7816692

LUPO, JOSEPH ANTHONY
A SIGNAL DETECTION COMPARISON OF TWO-PULSE MEASURES OF VISUAL TEMPORAL RESOLUTION.

CITY UNIVERSITY OF NEW YORK, PH.D., 1978

University Microfilms International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

A SIGNAL DETECTION COMPARISON OF TWO-PULSE MEASURES OF VISUAL TEMPORAL RESOLUTION

by

JOSEPH A. LUPO

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.

1978

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.

May 18, 1978 — date
Mitchell L. Kietzman — Chairman of Examining Committee

May 18, 1978 — date
Florence L. Denmark — Executive Officer

Mitchell L. Kietzman

Samuel Sutton

Wilfred A. Gibson

Supervisory Committee

The City University of New York

Abstract

A SIGNAL DETECTION COMPARISON OF TWO-PULSE MEASURES OF VISUAL TEMPORAL RESOLUTION

by

Joseph A. Lupo

Adviser: Professor Mitchell L. Kietzman

The purpose of this investigation was to compare signal detection rating scale measures for both sensitivity and response bias for two-pulse discrimination as "one-two" flash responses (the "one-two" condition) and as same-different" responses measured with reference to a single pulse comparison stimulus ("the same-different" condition). The "one-two" condition retained the most salient characteristics of classical psychophysical studies of two-pulse visual temporal resolution, i.e., a single observation interval, no comparison stimulus, "one-two" flash responses, and Class B observations. Unlike the classical psychopysical studies, feedback was given both single and double pulse stimuli were presented, and the sensitivity measures obtained were criterion-free. The "same-different" condition retained the most prominent features of forced-choice studies of two-pulse

resolution studies, i.e., at least one comparison and one test stimulus were presented on a single trial, Class A observations, and criterion-free sensitivity measures. Previous two-pulse temporal resolution studies have reported that the classical psychophysical threshold was more than double the value of the forced-choice threshold. The present study provided a signal detection analysis of these "threshold" measures.

Two trained observers participated in the main experiment. A single channel binocular free viewing optical system was used to present light pulses of 20 min visual angle to the dark-adapted observer. The luminance level (1.37 cd/m^2) was determined during preliminary testing using the UDTR procedure and was the level used in all aspects of this study. For the "one-two" condition the observer was presented with either a single 4 msec light pulse or two 2 msec light pulses separated by an interpulse interval. The observer was to report whether he saw "one" or "two" light flashes and to report the certainty of his response. For the "same-different" condition the observer was presented a 4 msec light pulse (the comparison stimulus) followed either by another 4 msec light pulse or two 2 msec light pulses separated by an interpulse interval. The observer was to report whether the second stimulus was "same" or "different" from the first stimulus. Interpulse interval durations of 20, 40, 60, 80 msec were used in both conditions. Each observer participated in 16 sessions of 256 trials.

For each observer two families of ROC curves were constructed. All of the curves were characterized by straight lines of nonunity slope. There were no significant differences in sensitivity and overall criterion between conditions. The minimum criterion setting for the report of "two" flashes remained constant as the interpulse interval was increased. This was interpreted as an indication that the observers were following the instructions given to them. The minimum criterion setting for the report of "different" for the "same-different" condition changed as the interpulse interval was increased; both observers became increasingly more lax as the interpulse interval was increased. This was interpreted as an effect of multiple cues for discrimination interacting with criterion. Sensitivity measures comparable in value to those obtained in the previous forced-choice studies were obtained from the "same-different" procedure, suggesting that these procedures were equivalent. This hypothesis was supported in a control study where sensitivity measures obtained from a two interval forced-choice procedure were compared directly and no significant differences were found. When the strictest criterion setting of the "one-two" condition was adjusted to provide "threshold" data comparable to that obtained in classical psychophysical studies the same "high threshold" was found. It was concluded that the large differences in threshold magnitude previously reported was due to an effect of criterion and not to differences in sensitivity.

Acknowledgements

I thank my dissertation sponsor, Mitchell L. Kietzman, for the innumerable and invaluable hours spent with me from the conception of this study to its completion. Certainly, he is most responsible for the completion of this study.

I also thank other members of my dissertation committee, Wilfred A. Gibson, Samuel Sutton, W. Crawford Clark, and Barbara Gillman for their advice and criticisms from which I have benefited.

A number of others also assisted me in this work: architect Micheal P. DiBella drafted the figures, Leslie Burton assisted in testing, Monica Graham was a very patient and understanding observer, Ray Simon and Roger Chin assisted with the apparatus, and Norma Byer provided useful advice for the typing of the manuscript.

I also thank my wife, Fina, and my parents, Joseph P. and Antoinette E. Lupo for their assistance, understanding, support, and love.

Table of Contents

List of Tables

List of Figures

I. Introduction

The major purpose of this investigation was to compare signal detection measures for both sensitivity and response bias for two-pulse discrimination as "one-two" flash responses using a rating scale procedure (without as comparison stimulus) and for two-pulse discrimination as "same-different" responses measured with reference to a single pulse comparison stimulus using a rating scale procedure. A comparison of signal detection measures for two-pulse discrimination as "same-different" responses measured with reference to a single pulse comparison stimulus using a rating scale procedure and two-pulse discrimination as "different" responses from single pulse comparison stimuli in a two interval forced-choice procedure was done in a control experiment in order to determine if these procedures would produce similar results. As will be seen, the procedures used in the present study had certain features in common with classical psychophysical and forced-choice studies of two-pulse visual temporal resolution.

Visual temporal resolution represents the discriminative power of the visual system in the time dimension. Psychophysical measures of the power have included the two-flash threshold and the comparison stimulus two-pulse threshold. The best distinction that can be made between these measures is one that has resulted from certain procedural aspects of past research. In studies that have used classical psychophysical methods, such as the method of limits (Dunlap, 1915; Mahneke, 1958; Purcell and Stewart, 1971), temporal resolution has been defined as the minimum period of time between two pulses of light for an observer to report them as two separated flashes. In these studies, the observer was presented a single stimulus interval (without a comparison) consisting of two pulses of light separated by an interpulse interval of varying duration. The point representing the minimal interpulse interval at which the observer reported "two-ness" 50% of the time was taken as the two'flash threshold.

In other studies that have employed more sophisticated psychophysical techniques such as the temporal forced-choice procedure (Kietzman, 1967; Kietzman & Sutton, 1968; Mertens & Lewis, 1972) or its adaptations (Lewis, 1967; Lewis, 1968; Lewis & Mertens, 1971), the observer was not asked to report whether or not two distinct flashes were seen but to report in which temporal interval some "difference" from the stimuli in the other intervals was observed. Here, temporal resolution was operationally defined as the point representing the minimal interpulse interval at which the observer could correctly discriminate a difference between a two-pulse test stimulus at a longer or shorter interpulse interval than a one-pulse comparison stimulus. In some studies (e.g., Lewis, 1967) a two-pulse comparison stimulus with a very brief, one msec interpulse interval was used in place of the one-pulse comparison stimulus. The 50% correct discrimination point (adjusted for chance) was taken as the two-pulse threshold.

In the forced-choice procedure, the observer need not actually see two distinct flashes of light; some other aspects of the test stimulus, other than its quality of "twoness" may have been inportant for the discrimination between stimuli, In the more classical procedures, the observer was asked to make a more perceptual response, based upon his sufjective ability to discriminate "one" from "two" light flashes. The distinction made here between the different kinds of response required of the observer from the two procedures is in accordance with the distinction of the terms "pulse" and "flash" which was first proposed by Bartley (1958, p. 99) who suggested that pulse refer to the stimulus and flash refer to the percept. The same terminology was also adopted in the resolution studies of Kietzman (1967) and Kietzman and Sutton (1968) to emphasize the fact that with the forced-choice procedure, it was not necessary for the observer to actually see two flashes in order to discriminate between stimuli.

This distinction of the kinds of response required of the observer in the clasical psychophysical and in the forced-choice studies of temporal resolution is related to and develops certain implications from Brindley's (1960, pp. 144-150) classification of Class A and Class B observations. Class B observations are those in which the observer must subjectively report on some quality or intensity of sensation. In the classical studies of temporal resolution, Class B observations are made because as the interpulse interval is manipulated the observer must report whether or not the stimulus presented produced the quality of "twoness". Class A observations are those in which the observer must report under certain experimental conditions whether or not a particular stimulus produced the same sensation or a different sensation. When the observer is permitted to use all possible cues to report a "difference" the observations made in the comparison stimulus two-pulse studies may be classified as Class A. The major implication, according to Brindley, is that only Class A observations can be used directly as a basis for analyzing the eye and the visual pathway. Class B observations are more likely to be confounded with subject variables such as response bias than Class A observations and are not as useful as Class A observations in providing observations which represent fundamental relations between stimuli and sensation. Class B observations become maximally useful only after they have been found to give the same results as Class A observations or when they are converted to Class A observations by being used under the same experimental conditions in which Class A observations were obtained.

