannuzza et al. study may partially explain the different results between that study and the present study.

Comparison Between the Rating Scale and Response Time Procedures

The finding that the three subject groups do not differ in visual flicker sensitivity is true of both the rating scale and the response time procedures. All mean d'e values were lower in the response time procedure (Table 2). However, none of the t-tests for related measures between rating scale and response time d'es was significant (t (10) = 0.82; t (4) = 2.62; t (8) = 2.28; for the major depressive patients, dysthymic patients, and normal control subjects, respectively. All p values were greater than .05). The response time procedure yielded lower d'e values in 6 of 11 major depressive patients, 4 of 5 dysthymic patients, and 8 of 9 normal subjects. This result is consistent with other studies which found that the response time procedure tended to yield lower sensory sensitivity scores (d'e or equivalent) in some subjects (e.g., Katz, 1970; Moss et al., 1979; Norman & Wickelgren, 1969; Weintraub & Fidell, 1979). The similarity of conclusions concerning visual flicker sensitivity and depression, the finding that the order of mean d'e values are identical (Table 2), and the high correlation between rating scale and response time d'e indicate a close similarity between these two d'e measures of flicker sensitivity.

Based on the results of the present study and those of Moss et al. (1970) and Katz (1970), it is concluded that the response time procedure is a useful and reliable method for determining sensory sensitivity values. For certain clinical populations it may be desirable to generate response time ROC curves. This procedure has the virtue of using a simple motor response (e.g., a finger lift) and not requiring the subject to remember response categories. It also has the virtue of facilitating data analysis because omitted or unused response categories occur less frequently than in the rating scale procedure. For example, in the present study the rating scale ROC curve for major depressive patient #120 was based on only one data point whereas her response time ROC curve was based on two points. This is important because an ROC curve based on two points much more accurately describes the subject's behavior than a curve based on only one point (McNicol, 1972).

The issue of the comparability of the rating scale and response time procedures with respect to measures of response criterion raises the issue of what is the best measure of response criterion. Previous studies are not helpful on this point because none of them directly compared response criterion measures derived from rating scale and response time ROC curves. The false alarm rate is simply the number of times subjects respond "flicker" to the fused stimulus, regardless of their response confidence. Therefore, the false alarm rate will always be identical for the two procedures.

In the present study, none of the measures of response criterion showed significant differences between the dysthymic patients and the normal control subjects. With respect to the rating scale data, major depressive patients had significantly higher log beta values than normal control subjects only at the third data point. At no data point did the major depressive patients differ from the dysthymic patients. With respect to the response time data, major depressive patients had significantly higher log beta values than normal subjects at five data points (including data point 10, which directly corresponds with data point 3 on the rating scale ROC curve) and higher log beta values than dysthymic patients and normal subjects at three data points. The false alarm rate of the major depressive patients was significantly lower than both other groups.

The finding that the major depressive patients adopt a significantly more conservative response criterion than normal subjects is clear. However, do the major depressive patients adopt a more conservative criterion than dysthymic patients? The false alarm rate and virtually all log beta values for the rating scale and response time ROC curves indicate that the answer to this question is yes (although statistical significance was achieved only for the false alarm rate and for only three response time log beta values). Therefore, further research on the relationship between major depression and dysthymia is needed. For example, such research could manipulate response criterion by different instruction conditions, by changing the probability of stimulus occurrence, or by varying payoffs for correct responses. Until more research clarifies the relationship between the different types of depressive disorders and response criterion, it seems best to report both the false alarm rate and the log beta values at each data point on an ROC curve.

Test-retest Analysis

For both the normal control subjects and the major depressive patients, comparisons of the differences of the rating scale d'e and log beta 3 measures for Sessions 1 and 2 were not statistically significant. (Statistical tests were not appropriate for the dysthymic patients because only two of them returned for Session 2.) Inspection of each individual's data indicated that, for most normal subjects, visual flicker performance on both these measures remained relatively stable over two experimental sessions. In contrast, for the major depressive patients, comparisons of these measures between Sessions 1 and 2, though not statistically significant (a result consistent with the findings of Holmberg, 1981 and Karp & Pollack, 1963), do show some interesting trends (Table 9). Interesting Session 1-Session 2 trends were also shown by the two dysthymic patients.