Cornsweet and Pinsker (1965) have more recently added another standard for observations that can be used as a basis for vigorously testing physiological aspects of the visual system. They suggest that observations must reflect as closely as possible the effect of stimulus manipulations, and not be confounded with observer biases, as much as this can be controlled by the experimenter. One of the ways to achieve this is through the use of feedback. They state that feedback must be given to the observer immediately after he has made a Class A observation, on a trial-to-trial basis. Without feedback the data reflect the relation between stimulus manipulations and the observer's subjective interpretation of "difference" in a discrimination task. Although feedback does not eliminate this criterion, when it is given, the observer's criterion becomes whatever aspects of his experience are more reliably correlated with the stimulus manipulations. In applying Cornsweet and Pinsker's suggestion to Brindley's method of maximizing the use of Class B observations, it is apparent that Class B observations when obtained without giving feedback to the observer would be of less theoretical significance with regard to visual functioning. In this sense, results of the classical psychophysical studies of two-pulse temporal resolution would be of little theoretical value with regard to the functioning of the visual system for several reasons. First, there was no objective way for feedback to be given because on every trial a two-pulse stimulus was presented. If the observer said "two" flashes the experimenter had to rely on the observer's subjective report that two flashes were actually seen. Second, there was no clear evidence whether or not these Class B observations would produce results similar to studies in which Class A observations were obtained, i.e., the forced-choice studies. Third, the Class B observations obtained in the classical psychophysical studies had never been obtained under identical experimental conditions as the Class A observations obtained in the forced-choice studies. Furthermore, using Cornsweet and Pinsker's (1965) line of reasoning, the classical psychophysical studies of two-pulse temporal resolution provided more information regarding subjective aspects of "twoness" discrimination than the ability of the visual system to resolve two pulses of light.

In the main experiment of the present investigation signal detection measures of two-pulse discrimination as "one-two" flash responses using a rating scale procedure (the"one-two" procedure) and as "same-different" responses measured with reference to a single pulse comparison stimulus using a rating scale procedure (the "one-two" procedure) and as "same-different" responses measured with reference to a single pulse comparison stiumlus using a rating scale procedure (the "same-different" procedure) were compared. In the control study sensitivity measures for the "same-different" procedure and for two-pulse discrimination as "different" responses from single pulse comparison stimuli in a two interval forced-choice procedure (the "different" procedure) were compared. The "one-two" procedure produced Class B observations comparable to those obtained in the clasical psychophysical studies of two-pulse temporal resolution. Both the "same-different" and "different" procedures produced Class A observations comparable to those obtained in the previous forced-choice studies. The two procedures in the main experiment and the two procedures in the control experiment were conducted under identical viewing conditions and for the same sets of observers. Feedback was given in both the main and control experiments for all procedures used. In view of the ways Class B observations may become maximally useful, the Class B observations produced from the "one-two" procedure were designed to conform to several criteria for conversion to Class A observations.

Although the distinction between Class A and Class B observations may be of methodological and theoretical importance Brindley has noted that while some investigators believe that only Class A observations were of any value, others believe that the distinction useless; they believe that both Class A and Class B observations provide an equally valid means of analyzing the eye and the visual pathway. This controversy is evident in the visual temporal resolution literature. Most studies (e.g., Lewis, 1967; Lewis, 1968; Lewis & Mertens, 1971; Mertens & Lewis, 1972; Purcell and Stewart, 1971; Theodor, 1972) have made no distinction between the kind of observations that were obtained from the classical psychophysical and forced-choice procedures and have discussed the measures of temporal resolution obtained from these procedures as if they were identical. Other studies, particularly those of Kietzman (1967) and Kietzman and Sutton (1968), have made a clear distinction between the observations that were obtained from these two procedures and have suggested that the measures of temporal resolution obtained from the two procedures were actually rather different.

Perhaps one of the reasons why there was such controversy was that the two procedures had never been compared within a single experiment with identical viewing conditions for the same sets of observers. Kietzman (1967) and Kietzman and Sutton (1968) reported from a series of separate studies that the comparison stimulus two-pulse threshold obtained from the forced-choice procedure was less than half the value of the two-flash threshold, obtained from a classical psychophysical procedures (i.e., method of constant stimuli). The fact that the forced-choice procedure produced threshold reduction was not surprising (Blackwell, 1953), but what was rather surprising was the drop in threshold of such a large magnitude (from about 73 msec to about 31 msec, Kietzman & Sutton, 1968). Kietzman and Sutton (1968) stated that the forced-choice procedure facilitated discrimination by allowing the observer to respond to characteristics of the stimulus other than the presence of the interpulse interval. One of these characteristics was the total stimulus duration;

as the interpulse interval was increased the total stimulus duration was increased as well. Kietzman and Sutton (1968) attempted to remove this presumable facilitating effect by packaging stimuli so as to remove duration differences between a two-pulse test stimulus and a one-pulse comparison stimulus. The comparison stimulus was equal in duration to the interpulse interval of the test stimulus plus a constant of 4 msec (total light time of the test stimulus). Also, as comparison pulse duration was increased, its luminance was decreased in order to keep the energy constant and equal to that of the test stimulus. For any interpulse interval (12-60 msec), the elimination of the duration differences between comparison and test stimuli resulted in reduced levels of discrimination, as compared to comparable data in which the comparison stimulus was not matched to the duration of the test stimulus. Kietzman and Sutton concluded that the total stimulus duration was an important contributing factor in two-pulse discriminations.

To show that not only the interpulse interval but also the total stimulus duration was a significant factor in two-pulse discriminations, Kietzman and Sutton (1968) compared the accuracy of discrimination of a 4 msec comparison pulse with either a test double-pulse of varying interpulse interval (2.5-40 msec) or a test single-pulse having the same total stimulus duration as that of the test double-pulse. All stimuli were equal energy stimuli. Results showed that as the total stimulus duration was increased the accuracy of discrimination for both sets of test stimuli was also increased; moreover, the rate of increased discrimination was almost identical. The studies of Kietzman and Sutton have shown that the processes underlying two-pulse discriminations can be extremely complex. They suggested that the large difference in threshold magnitude between the classical psychophysical and forced-choice temporal resolution studies could not be explained by the use of the forced-choice procedure alone, but by some interaction between the forced-choice method and the cues that are allowed by using it.

One of the problems often encountered in experimental studies in which there are multiple cues for discrimination is the stimulus attributes problem (see Stevens, 1934). The stimulus attributes problem refers to the the fact that a single stimulus may contain many discriminable attributes (cues) and that the attributes and not the sensations under investigation become the object of observational attention (Boring, 1963, p. 21). It is a methodological problem in the sense that the cues used by the observer in making his discriminations are not necessarily the ones that the experimenter is interested in or believes he is examing (Boynton, 1972). With respect to two-pulse discrimination, Kietzman and Sutton (1968) have demonstrated that in addition to the total stimulus duration of the stimulus package, a variety of other cues may be used by the observer in making his discriminations, i.e., the cues used by the observer in forced choice studies of temporal resolution are multiple cues. Kietzman and Sutton found that observers took advantage of the freedom permitted in the forced choice procedure to used a variety of cues such as "flick", "color", "shorter", "longer", and "other cues", as well as cues which the observers could not verbalize, "no report", to discriminate between stimuli, The forced-choice instructions did not specify the nature of the difference to be used by the observer in making his judgements. The forced-choice instructions actually encouraged multiple cues to be used and in this sense was tolerant of the stimulus attributes problem. On the other hand, in the classical psychophysical studies the observer was instructed to report on a specific sensory dimension, i.e., the "twoness" of the stimulus package; the use of any other stimulus attribute was not encouraged. However, the actual cues used by the observer are unknown to the experimenter since there is no guarantee that the instructions are followed by the observer (Kietzman & Sutton, 1968). A related problem with respect to stimulus attributes and the classical studies is that a single pulse of light may be seen as double (Dunlap, 1915; Bartley & Wilkinson, 1953).

Quantity of light. There is a large body of scientific literature concerned with the effects of manipulating pulse luminance or duration on two-pulse temporal resolution measures. One would tend to think that a survey of this literature would reveal some justification as to whether or not the distinction made between the measures of temporal resolution obtained from the classical psychophysical and forced-choice procedures in necessary. Unfortunately, because of the variety of experimental conditions under which these effects have been studied, comparison between studies is difficult. As will be shown in the survey that follows, very little is known about the effects of changing pulse luminance or duration on two-pulse measures of temporal resolution. Moreover, it cannot be determined from these studies whether or not these experimental manipulations produce different effects on measures of two-pulse temporal resolution obtained from classical psychopysical procedures and from forced-choice procedures.

One of the first investigations of visual temporal resolution was reported by Dunlap (1915). Using rather crude electronic equipment, Dunlap attempted to study the effect of pulse duration, pulse intensity, and dark adaptation on the two-flash threshold. His results may be summarized as follows: (a) increasing the duration of each of the two pulses equally lowered the threshold; (b) increasing the duration of the first pulse which keeping the duration of the second pulse constant reduced the threshold almost as effectively as (a) above; (c) increasing the duration of the second pulse while keeping the duration of the first pulse constant had practically no effect; (d) the effects of increasing intensity (brightness) were variable and it could not be concluded that increasing the intensity of the pulses lowered the two-flash threshold and (e) the threshold was found to be lower for the light-adapted eye than for the dark-adapted eye.

Mahneke (1958) repeated some of Dunlap's experiments using more sophisticated equipment such as a glow modulator as a light source and an electronic generator which permitted greater accuracy control of pulse duration. Mahneke's chief concern was the effect of manipulating pulse duration of the two-flash threshold. Three different sets of pulse combinations were used. In increasing the durations of the first and second pulses equally from 1 to 300 msec the threshold decreased from 58.6 to 3 msec. Transformation of this data indicated that interpulse interval duration approximated a linear function of log pulse duration. Increasing the duration of the first pulse while keeping the duration of the second pulse constant at 1 msec produced somewhat more complicated results. When the first pulse was increased in duration from 1 to 300 msec the threshold was increased from 57 to 79 msec; further increases in the duration of the first pulse decreased the threshold. Increasing the duration of the second pulse from 1 to 50 msec while keeping the duration of the first pulse at a constant duration greater than about 3 msec reduced the threshold from 56.5 to 16.1 msec; the two-flash threshold was independent of the second pulse duration when the first pulse duration was less than about 3 msec. These results led Mahneke to conclude that the capacity of the visual system to percieve a dark interval between two successive pulse of light was increased if within a given time interval the total quantity of light was increased. Thus, according to Mahneke, any increase in the total energy produced, such as an increase in pulse duration or in pulse luminance would improve discrimination. This hypothesis appeared to be supported with respect to Mahneke's findings concerning pulse duration but not supported when one considered Dunlap's (1915) findings concerning pulse luminance. As a result of finding that interpulse interval duration approximated a linear function of log interpulse interval duration, Mahneke suggested the applicability of the Ferry-Porter law to two-flash resolution thresholds. The Ferry-Porter law represents a formulation of the relationship between the quantity of light (total energy produced) and the briefest interpulse interval between a large number of successive pulses that could be discriminated or perceived as fused. Specifically, the law states that there is a linear relationship between the flicker fusion frequency and the log of the luminance ($f = \log I$).

In a general way, the findings of Clark (1963) tend to support Mahneke's "quantity of light" hypothesis. Clark found for both single dark interval (two-pulse) and multiple dark interval (flicker) stimuli, that as the stimulus luminance increased the duration of the just noticeable dark interval decreased. Furthermore, the improvement in temporal resolution was linearly related to log luminance. The slopes of the two-pulse and flicker functions were different; a given stimulus intensity became more effective at "revealing" a single dark interval for two-pulse stimuli, than at "revealing" multiple dark intervals for flicker stimuli, because the latter stimuli were much briefer.

The results of several other investigations have not supported the "quantity of light" hypothesis. Ireland (1954) examined the effect of varying three different light intensities on the two-flash threshold. The highest intensity was found to reduce the threshold insignificantly (by 10 msec), and two of the five observers were uneffected by the intensity manipulations. Ireland concluded that the Ferry-Porter law, applicable to certain ranges of luminance for CFF, was essentially not applicable to the two-flash threshold measure of resolution. Mahneke (1958) also failed to consider certain important implications of his hypothesis (Kietzman, 1967). Mahneke increased the total energy by increasing pulse duration but this did not necessarily mean that the amount of energy integrated by the nervous system was increased; the total stimulus duration or the duration of a single pulse of light may have been beyond the critical duration of Bloch's law. Also, the fact that shorter duration light pulses may appear brighter than longer duration light pulses of the same luminance (Broca-Sulzer effect) further complicated the situation in that the observer may have been responding on the basis of apparent brightness. Kietzman (1967) further investigated the two-flash threshold as a function of stimulus energy by comparing the effect of increasing pulse duration with the effect of increasing energy by increasing pulse luminance. Two observers were presented with equal stimuli (i.e., total stimulus energy, luminance x duration, of the various pairs of stimuli was the same) having durations and luminances of 4, 27, and 62 msec, and 40, 269, and 612 mL. (127, 856, and 1948 cd/m^2). One observer also viewed a 993 mL. (3160 cd/m^2) stimulus. For both observers as the duration of the stimulus increased at a fixed luminance the threshold dropped markedly (e.g., for one observer the reduction was from 73.8 to 22.2 msec at a luminance of 856 cd/m^2). However, an increase in pulse luminance produced only a slight decrease in the threshold (the greatest reduction was only 14.9 msec). The fact that these two operations, both designed to increase stimulus energy, produced different results placed further doubt upon the quantity of light hypothesis.

The effect that pulse duration may have on the two-flash threshold has been more thoroughly investigated in a series of experiments by Purcell and Stewart (1971). The pulse luminance was kept constant at 40 fL. (137 cd/m^2) throughout the series. In one experiment the threshold obtained from two 5 msec pulses was compared with that obtained from two 60 msec pulses. It was found, as Dunlap (1915), Mahneke (1958), and Kietzman (1967) had found, that increasing the duration of each pulse equally significantly reduced the threshold. In another experiment the effects of unequal pulse durations were investigated. Increasing the duration of the first pulse from 30 to 90 msec significantly reduced the threshold when the duration of the second pulse was constant at 30, 60, and 90 msec. This finding appeared to be in agreement with the results of Dunlap (1915) and Mahneke (1958). Increasing the duration of the second pulse when the duration of the first pulse was constant at 30, 60, and 90 msec also reduced the threshold in agreement with Mahneke's (1958) findings but not with Dunlap's (1915) findings. In addition, Purcell and Stewart found that decreasing the duration of the second pulse from 90 to 5 msec when the duration of the first pulse was 90 msec also reduced the threshold.

Kietzman (1967) was one of the first investigators to apply the temporal forced-choice procedure to studies of two-pulse resolution. Kietzman studied the effect of increasing pulse luminance on the comparison stimulus two-pulse threshold. The stimuli consisted of a 10 msec comparison single-pulse and a test double-pulse with a variable interpulse interval; each pulse had a duration of 5 msec. There were four observation intervals. In the first interval the comparison stimulus was always presented so that the observers task was to report in which interval, the second, the third, or the fourth, the test stimulus appeared; the position of the test stimulus was randomly varied from trial to trial. The other two observation intervals contained 10 msec single pulse stimuli. The comparison stimulus two-pulse threshold, as stated previously, was specified as the interpulse interval that produced 50% correct judgements after correction for chance. The experimental parameter was pulse luminance which was tested at 2.5, 40, 62, and 612 mL. (8.0, 127, 197, and 1948 cd/m^2). Threshold values for five observers were found to range from 11 to 31 msec and no systematic change in threshold was found as pulse luminance was increased.

Lewis (1967) also used a temporal forced-choice procedure to determine the effect of increasing pulse luminance on the comparison stimulus two-pulse threshold. There were two observation intervals, one in which a comparison stimulus consisting of a pair of 1 msec pulses separated by a 1 msec interpulse interval was presented, the other in which a test stimulus consisting of a pair of 1 msec pulses separated by a variable interpulse interval. The temporal position of the comparison and test stimuli was randomly varied from trial to trial. The pulse luminances were .32, 1, 10, 100, and 1000 mL. (1.02, 3.18, 31.82, 318.27, and 3182.70 cd/m^2). The observer's task was to report the temporal position of the "longer" stimulus. Lewis reported that the two-pulse threshold was a negatively accelerated function of increasing luminance. The largest reduction in threshold was found to occur between 1.02 and 3.18 cd/m^2; above 3.18 cd/m^2; there was little or no effect.

There may have been several reasons for the discrepancy between the results obtained by Kietzman (1967) and Lewis (1967). Kietzman used longer duration pulses than Lewis and investigated luminance changes over a different range. Pulse duration and pulse luminance were both variables that were known to effect the total stimulus energy level; hence, Kietzman and Lewis investigated the two-pulse threshold at different energy levels, at fact that has made their studies somewhat incomparable. Kietzman asked his observers to report on the stimulus which was "different" (a Class B observation). Furthermore, Kietzman did not inform his observers as to the accuracy of their responses while Lewis did. Finally, there were minor procedural differences that may have contributed to the different results. Lewis used an adaptive forced-choice procedure (BUDTIF; Campbell, 1963) whereas Kietzman did not.

Kramer (1969) employed the same procedure and equipment used by Kietzman (1967) and extended the luminance range from .013 to 788 mL. (.041 to 2508 cd/m^2) to include those luminances studied by Lewis (1967). The results were inconclusive in that one of the two observers showed a decrease in comparison stimulus two-pulse threshold with increasing luminance, while the other observer showed no significant change in threshold with increasing luminance.