Two of the four major depressive patients taking phenelzine had higher rating scale d'e values in Session 2 (patient # 113 went from a d'e of -0.25 in Session 1 to one of 1.99 in Session 2. This represents a large change in performance) and all four patients taking imipramine had higher d'e values in Session 2.

With one exception (patient #113), all major depressive patients, regardless of medication, had lower log beta 3 values (were less conservative) in Session 2, possibly due to some unknown effect of antidepressant medications on response criterion.

Interestingly, the rating scale d'e values of the two dysthymic patients decreased in Session 2 (patient #118 went from a d'e value of 0.81 in Session 1 to one of -0.04 in Session 2, a considerable drop in performance; patient #124 went from a d'e of 0.86 to 0.57). Both dysthymic patients were taking imipramine prior to Session 2, thus the trend in flicker sensitivity of the dysthymic patients is in the opposite direction of the trend of the major depressive patients. Patient #118 had a log beta 3 value of .05 in Session 1 and a value of -.02 in Session 2, consistent with the trend of the major depressive patients taking imipramine. Patient #124 had a log beta 3 value of -.21 in Session 1 and a value of -.10 in Session 2, a result that is inconsistent with the trend of the major depressive patients taking imipramine.

There is some evidence that neurotic or atypical depressive patients are more likely to respond clinically to monoamine oxidase inhibitors such as phenelzine while endogenous depressives may be more likely to respond to tricyclic antidepressants such as imipramine (Quitkin, Rifkin, & Klein, 1979). The effect of imipramine and phenelzine on visual flicker sensitivity and response criterion in depressed patients has never been systematically investigated. Given the different trends in d'e and log beta 3 values between major depressive and dysthymic patients found in the present study with respect to imipramine, and given the different trends found between imipramine and phenelzine, the signal detection visual flicker technique may prove to be a useful tool in elucidating the behavioral and physiological actions of these medications on major depressive and dysthymic patients.

Suggestions for Future Research

The fact that in the initial sample, 39% of the major depressive patients and 40% of the normal control subjects had to be eliminated because their flicker discrimination was at a chance level of performance was unfortunate and disturbing. The present study was designed to yield d'e values in the 1.0 to 1.5 range, but clearly the task was too difficult for a great number of subjects. The large individual differences in visual flicker sensitivity found in the main experiment were not anticipated from the data of the pilot study. In retrospect, a better design might have been to test more pilot subjects (including some depressed patients) and also, prior to each experimental session, to determine a criterion-free measure of flicker sensitivity using an adaptive forced-choice procedure similar to the visual threshold procedure described in this study. The frequency values of the signal detection discrimination task could then be adjusted re sensation level such that a sufficient number of hits and false alarms are generated for an ROC curve which will indicate less than perfect but greater than chance performance. Note that a flicker sensitivity measure can be derived from both the adaptive forced-choice and the signal detection procedures but that only the signal detection procedure yields a measure of response criterion.

A variable worth exploring in future studies is the effect of speech or psychomotor retardation on visual flicker performance. In two studies by Bruder et al. (1975; 1980) depressed patients were found to be less sensitive in detecting the presence of a transient auditory signal than were schizophrenic patients or non-patient control subjects. That finding could not be explained in terms of group differences in response criterion because a criterion-free forced-choice method was used. In a third study by Bruder et al. (1981) depressed patients did not differ from normal control subjects in their thresholds for detecting a single click in either ear. The apparent discrepancy among the findings of those studies was discussed in terms of the symptom of speech retardation. Bruder et al. (1981) pointed out that the depressed patients in the earlier studies showed high symptom ratings of speech retardation whereas the patients in the later study displayed less evidence of that symptom. Other symptoms were not correlated with poor detection performance.