Nilsson (1969) argued that Lewis (1967) found a marked luminance effect because his stimuli were all in the scotopic range of vision, whereas Kietzman's (1967) stimuli were all photopic. Nilsson reported that the comparison stinulus two-pulse difference threshold was not effected by changes in luminance when pulse luminance was in the photopic range of 50 to 2000 mL. (159 to 6365 cd/m^2) and pulse durations were each 1 msec.

Lewis and Mertens (1971) found however, a marked decrease in threshold values as the luminance was increased from 10 to 3000 mL. (31.8 to 9548 cd/m^2); a much larger range than was investigated by Nilsson. Mertens and Lewis (1972) replicat ed their findings for the luminance range of 31.8 to 3183 cd/m^2.

The manipulation of pulse luminance may have a different effect on the comparison stimulus two-pulse threshold than on the two-flash threshold. Kietzman and Smith (1965) applied the temporal forced-choice procedure to determine the effect of pulse duration on two-pulse resolution. The experimental conditions consisted of three test packages differing in total light times of 20, 34, and 56 msec. Each test package consisted of a one-pulse comparison stimulus and two-pulse test stimuli of the same luminance and total light time. Three different test stimuli within each test package were used: one consisted of a short duration first pulse and a long duration second pulse, another consisted of a long duration first pulse and a short duration second pulse, and another consisted of equal duration fcirst and second pulses. The independent variable was the interpulse interval. Kietzman and Smith found that within a test package a short duration first pulse and a long duration second pulse produced better discrimination than pulses of equal duration and that pulses of equal duration produced better discrimination than when the first pulse was of long duration and the second pulse was of short duration. The test condition with equal duration pulses produced a 50% adjusted correct discrimination level and therefore a comparison stimulus two-pulse threshold was in effect determined for this condition. When this condition was examined in all three test packages, the effect of increasing pulse duration on the threshold was determined. Increasing the duration of each pulse from 10 to 17 msec reduced the threshold from 30 to 20 msec; however a further increase in pulse duration from 17 to 28 msec did not effect the threshold value. These results were different from those reported by Dunlap (1915), Mahneke (1958), and Purcell and Stewart (1971), all who used the two-flash threshold as a measure of resolution.

Theodor (1972) studied the detectability of a "gap" (i.e., a fixed interpulse interval at durations of 10, 14, and 16 msec) for various first pulse - second pulse combinations (450 - 50 msec, 350 - 150 msec, 250 - 250 msec, 150 - 350 msec, and 50 - 450 msec) with a constant luminance of 3 fL. (10.27 cd/m^2). The total light time was 500 msec, a duration much longer than previous studies. Using the signal detection rating scale procedure, and unbiased estimate of observer's ability to detect the gap within each of the pulse duration combinations. In half of the trials a 500 msec single pulse was presented and in the other half two pulses were presented. The observer's task was to state whether or not he detected a gap (interpulse interval) and to rate the confidence of his decision. It was found that when the duration of the first pulse had a much longer duration than that of the second pulse detectability was better; a finding which seemed to be consistent with the findings of dunlap (1915), Mähneke (1958), and Purcell and Stewart (1971) who found a reduction in the two-flash threshold when the duration of the first pulse was increased. Theodor also found that when the second pulse duration was very "brief" (150 or 50 msec) detectability was reduced; a finding which seemed consistent with the findings of Purcell and Stewart (1971) who found an increase in the two-flash threshold when the duration of the second pulse was decreased. None of these findings were in agreement with the comparison stimulus two-pulse findings of Kietzman and Smith (1965) that were previously mentioned.

Both two-flash and two-pulse measures of temporal resolution have been studied under a variety of experimental conditions, e.g., different pulse luminances and pulse durations (Boynton, 1972) a fact that has made comparision between studies extremely difficult. This difficulty is compounded by the fact that no one theory, be it the "quantity of light" (Mahneke, 1958), visual persistence (Purcell and Stewart, 1971), or critical duration (Davy, 1952; Kietzman, 1967; Lewis, 1967) can adequately explain and predict the complexity of results. It was apparent that very little was known about the effecgs of changing pulse luminance or pulse duration on two-pulse measures of temporal resolution, in the face of a large body of scientific investigation. Furthermore, it could not be determined whether or not the various experimental manipulations produced different effects on the two-flash and comparison stimulus two-pulse measures. Hence, the results of these studies could not be used as a means to determine whether or not the distinction between the classical psychophysical and forced-choicd two-pulse temporal resolution measures was necessary or justified.

Signal detection theory. Prior to the present study there was no known two-pulse temporal resolution study that applied the signal detection rating scale procedure. In addition, there was no single study that had attempted to compare signal detection measures for sensitivity and response bias from procedures which maintain the characteristics of the previous classical psychophysical and forced-choice studies of two-pulse temporal resolution. There are however, signal detection yes-no procedures that have many characteristics in common with both the classical psychophysical procedures (Tong, 1972; Gruzelier &Venables, 1974) and the forced-choice procedures (Mertens & Lewis, 1972). Signal detection measures of sensitivity have the advantages of being independent of a measurable decision criterion and psychophysical method used in making the measurements (Swets, 1973). In addition to these advantages the signal detection rating scale procedure has over other signal detectfion procedures. In the yes-no task the observer is usually allowed only two responses, "yes" a signal was presented, and "no" a signal was not presented. In a rating scale procedure the observer gives more information about his decision, i.e., the certainty of which he believes that a signal was presented; there is nothing within decision theory that requires an observer to make only a binary decision (Green & Swets, 1966,p. 40). Furthermore, a trained observer can perform as well on a rating task as on a yes-no task (Egan, Schulman & Greenberg, 1959).[1]

It is assumed in the rating scale experiment that the observer is able to maintain several decision criteria simultaneously; each point on the rating scale is considered a different criterion. A strict criterion corresponds to points on the rating scale which indicate a high certainty of response and an expected low false alarm rate. A lax criterion corresponds to points on the rating scale which indicate uncertainty and an expected high false alarm rate. Although in theory any number of categories on a rating scale may be used, in practice usually only 4 to 10 categories are used (McNicol, 1972, pp.101-104). The range of the scale is from high certainty that a signal was presented (corresponding to the strictest criterion) to high certainty that noise was presented

---

[1]Several other studies have shown that comparable sensitivity measures are obtained from rating scale and yes-no procedures (e.g., Swets,1959; Emmerich, 1968). However, this relation may be specific to visual and auditory stimuli. For example, Clark and Mehl (1971) and Clark and Dillon (1973) have reported differences in sensitivity measures between the yes-no and rating scale procedures when thermal stimuli were used.

(corresponding to the laxest criterion). It is apparent that when an observer reports high certainty that noise was presented it is the same as reporting high uncertainty that a signal was presented. The raw data is converted into a set of hit and false alarm probabilities.[2] From these probabilities the relative operating characteristic (ROC) curve may be constructed. The probability of a false alarm is usually plotted as the abscissa and the probability of a hit is plotted as the ordinate. The probability of a hit corresponds to the "power" of a statistical test (1-the probability of a Type I error) and the probability of a false alarm corresponds to the probability of a Type II error. With n rating categories, n-1 points on the ROC curve may be constructed when the observer uses all categories. In the yes-no procedure for an ROC curve to be constructed from n points n separate experimental conditions must be used.

---

[2]The discussion of signal detection rating scale procedures is presented here to facilitate the reader's understanding of the problem. The terminology and symbols used here represent their traditional usage. Another system which gives justice to the actual responses of the observer and not to inferred responses will be introduced at a later point in this presentation and an explanation of it is presented in the Appendix.

Hence, another advantage of the rating scale procedure is that it is economical as well as efficient.

In most psychophysical experiments, the experimenter has no apriori knowledge about the signal and noise distributions. When plotted on double probability scales,i.e., in deviation units for a standard normal distribution, the form of the ROC curve may provide information regarding the shape of the signal and noise distributions. If the underlying signal and noise distributions are Gaussian then the ROC curve will be a straight line. If the variances of the two distributions are equal then the slope of the line should be equal to unity. When both the signal and noise distributions are normal with equal variance, the distance between the means of these distributions, denoted d' (Green & Swets, 1966, pp. 50-68). Intuitively, this means that the further the distance from the mean of the signal distribution to the mean of the noise distribution, the easier it is for the observer to discriminate signal from noise. Also, when the signal distributions are normally distributed with equal variance, the decision criterion measure, is the value of the likelihood ratio at which the criterion has been set to yield a particular point on the ROC curve. Mathematically, empirical β is the ratio of the ordinate of the signal distribution to the ordinate of the noise distribution at the criterion setting.

When the variance of the signal distribution is greater than that of the noise distribution, the slope of the ROC curve (plotted in standard deviation units) will be less than unity; when the variance of the signal distribution is less than that of the noise, the slope will be greater than unity. When the slope of the ROC curve is not equal to unity, the difference between z(S/n) and z(S/s) changes as different points along the ROC curve are selected. Hence, d' cannot be used as a measure of sensitivity for the unequal variance case.

Several alternate measures of sensitivity have been used when the signal and noise distributions were of unequal variance. The measure D (Δm,s), suggested by Green and Swets (1966), pp. 96, 407) is one such measure; where Δm, is the difference between the means of the signal and noise distributions measured in standard deviation units of the noise distribution. It is equal to z(S/n) at the point where z(S/s) = 0. In this case, s is the slope of the ROC curve. Egan and Clarke (1966) were the first to suggest another widely used measure, d'e defined as twice the value of z(S/s) or z(S/n), ignoring signs, at the point where the ROC curve intersects the negative diagonal. The d'e measure has the advantage of giving equal weight to the signal and noise standard deviations and therefore is an appropriate measure when it is suspected that both the signal and noise variance has changed over a series of treatments. Another widely used measure is P(A), the area under the ROC curve (Green & Swets, 1966, pp. 45-50). The P(A) is nonparametric measure of sensitivity and no assumptions need to be made about the form of the underlying signal and noise distributions. (Both D (m,s) and d'e require that the underlying signal and noise distributions be normally distributed). The area under the ROC curve is an index of the degree of overlap between the signal and noise distributions. A high degree of overlap indicates low sensitivity; little overlap indicates high sensitivity. Several other measures of sensitivity have been suggested (see Simpson & Fritter, 1973). It is apparent that the selection of a sensitivity measure is rather arbitrary; what is most important is that the convention used be made explicit and adhered to in the experiment in order that comparisons of sensitivity measures under different conditions to be valid (McNicol, 1972, p. 90).

When the variances of the signal and noise distributions are not equal to unity, the underlying signal and noise distributions may intersent at more than one point where is equal to 1 (Green & Swets, 1966, pp. 62-65). McNicol, (1972, pp. 93-96) has suggested a method for calculating for the unequal variance case which makes the usual assumption that the noise distribution has a mean of zero and a standard deviation of 1, and also takes into account the slope of the ROC curve.

The two-flash threshold obtained from classical psychophysical studies of visual temporal resolution is a noncriterion-free measure of discrimination. On the other hand, the comparison stimulus two-pulse threshold measure derived from the forced-choice studies is a criterion-free measure of discrimination. It was thought that a signal detection analysis of these measures, that would provide independent measures of sensitivity and criterion, would help to clarify whether or not the differences in "threshold" magnitudes represent sensitivity differences or differences due to criterion. The most salient characteristics of the classical psychophysical procedures that have been used to measure two-pulse visual temporal resolution that were retained in the "one-two" procedure of the present study included a single observation interval, no comparison stinulus, "one-two" responses, and Class B observations. Unlike the classical psychophysical studies, feedback was given to the observers and the sensitivity measures obtained are criterion-free. The most prominent features of the forced-choice studies which were retained in the "same-different" procedure of the present study included the presentation of at least one comparison and test stimulus on a single trial, Class A observations, and the sensitivity measures obtained were criterion-free. Differences between the forced-choise procedure and the "same-different" rating scale procedure made it necessary to determine if these procedures would produce similar results in a control experiment. For the main experiment it was thought that the extra information given by the comparison stimulus in the "same-different" condition would improve sensitivity and that sensitivity differences would be found between the "same-different" and "one-two" conditions.

II. Main Experiment: Method

Observers

There were two observers. Observers MG was a female undergraduate, age 20, who had normal visual acuity with correction and was paid for her services. She was informed prior to the study of the tasks she would have to perform but was told nothing about the specific stimulus conditions, design, and purpose of the study. Observer JL, age 26, the author, had normal visual acuity without correction.

Apparatus

A single channel binocular free vieing optical system was used to present the light pulses and the fixation lines. The light source was a glow modulator gas-discharge tube (Sylvania R1131C) operated by a constant current and activated by timing and gating circuits. An argon ultraviolet lamp (General Electric AR-4) irradiated the glow modulator tube to provide short and stable ionization times (20-40 microsec). Pulse durations and the timing of other events within each trial within each trial were controlled by a transistorized digital time generator (Analog and Digital Instruments Company TG700). After passing through neutral density filters (Kodak Wratten) the light output of the glow modulator impinged upon a homogeneous white lucite disc. The size of the tatrget viewed by the observer was 3 mm in diameter and subtended a visual angle of about 20 min. At a distance 3.7 mm above, below, and to the left and right of the target were four dim red fixation lines (2.93 x 1.70 mm) illuminated b L.E.D. bulbs and powered by a separate power supply. The fixation lines subtended a visual angle of 1 deg 47 min to produce foveal fixation.

Before and after each session of data collection the intensity of the glow modulator was monitored on an oscilloscope (Tektronix 532) display of the output of a phototube (RCA 929) powered by another power supply. This was done to control fo the stability of the light output from session to session and to be certain that the glow modulator did not heat up sufficiently during the session to change its output. Throughout the data collection period the reading at the beginning of a session was identical to that at the end of the session. The variability of the glow modulator output on a session to session basis was less than 7%.

Due to the unavailability of a photometer, an absolute lúminance calibration was done only once at the end of the study with a Spectra Prichard Photometer.

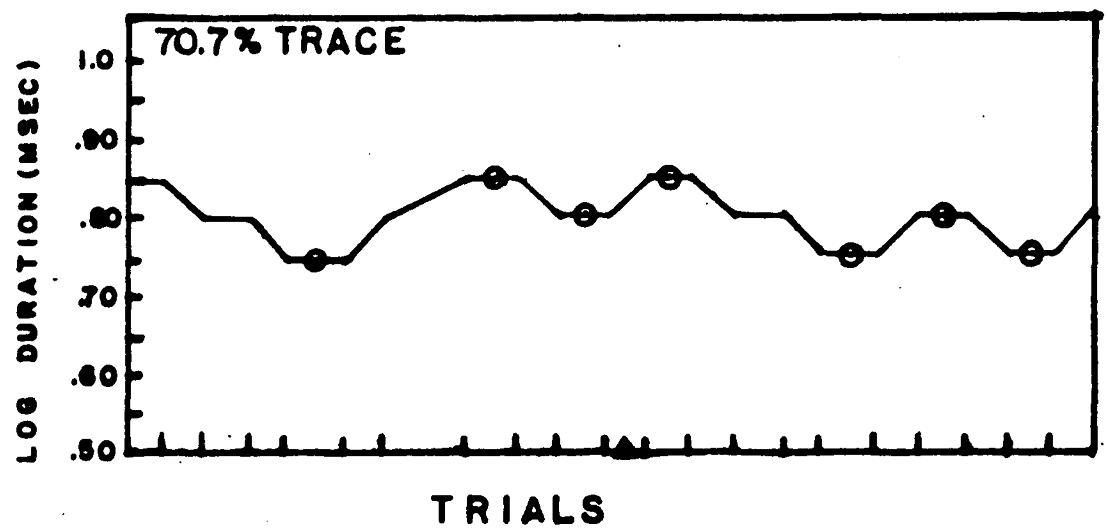
<u>Procedure</u>

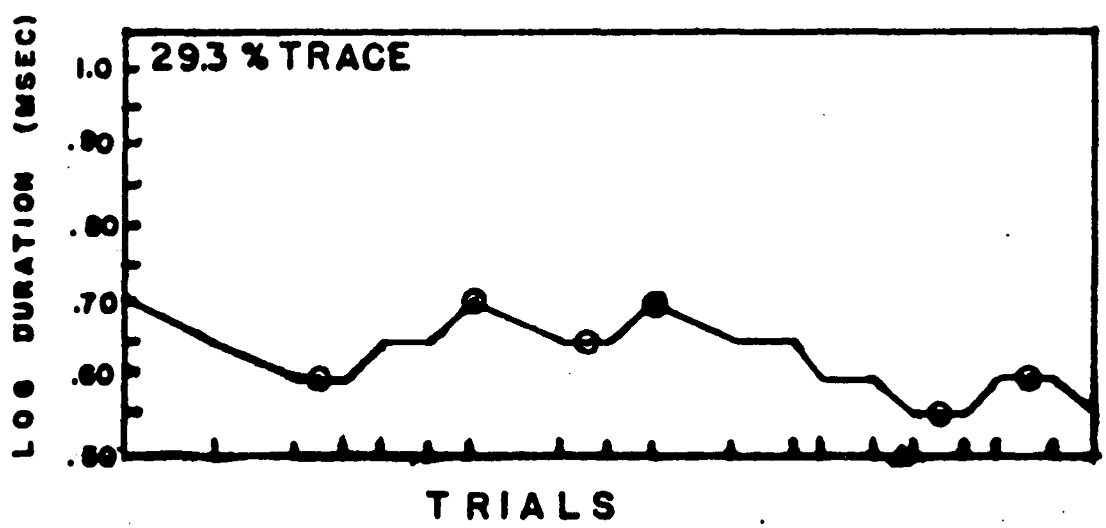
<u>Preliminary testing</u>. In order to determine an appropriate luminance level for a single light pulse an Up-Down-Transformed-Response (UDTR) procedure was employed to estimate the 29.3% and 70.7% points on the psychometric function (Witherill & Levitt, 1965; Levitt & Treisman, 1969).