Speech retardation was not measured in the present study but a careful review of each patient's chart and general observation of each patient before and after testing failed to show evidence that speech retardation was a prominent symptom of any of the patients tested. Psychomotor retardation was measured in this study by response time and reaction time. Compared to the normal control subjects, no evidence of increased psychomotor retardation was found in either group of depressed patients. Significantly slower reaction times in depression is characteristic of only a few endogeneously depressed patients (Miller, 1975). Glass et al. (1981)

found no difference in speed of a finger lift reaction time to the word "LIFT" displayed on a cathode ray tube between a group of depressed outpatients similar to the ones tested in the present study and a normal control group.

Perhaps visual flicker sensitivity differences, like auditory detection differences, will be discovered between depressed patients with and without psychomotor retardation and between normal subjects and depressed patients with that symptom.

McAllister (1981) in a review of a study by Henry, Weingartner, and Murphy (1973) suggested that manic patients make more errors of commission than of omission on a variety of memory function tasks. That is, they tend to make many false alarms indicating a liberal response criterion. A signal detection study of visual flicker comparing major depressive patients to manic patients may yield very interesting data with respect to sensory processing in these two patient groups.

Finally, future investigations which replicate and extend the findings of the present study by using other measures of temporal resolution (e.g., the two-flash threshold) and different stimulus materials (e.g., signal detection analysis of "affective" and "neutral" word recognition memory) might help to elucidate some of the mechanisms underlying depressed patient-normal subject differences in information processing.

Conclusion

The major conclusion of the present study is that previously reported visual flicker differences between depressed patients and normal control subjects were probably due the different decision-making criterion used by the patients and not due flicker sensitivity differences between the groups.

Another conclusion is that, when groups of subjects are compared to each other, the rating scale and response time procedures yield identical conclusions with respect to visual flicker sensitivity for both normal subjects and psychiatric patients. Also, in most subjects, the response time procedure tends to yield lower flicker sensitivity scores than the rating scale procedure. The comparability of these two procedures with respect to response criterion is less clear. If future researchers are to substitute the response time for the rating scale procedures because the former is an easier task, then it is suggested that the false alarm rate as well as log beta values at each data point be reported.

Appendix

Filter calibration. Calibration of the filter combinations provided by the intensity programmer and fixed filter holder were determined within the optical system using photomultipliers 1 and 2 (see Apparatus section above). Each photomultiplier was connected to its own resistor box (Shallcross) such that photomultiplier voltage gain was proportional to resistance. The resistance boxes provided inputs to an oscilloscope high-gain differential preamplifier (Tektronix, Type D).

The glow modulator tube was pulsed on for 10 msec every second. With no filters between the two photomultipliers, the resistance of photomultiplier 1 was set at 100 kilohms and the resistance of photomultiplier 2 was adjusted to "null" the voltage difference between the outputs of the photomultiplier resistor boxes. After filters were introduced between the photomultipliers, the resistance of photomultiplier 1 was reduced to again "null" the voltage difference. This procedure was repeated three times for each filter combination.

The ratio of final resistance of photomultiplier 1 to 100 kilohms (multiplied by 100) indicates the per cent transmission for that filter combination. Per cent transmission was corrected for the input impedance of the oscilloscope preamplifier using the following formula (R equals resistance in ohms):

$$\text{per cent transmission} = 1.0125/((1/R \times 10^5) + .0125) \times 100$$

$$\text{density} = \log (100/\text{per cent transmission})$$

Footnotes

1. Note that all normal control subjects were paid but neither the major depressive nor the dysthymic patients were paid. This difference in payment structure for the two groups probably did not effect the performance of the patients because, as a group, major depressive patients actually had higher flicker sensitivity values than did the normal control subjects (see Results chapter). The flicker sensitivity scores of dysthymic patients were not significantly different from the other groups. However, lack of payment may have contributed to the reluctance of many of the patients to return for a second session.