In view of the fact that the experimental apparatus did not provide a convenient means of changing luminance advantage was taken of Bloch's law, which states that for threshold detection the duration and the intensity of a light pulse are reciprocally related up to a critical duration (Graham, 1965). Therefore, at a constant luminance the duration of a single pulse of light was varied according to the "stepping rule" in approximately .05 log unit steps rangong from 3.6 to 10 msec. In order to estimate the 70.7% point the duration was reduced by a single step only after two consecutive "yes" (i.e., seen) responses, and increased by a step after each "no" response. The stragedy uased to estimate the 29.3% point was exactly the reverse. The control charts provided by Levitt and Treisman (1969) were used to facilitate presentation and recording of the data.

Observer MG was seated, positioned using a chin rest, and dark-adapted for 8 min. Each trial began with a warning click which signaled the observer to fixate on the hypothetical intersect of the four fixation lines. On 90% of the trials the warning click was followed by a light pulse after a 1.5 sec fixed foreperiod (signal trials). On the other 10% of the trials, no light pulse was presented after the foreperied (noise trials). The observer's task was to report "Yes" if she saw the light and "No" if she did not see the light. Two practice sessions of 112 each were given to familiarize the observer with the task. Then three testing sessions of 112 trials each were run to estimate the desired psychometric points. The average of the "peaks" and the "troughs" for both the 70.7% trace and the 29.3% trace were calculated; their average (grand mean) provided an estimate of the 50% (threshold) point. Figure 1 shows sample data obtained using this procedure. Appropriate calculations were made in order to find a filter combination which when used with a 2 msec pulse would provide an energy

Figure 1. Sample data obtained from the UDTR procedure. Circles indicate points which contributed to the determination of the 50% threshold point. Triangles indicate trials where no light pulse was presented.

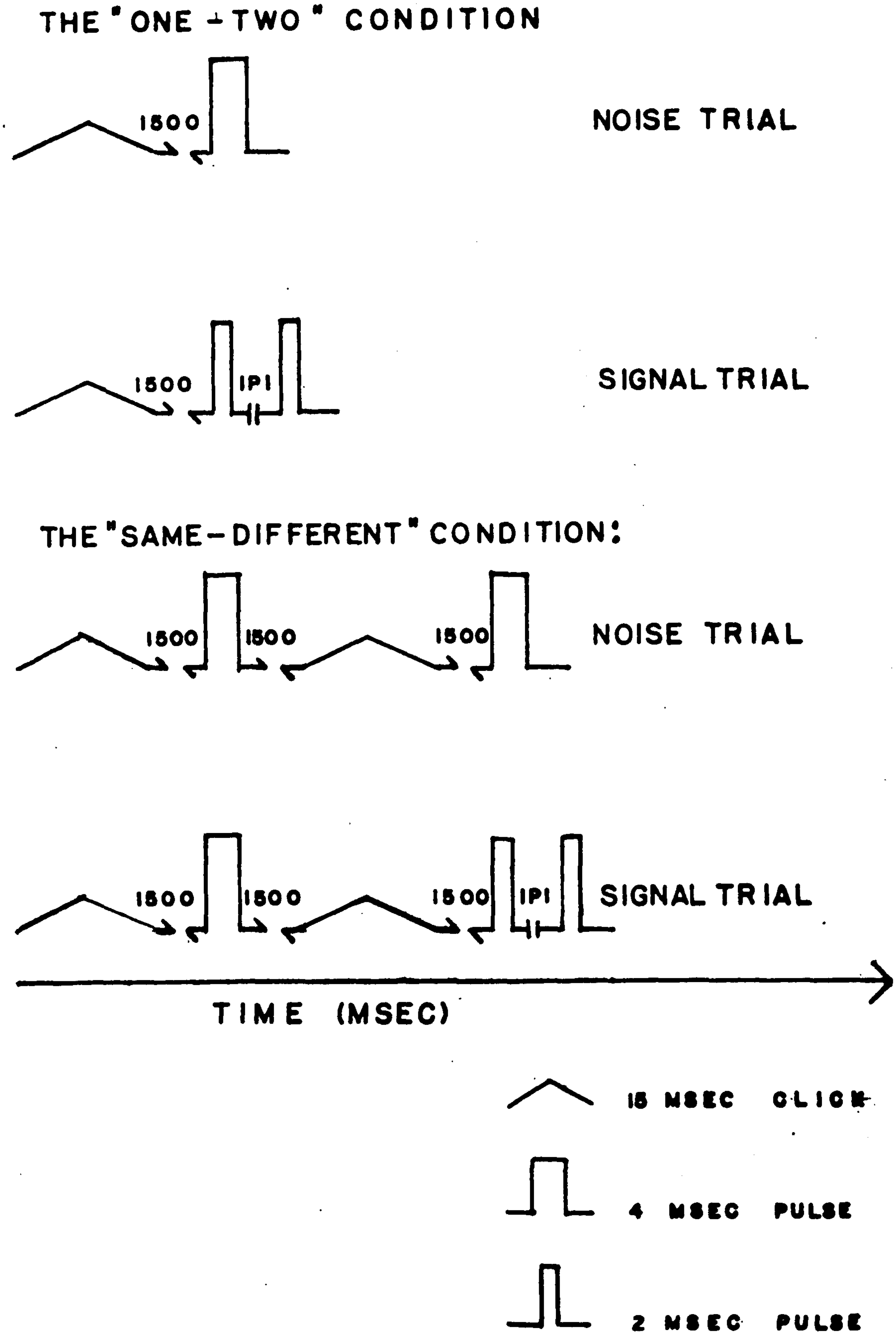

70.7 % TRACE
LOG DURATION (MSEC)
1.0
.90
.80
.70
.60
.50
TRIALS


29.3 % TRACE
LOG DURATION (MSEC)
1.0
.90
.80
.70
.60
.50
TRIALS

level which would be .6 log units above the 50% (threshold) point. The absolute luminance of the target with this filter combination was later found to be .43 (1.37 cd/m^2) and was the luminance used throughout the rest of the study for both observers.

Main experiment. Two signal detection procedures provided the experimental conditions for the present study. One procedure was designed to measure two-pulse discrimination as "one-two" flash responses using a rating scale procedure (the "one-two" condition). The other procedure was designed to measure two-pulse discrimination as "same-different" responses measured with reference to a single pulse comparison stimulus using a rating scale procedure (the "same-different" condition). The sequence of events for both the "one-two" and "same-different" conditions is shown in Figure 2. A noise trial for the "one-two" condition consisted of a 15 msec warning click, followed by a fixed 1.5 sec foreperiod and then a 4 msec light pulse (the noise stimulus); a signal trial consisted of a 15 msec warning click, followed by a 1.5 sec fixed foreperiod and then two equal-energy 2 msec light pulses separated by an interpulse interval (the signal stimulus). A noise trial for the "same-different" condition consisted of a 15 msec warning click followed by a 1.5 sec foreperiod and then a 4 msec light pulse (the comparison stimulus); after 1.5 sec this sequence was repeated with the second 4 msec light pulse being the noise stimulus. A signal trial for the "same-different" condition consisted of a 15 msec warning click followed by a 1.5 sec foreperiod and

Figure 2. Sequence of events for signal and noise trials for both conditions. IPI indicates an interpulse interval duration of 20, 40, 60, or 80 msec. See text for details.


THE "ONE - TWO" CONDITION
1500
NOISE TRIAL
1500
IPI
SIGNAL TRIAL
THE "SAME-DIFFERENT" CONDITION:
1500
1500
1500
NOISE TRIAL
1500
1500
1500
IPI
SIGNAL TRIAL
TIME (MSEC)
15 MSEC CLICK
4 MSEC PULSE
2 MSEC PULSE

then a 4 msec light pulse (the comparison stimulus); after 1.5 sec the 15 msec click was resounded followed by a 1.5 sec foreperiod and then rwo equal-energy 2 msec light pulses separated by an interpulse interval (the signal stimulus). Four interpulse interval durations were used for both conditions: 20, 40, 60, and 80 msec.

During each testing session the observer was seated and positioned on the chin rest and was given 8 min to dark-adapt. For the "one-two" condition the observer was told that on each trial a warning click would be presented, which meant that the observer should fixate between the four fixation lines (which were previously adjusted so that they were just visible), and that following the click a stimulus would be presented. The observer was told that sometimes the stimulus would consist of two flashes of light and sometimes only one flash. The observer was to report whether he saw "one" or "two" light flashes. Also, he was to report the certainty of his response using the following code: "1" signified "very certain", "2" signified "moderately certain", and "3" signified "very uncertain". Therefore, after each trial the observer reported whether he saw one or two flashes, and then reported the code number indicating the certainty of his response. For the "same-different" condition the observer was told whether the stimulus following the second click was "same" or "different" from the stimulus following the first click; and also was told to report the certainty of his response using the same system as in the "one-two" condition. The observers were permitted to use any and all possible cues to determine a "difference".

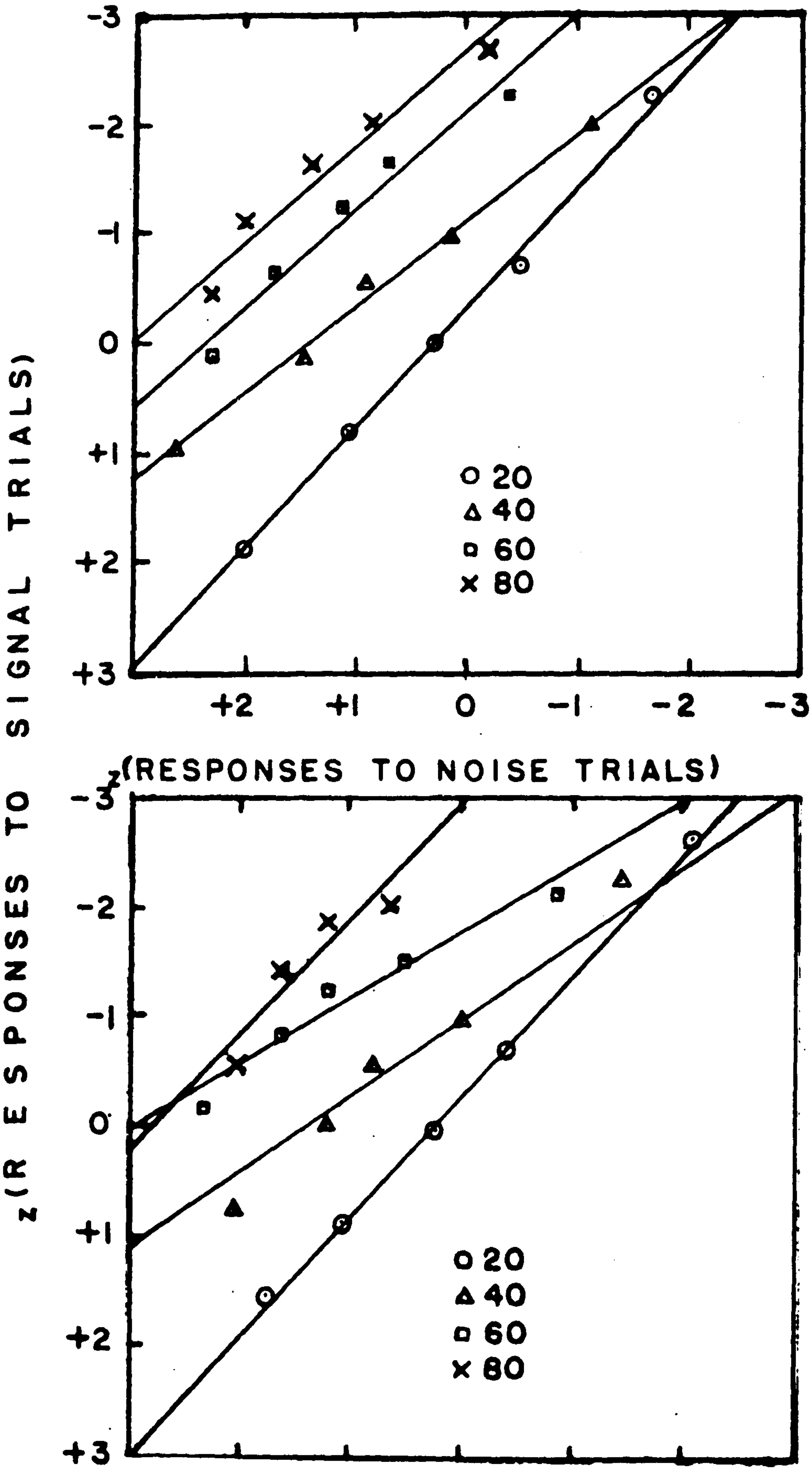
After each trial in which the observer made a correct decision, i.e., correctly reported "one" or "two" or "same" or "different", regardless of of the certainty of his response, the observer was told "Yes" by the experimenter; therefore, feedback was given.

The "one-two" and "same-different" conditions were counterbalanced throughout a session (ABBA). Each component of the counterbalancing (e.g., A) contained four blocks of 16 trials. Within each block there were 50% signal and 50% noise trials, all at the same interpulse interval. The interpulse interval used throughout a block was quasi-randomly selected. There were 256 trials per session; 128 per condition. Each observer participated in 16 sessions.

In addition, both observers were given sufficient practice; the practice sessions were conducted in an identical manner to the experimental proper. Obwerver MG was given three sessions (768 trials) of practice. Observer JL, a more experienced observer, was given two sessions (512 trials) of practice.

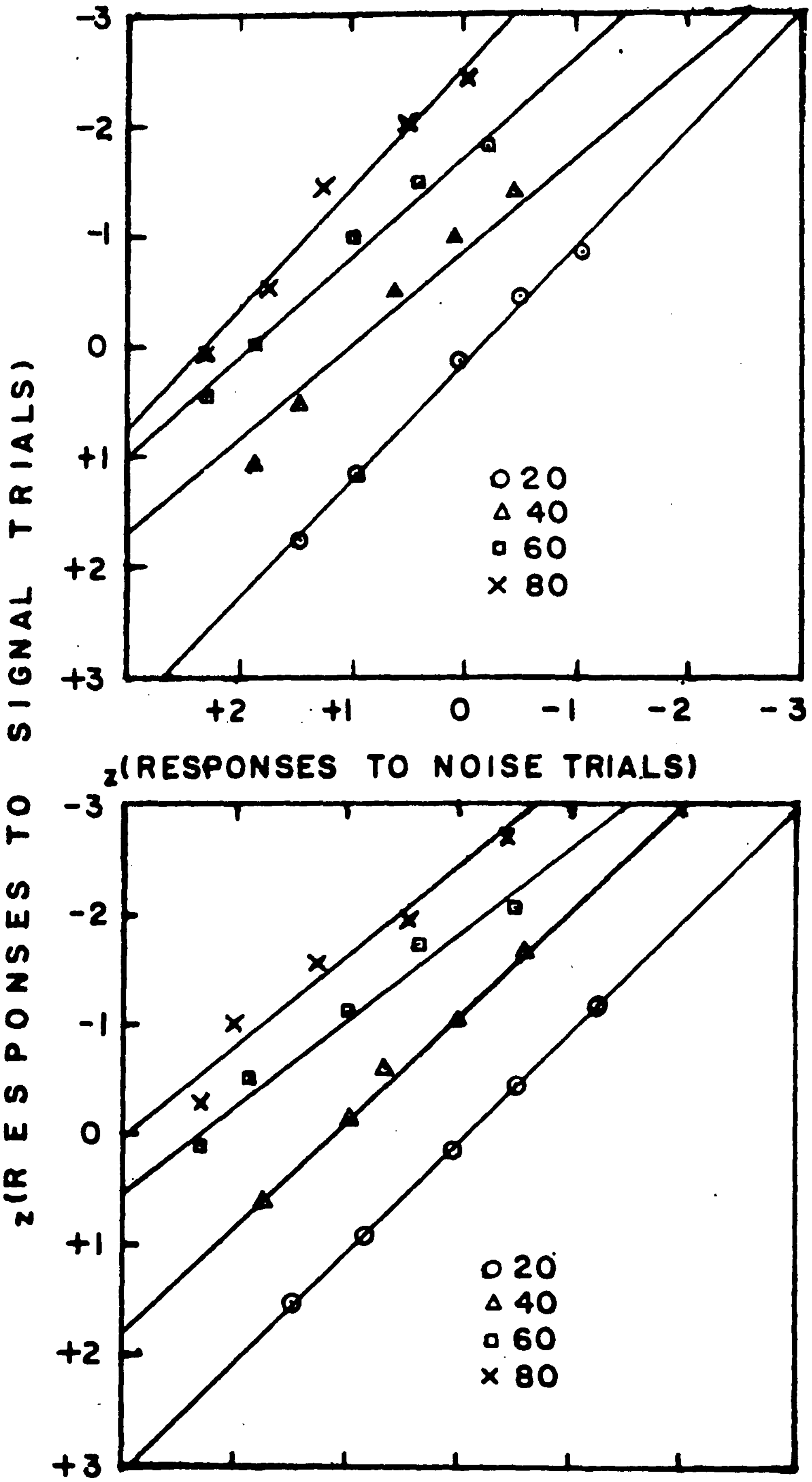
Treatment of data. For each observer two families (one for each condition) of ROC curves were constructed and are shown in Figures 3 and 4. Each curve was constructed from the commulative probabilities of responses to signal and noise trials z(S/s) or z(N/s) and z(s/n) or

Figure 3. Family of ROC curves for the "one-two" condition (top) and "same-different" condition (bottom) for observer MG. Circles, triangles, squares, and crosses indicate criterion points for interpulse intervals of 20, 40, 60, and 80 msec respectively.