2. The average luminance of the two stimuli were equated using the Talbot-Plateau law as defined by the following equation:

$$AL = L \times (tl / tl + td)$$

where:

AL = average luminance over time
L = absolute luminance of light
tl = duration of each separate light flashes
td = duration of the dark interval between light flashes

(see Riggs, 1971).

Note that, when asked, none of the subjects reported being able to discriminate the 16 Hz from the 116 Hz stimulus on the basis of brightness.

3. Log beta 3 values derived from rating scale ROC curves and log beta 10 values derived from response time ROC curves are equivalent (Tables 3 and 6) because both points are obtained by cumulating all of the "flicker responses" (separately from 16 Hz and 116 Hz trials); both represent the point just before the "fused" response (Norman & Wicklegren, 1969). Log beta 3 and log beta 10 are equivalent to beta values obtained from binary signal detection data (i.e., yes-no data) where ROC curves are not generated and d' and beta are based on a single data point.

Reference Notes

1. Clark, W. C., & Mehl, L. *Thermal pain: Sensitivity (d') and criterion (Lx differences between psychiatric patients and normals*. Paper presented at the meeting of the International Congress of Psychology, Paris, July, 1976.

2. Krooss-Glover, B. *Screening evaluation form*. Unpublished evaluation form, 1979. (Available from Dr. M. L. Kietzman, Department of Psychophysiology, New York State Psychiatric Institute, 722 West 168 Street, New York, N. Y. 10032).

3. Bruder, G. E. Personal communication, May, 1981.

Bibliography

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (3rd Edition). Washington, D. C.: Committee on Nomenclature and Statistics, 1980.

Andreasen, N. C. Concepts, diagnosis, and classification. In E. S. Paykel (Ed.), *Handbook of affective disorders*, New York: Guilford, 1982.

Anisman, H. & Zacharko, R. M. Depression: The predisposing influence of stress. *Behavioral and Brain Sciences*, 1982, *5*, 89-137.

Barry, J. R. Prognosis and patient response to flickering light. *Journal of Clinical Psychology*, 1962, *18*, 447-450.

Beck, A. T., Ward, C. H., Mendelson, M., Mack, J., & Erbaugh, J. An inventory for measuring depression, *Archives of General Psychiatry*, 1966, *4*, 561-571.

Black, S., Franklin, L. M., deSilva, F. P. R., & Wijewickrama, H. S. D. S. The flicker-fusion threshold in schizophrenia and depression. *New Zealand Medical Journal*, 1975, *81*, 244-246.

Bobon, D. P., Ott, H., & Holmberg, G. (Eds.), Critical flicker-fusion frequency in man: Validity, reliability, and sensitivity in psychopharmacology. *Pharmacopsychiatria*, 1982, *15*, Special Issue 1.

Bross, M., Harper, D. & Sicz, G. Visual effects of auditory deprivation: Common intermodal and intramodal factors. *Science*, 1980, *207*, 667-668.

Bruder, G., Spring, B., Yozawitz, A., & Sutton, S. Auditory sensitivity in psychiatric patients and non-patients: monotic click detection. *Psychological Medicine*, 1980, *10*, 133-138.

Bruder, G. E., Sutton, S., Babkoff, H., Gurland, B. J., Yozawitz, A., & Fleiss, J. L. Auditory signal detectability and facilitation of simple reaction time in psychiatric patients and non-patients. *Psychological Medicine*, 1975, *5*, 260-272.

Bruder, G., Sutton, S., Berger-Gross, P., Quitkin, F., & Davies, S. Lateralized auditory processing in depression: Dichotic click detection. _Psychiatry Research_, 1981, _4_, 253-266.

Buhler, R. A. Correlates of perceptual change: C. Stress. In W. H. Ittelson & S. B. Kutash (Eds.), _Perceptual changes in psychopathology_. New Brunswick, N. J.: Rutgers University Press, 1961.

Byrne, D. G. Affect and vigilance performance in depressive illness. _Journal of Psychiatric Research_, 1977, _13_, 185-191.

Campbell, R. A. Detection of a noise signal of varying duration. _Journal of the Acoustical Society of America_, 1963, _35_, 1732-1737.

Cegalis, J. A., & Deptula, D. Attention in schizophrenia: Signal detection in the visual periphery. _Journal of Nervous and Mental Disease_, 1981, _169_, 751-760.

Charles, S. C. Pitfalls in the diagnosis of patients with affective disorders. _American Journal of Psychotherapy_, 1982, _36_, 4-13.

Christie, M. J., Little, B. C., & Gordon, A. M. Peripheral indices of depressive states. In H. M. van Praag (Ed.), _Handbook of biological psychiatry_. New York: Dekker, 1980.

Clark, W. C. The psyche in psychophysics: A sensory-decision analysis of the effect of instructions on flicker sensitivity and response bias. _Psychological Bulletin_, 1966, _65_, 358-366.

Clark, W. C., Brown, J. C., & Rutschmann, J. Flicker sensitivity and response bias in psychiatric patients and normal subjects. _Journal of Abnormal Psychology_, 1967, _72_, 35-42.

Clark, W. C., Rutschmann, J., Link, R. & Brown, J. C. Comparison of flicker-fusion thresholds obtained by the methods of forced-choice and limits on psychiatric patients. _Perceptual and Motor Skills_, 1963, _16_, 19-30.

Davis, G. C., Buchsbaum, M. S., & Bunney, W. E. Analgesia to painful stimuli in affective illness. _American Journal of Psychiatry_, 1979, _136_, 1148-1151.

Douthwaite, W. A., & Williams, P. Critical flicker frequency technique and interpretation using a new flickering source. British Journal of Physiological Optics, 1974, 29, 30-42.

Duffy, E. & Lacey, O. L. Adaptation in energy mobilization: Changes in general level of palmer conductants. Journal of Experimental Psychology, 1946, 36, 437-452.

Dusoir, A. E. Treatments of bias in detection and recognition models: A review. Perception & Psychophysics, 1975, 17, 167-178.

Emmerich, D. S., Gray, J. L., Watson, C. S., & Tanis, D. C. Response latency, confidence, and ROCs in auditory signal detection, Perception & Psychophysics, 1972, 11, 65-82.

Endicott, J., & Spitzer, R. L. A diagnostic interview--the Schedule for Affective Disorders and Schizophrenia. Archives of General Psychiatry, 1978, 35, 837-844.

Falk, J. L., & Kline, D. W. Stimulus persistence in CFF: Overarousal or underactivation? Experimental Aging Research, 1978, 4, 109-123.

Frith, C. D. Consciousness, information processing and schizophrenia. British Journal of Psychiatry, 1979, 134, 225-235.

Frith, C. D. The interaction of noise and personality with critical flicker fusion performance. British Journal of Psychology, 1967, 58, 127-131.

Ginsberg, N. Flicker fusion bibliography, 1953-1968. Perceptual and Motor Skills, 1970, 30, 427-482.

Glass, R. M., Uhlenhuth, E. H., Hartel, F. W., Mayuzas, W., & Fischman, M. W. Cognitive dysfunction and imipramine in outpatient depressives. Archives of General Psychiatry, 1981, 38, 1048-1051.

Goldman, P. S., Lodge, A., Hammer, L. R., Semmes, J., & Mishkin, M. Critical flicker frequency after unilateral temporal lobectomy in man. Neuropsychologia, 1968, 6, 355-363.

Grandjean, E., Baschera, P., Martin, E., & Weber, A. The effects of various conditions on subjective states and critical flicker frequency. In R. R. Mackie (Ed.), Vigilance: Theory, operational performance, and physiological correlates. New York: Plenum, 1977.

Gruzelier, J. H., & Venables, P. H. Two-flash threshold, sensitivity and beta in normal subjects and schizophrenics. Quarterly Journal of Experimental Psychology, 1974, 26, 594-604.

Hamilton, M. A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry, 1960, 23, 56-62.

Harper, D. W. Signal detection analysis of effect of white noise intensity on sensitivity to visual flicker. Perceptual and Motor Skills, 1979, 48, 791-798.

Henry, G. M., Weingartner, H., & Murphy, D. L. Influence of affective states and psychoactive drugs on verbal learning and memory. American Journal of Psychiatry, 1973, 130, 966-971.

Hollister, L. E. Psychiatric disorders. In G. S. Avery (Ed.), Drug treatment (2nd Edition). Balgowlah, Australia: ADIS, 1980.

Holmberg, G. Critical flicker fusion (CFF) test for sedative effect of antidepressants. Acta Psychiatrica Scandinavica, 1981, 63 (supplementum 290), 289-301.

Irvine, R. P. Critical flicker frequency for paretics and schizophrenics. Journal of Abnormal and Social Psychology, 1954, 49, 87-88.

Johannsen, W. J., Friedman, S. H., & Liccione, J. V. Visual perception as a function of chronicity in schizophrenia. British Journal of Psychiatry, 1964, 110, 561-570.

Johnson, F. N. Depression: Some proposals for future research. Diseases of the Nervous System, 1975, 36, 228-232.

Johnson, F. N. Psychological approaches to lithium therapy. In T. B. Cooper, S. Gershon, N. S. Kline, & M. Shou (Eds.), Lithium: Controversies and unresolved issues. Amsterdam: Excepta Medica, 1979. (a)

Johnson, F. N. The psychopharmacology of lithium. Neuroscience & Biobehavioral Reviews, 1979, 3, 15-30. (b)

Johnson, F. N. The use of fish in studying the behavioural effects of lithium. Pharmacopsychiatria, 1981, 14, 202-212.

Karp, E., & Pollack, M. Comparative studies of chlorpromazine and imipramine. III. Critical flicker fusion thresholds in psychiatric patients. Psychopharmacologia, 1963, 4, 452-458.

Katz, L. A comparison of type II operating characteristics derived from confidence ratings and from latencies. Perception & Psychophysics, 1970, 8, 65-68.

Kerkhof, G. A., van der Schaaf, T. W., & Korving, H. J. Auditory signal detection : Effects of long-term practice and time on task. Perception & Psychophysics, 1980, 28, 79-81.

Kietzman, M. L., & Gillam, B. J. Visual temporal integration and simple reaction time. Perception & Psychophysics, 1972, 11, 333-340.

Kietzman, M. L., Spring, B., & Zubin, J. Perception, cognition, and attention. In H. I. Kaplan, A. M. Freedman, & B. J. Sadock (Eds.), Comprehensive textbook of psychiatry, III. Baltimore: Williams & Wilkins, 1980.

King, H. E. Two-flash and flicker-fusion thresholds for normal and schizophrenic subjects. Perceptual and Motor Skills, 1962, 14, 517-518.

Koerner, F., & Teuber, H. L. Visual field defects after missle injuries to the geniculo-striate pathway in man. Experimental Brain Research, 1973, 18, 88-113.

Landis, C. An annotated bibliography of flicker fusion phenomena covering the period 1740-1952. Ann Arbor, Mich.: University of Michigan Press, 1953.

Landis, C. Determinants of the critical flicker-fusion threshold. Physiological Reviews, 1954, 34, 259-286.

Laux, G., & Streichenwein, S. M. Treatment of therapy-resistant depressed patients with I.V. infusions of nominfensine controlled by galvanic skin resistance. In J. Obiols, C. Ballus, E. Gonzalez Monclus, & J. Pujol (Eds.), Biological psychiatry today. Amsterdam: Elsevier/North Holland Biomedical Press, 1979.

Lawrence, J. R., McEwen, J., Stonier, P. D., & Pidgen, A. W. Pupil size and critical flicker fusion threshold: A reevaluation. Drug Development Research Supplement, 1982, _1_, 67-75.

Leigh, G. M. Alcohol and cigarette smoking effects on critical flicker frequency (Doctoral dissertation, University of Waterloo, 1980). _Dissertation Abstracts International_, 1980, _41_, 1142B-1143B.

Levitt, H., & Treisman, M. Control charts for sequential testing. _Psychometrika_, 1969, _34_, 509-518.

Maley, M. J. Two-flash threshold, flicker-fusion threshold, and skin conductance. _Psychonomic Science_, 1967, _9_, 633-634.

Malone, J. R. L., & Hemsley, D. R. Lowered responsiveness and auditory signal detectability during depression. _Psychological Medicine_, 1977, _7_, 717-722.

Mannuzza, S., Spring, B. J., Gottlieb, M. D., & Kietzman, M. L. Visual detection threshold differences between psychiatric patients and normal controls. _Bulletin of the Psychonomic Society_, 1980, _15_, 69-72.

Markowitz, J., & Swets, J. A. Factors affecting the slope of empirical ROC curves: Comparison of binary and rating responses. _Perception & Psychophysics_, 1967, _2_, 91-97.

Maruyama, K. Stimulus wave form and CFF. _Tohoku Psychologica Folia_, 1976, _35_, 115-121.

Matin, L. Use of the glow modulator tube for visual research. _American Journal of Psychology_, 1964, _77_, 650-651.

McAllister, T. W. Cognitive functioning in the affective disorders. _Comprehensive Psychiatry_, 1981, _22_, 572-586.

McDonough, J. M. Critical flicker frequency and the spiral aftereffect with process and reactive schizophrenics. _Journal of Consulting Psychology_, 1960, _24_, 150-155.

McNicol, D. _A primer of signal detection theory_. London: George Allen & Unwin, 1972.

Medina, R. F. Frontal lobe damage and flicker fusion frequency. _Archives of Neurology and Psychiatry_, 1957, _77_, 108-110.

Miller, E., & Lewis, P. Recognition memory in elderly patients with depression and dementia: A signal detection analysis. Journal of Abnormal Psychology, 1977, 86, 84-86.

Miller, J. G. Adjusting to overloads of information. Disorders of Communication, 1964, 42, 87-100.

Miller, W. R. Psychological deficit in depression. Psychological Bulletin, 1975, 82, 238-260.

Milner, A. D., Beech, H. R., & Walker, V. J. Decision processes and obsessional behaviour. British Journal of Social and Clinical Psychology, 1971, 10, 88-89.

Misiak, H. The flicker-fusion test and its applications. Transactions of the New York Academy of Sciences, 1967, 29, 616-622.

Moss, S. M., Myers, J. L., & Filmore, T. Short-term recognition memory of tones. Perception & Psychophysics, 1970, 7, 369-373.

Neuchterlein, K. H., Parasuraman, R., & Jiang, Q. Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. Science, 1983, 220, 327-329.

Norman, D. A., & Wickelgrun, W. A. Strength theory of decision rules and latency in retrieval from short-term memory. Journal of Mathematical Psychology, 1969, 6, 192-208.

Nunn, C. M. H. A model of the functional psychoses. Acta Psychiatrica Scandinavica, 1980, 62, 75-84.

Parsons, O. A., & Huse, M. M. Impairment of flicker discrimination in brain-damaged patients. Neurology, 1958, 8, 750-755.

Perris, C. A study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses. VII: Studies in perception: b) Flicker and fusion threshold. Acta Psychiatrica Scandinavica, 1966, 42 (supplementum), 102-117.

Perris, C., & d'Elia, G. Flicker and fusion threshold and anxiety scores in alcoholics. Behavioral Neuropsychiatry, 1971, 3, 19-21.

Pieron, H. [Vision in intermittent light.] In W. D. Neff (Ed. and L. M. Hurvich & D. Jameson, trans.), Contributions to sensory physiology (Vol. 1). New York: Academic Press, 1965.

Price, R. H., & Eriksen, C. W. Size constancy in schizophrenia: A reanalysis. Journal of Abnormal Psychology, 1966, 71, 155-160.

Quitkin, F., Rifkin, A., & Klein, D. F. Monoamine oxidase inhibitors: A review of antidepressant effectiveness. Archives of General Psychiatry, 1979, 36, 749-760.

Rappaport, M., Hopkins, H. K., Silverman, J. & Hall, K. Auditory signal detection in schizophrenics. Psychopharmacologia, 1972, 24, 6-28.

Rees, J. N., & Botwinick, J. Detection and decision factors in auditory behavior of the elderly. Journal of Gerontology, 1971, 26, 133-136.

Riggs, L. A. Vision. In J. W. Kling & L. A. Riggs (Eds.), Woodworth & Schlosberg's experimental psychology (3rd ed.). New York: Holt, Rinehart, & Winston, 1971.

Riklan, M., Levita, E., & Misiak, H. Critical flicker frequency and integrative functions in parkinsonism. Journal of Psychology, 1970, 75, 45-51.

Roth, M., & Barnes, T. R. E. The classification of affective disorders: A synthesis of old and new concepts. Comprehensive Psychiatry, 1981, 22, 54-77.

Roth, M., & Mountjoy, C. O. The distinction between anxiety states and depressive disorders. In E. S. Paykel (Ed.), Handbook of affective disorders. New York: Guilford, 1982.

Sartorius, N., Jablensky, A., Gulbinat, W., & Ernberg, G. WHO collaborative study: assessment of depressive disorders. Psychological Medicine, 1980, 10, 743-749.

Siegal, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Simonson, E., & Brozek, J. Flicker fusion frequency background and applications. Physiological Reviews, 1952, 32, 349-378.

Smith, J. M., & Misiak, H. Critical flicker frequency (CFF) and psychotropic drugs in normal human subjects--A review. Psychopharmacology, 1976, 47, 175-182.

Sternberg, S. High speed scanning in human memory. Science, 1966, 153, 652-654.

Stewart, R. A., Groff, C., & Kiker, V. L. Stress and flicker perception: I. The effects of shock induced anxiety. Perceptual and Motor Skills, 1971, 33, 1211-1214.

Teasdale, J. D., Fogarty, S. J., & Williams, M. G. Speech rate as a measure of short-term variation in depression. British Journal of Social and Clinical Psychology, 1980, 19, 271-278.

Teuting, P. Special report on depression research. (DHHS Publication No. (ADM)81-1085). Washington, D. C.: National Institute of Mental Health, 1981.

Titcombe, A. F., & Willison, R. G. Flicker fusion in multiple sclerosis. Journal of Neurology, Neurosurgery, and Psychiatry, 1961, 24, 260-265.

Turner, P. Flicker fusion threshold in mental illness. New Zealand Medical Journal, 1975, 82, 98. (Letter)

Wagoner, R. A. Differences in response latency and response variability between high and low anxiety subjects in a flicker-fusion task. Journal of Abnormal and Social Psychology, 1960, 61, 355-359.

Watson, C. G., Thomas, R. W., Felling, J., & Andersen, D. Differentiation of organics from schizophrenics with trail making, dynamometer, critical flicker fusion, and light-intensity matching. Journal of Clinical Psychology, 1969, 25, 130-133.

Weber, A., Fussler, C., O'Hanlon, J. F., Gierer, R., & Grandjean, E. Psychophysiological effects of repetitive tasks. *Ergonomics*, 1980, *23*, 1033-1046.

Weintraub, D. J., & Fidell, L. S. A signal-detection approach to subception: Concomitant verbal and finger-latency responses in metacontrast. *Perception & Psychophysics*, 1979, *26*, 143-152.

Whitehead, A. Verbal learning and memory in elderly depressives. *British Journal of Psychiatry*, 1973, *123*, 203-208.

Winer, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

Wynne, L. C., Cromwell, R. L., & Matthysse, S. (Eds.), *The nature of schizophrenia: New approaches to research and treatment*. New York: Wiley, 1978.

Zealley, A. K., & Aitken, R. C. B. Measurement of mood. *Proceedings of the Royal Society of Medicine*, 1969, *62*, 993-996.