-3
-2
-1
0
+1
+2
+3
+2
+1
0
-1
-2
-3
○ 20
△ 40
□ 60
× 80
z(RESPONSES TO NOISE TRIALS)
z(RESPONSES TO SIGNAL TRIALS)
○ 20
△ 40
□ 60
× 80

Figure 4. Family of ROC curves for the "one-two condition (top) and "same-different" condition (bottom) for observer JL. Circles, triangles, squares, and crosses indicate criterion points for interpulse intervals of 20, 40, 60, and 80 msec respectively.


-3
-2
-1
0
+1
+2
+3
+2
+1
0
-1
-2
-3
z(RESPONSES TO NOISE TRIALS)
z(RESPONSES TO SIGNAL TRIALS)
○ 20
△ 40
□ 60
× 80
-3
-2
-1
0
+1
+2
+3
○ 20
△ 40
□ 60
× 80

z(N/n), respectively)[3] for the various criterion[4] settings at one interpulse interval and then applying the methos of mutual regression (Grice, 1966) to determine the line of best fit. The equations of these lines are presented in Table 1 in the form y = mx + b, where y = z(S/s), or z(N/s), x = z(S/n), or z(N/n), m = the slope, and b = z(S/s), or z(N/s), when z(S/n) or z(N/n) = 0.[5] The mutual regression analysis guarantees that the line of best fit will pass through the centroid (Gibson, Note 2).

All of the ROC curves were characterized by straight lines of nonunity slope. These characteristics were indications of underlying Gaussian signal and noise distributions of unequal variance. Therefore, appropriate

---

[3]The terminology used here and the way in which symbols are used to label the ordinate and abscissa of the rating ROC curve are different from the traditional terminolgy and use of symbols. As suggested by Kietzman (Note 1), the system used here gives justice to the actual responses of the observer and not inferred responses (see Appendix for details).

[4]Observer MG had failed to use the "same, very certain" rating category at an interpulse interval of 80 msec. This category was removed by combining it with the "same, certain" category. This method of combining categories has the advantage over other methods which retain categories by adding or subtracting constants in that there is less error in distorting the ROC curve when it is plotted (McNicol, 1972, pp. 105-108).

[5]Or y = mx + b, where y = z(S/s), x = z(S/n), m = the slope, and b = z(S/s) when z(S/n) = 0, in the traditional system.

Table 1

Mutual Regression ROC Lines for "One-Two" and "Same-Different" Conditions at each Interpulse Interval

Observer	IPI	"One-Two"	"Same-Different"
MG	20	y = 1.106 x + .341	y = 1.110 x - .320
	40	y = .807 x - 1.154	y = .727 x - 1.027
	60	y = .892 x - 2.124	y = .604 x - 1.802
	80	y = .855 x - 2.671	y = 1.083 x - 2.994
JL	20	y = 1.052 x + .141	y = 1.004 x + .059
	40	y = .836 x - .861	y = .962 x - 1.102
	60	y = .919 x - 1.760	y = .770 x 1.819
	80	y = 1.098 x - 2.546	y = .808 x 2.395

sensitivity measures, d'e, D(Δm,s), and P(A) were calculated. The d'e measure was calculated as twice the z(S/s) or z(N/s) value of the poit at which the ROC curve crossed the negative diagonal.[6] In calculating D(Δm,s), Δm was taken as the value of z(S/n) or z(N/n) at the point where z(S/s) or z(S/s) or z(N/s) = 0;[7] s, the slope was taken as the slope of the fitted ROC curve. The nonparametric measure, P(A), was calculated using the trapezoidal method (McNicol, 1972, p. 115).

To test for differences in sensitivity between the two conditions, a method suggested by Gibson (Note 2) was used. It was apparent that in the rating scale procedure that each criterion point could give an estimate of sensitivity by taking the difference of the z(S/s) or z(N/s) and z(S/n) or z(N/n) probability values[8] at each criterion point. The mean of these values (d'M) could then be taken as an overall measure of sensitivity. In a sense, each measure of sensitivity at a given criterion point was comparable to a d' measure when only one point on the ROC curve was known. For both conditions d'M was calculated at each interpulse

---

[6]Or as twice the z(S/s) value of the point at which the ROC curve crossed the negative in the traditional system (Egan & Clarke, 1966).

[7]Or Δm was taken as the value of z(S/n) at the point where z(S/s) = 0, in the traditional system.

[8]Or the difference of the z(S/s) and z(S/n) probability values in the traditional system.

interval. An appropriate method for testing the difference between d'M for the "one-two" and "same-different" conditions at each interpulse interval was the t-test for repeated measures.[9]

For each interpulse interval at every criterion setting, a measure of response bias, log β', was calculated according to the method suggested by McNicol 1972, pp. 93-96, 123) for the unequal variance case. Here, the ordinate for z(S/s) or z(N/s) was divided by the ordinate for z(S/n) or z(N/n),[10] multiplied by the slope of the ROC curve, and then the log of the result was taken. Log β' was reported instead of ' for two reasons. First, β' has a lower limit of zero and cannot assume negative values. In view of this fact, β' values may be skewed and tend to cluster near zero. Transformation of β' values to log β' values attempts to remove the skew and thus enables log β' values to be used in statistical tests requiring a roughly normal distribution of scores. Second, log β' values facilitate interpretation of the degree of bias. When there is a bias to signal responses log β' will be negative. When there is a bias to noise responses log β' will be positive. When bias is absent log β' will be zero.

---

[9]For observer MG the "one-very uncertain" at an interpulse interval of 80 msec was not included in statistical tests to provide an equal number of comparisons. It was thought that sacrificing one degree of freedom would outweigh the consequences of further distorting the ROC analysis by adding and subtracting constants.

[10]Or the ordinate for z(S/s) was divided by the ordinate for z(S/n), in the traditional system.

From each of the criterion settings the mean log β' (log β'M) was calculated at each interpulse interval between both conditions a series of t-tests were performed (Gibson, Note 2).

III. Main Experiment: Results

The results of the present investigation indicated that there were no overall differences in rating scale signal detection measures of sensitivity and criterion between the "one-two" and "same-different" conditions for both observers.

Appropriate sensitivity measures for underlying Gaussian signal and noise distributions of unequal variance, d'e, $D(\Delta m,s)$ and P(A) are shown in Table 2. For both the "one-two" and "same-different" conditions the value of each sensitivity measure increased as the inter-pulse interval was increased. Moreover, there appeared to be no difference between the value of each sensitivity measure between conditions at each interpulse interval. Table 3 shows the value of d'M for each interpulse interval and the value of t. For the two observers the null hypothesis of no difference between the procedures was not rejected at any of the interpulse interval values tested ($p > .01$ for all values of t).

Table 4 shows the value of log βM at each interpulse interval for both conditions and the obtained values of t. The null hypothesis was not rejected (at the $p > .01$ level) for all values of t.

Although there were no significant differences in criterion between the "one-two" and "same-different" conditions at each interpulse interval, the criterion point representing the minimal criterion value at which the observer

Table 2
Signal Detection Measures of Sensitivity
at each Interpulse Interval

		d'e		P(A)		D(Δm,s)	
Observer	IPI	O-T	S-D	O-T	S-D	O-T	S-D
MG	20	.33	.34	.55	.56	(.31, 1.10)	(.29, 1.11)
	40	1.30	1.16	.80	.81	(1.43, .81)	(1.41, .73)
	60	2.26	2.27	.94	.94	(2.39, .89)	(2.98, .60)
	80	2.96	2.86	.98	.97	(3.12, .86)	(2.76,1.08)
JL	20	-.13	-.06	.47	.49	(-.13,1.05)	(.06, 1.00)
	40	.93	1.16	.75	.79	(1.03, .84)	(1.14,1.10)
	60	1.86	2.14	.90	.92	(1.94, .92)	(2.36, .77)
	80	2.46	2.73	.95	.97	(2.32, 1.10	(2.96, .81)

Table 3

Comparision of d'M for the "one-Two" and "Same-Different"

Conditions and Obtained Values of t

		d'M				
Observer	IPI	O-T	S-D	t^a	df	p
MG	20	.312	.306	.260	4	.9
	40	1.313	1.174	2.528	4	.1
	60	2.245	2.189	2.776	4	.05
	80	2.956	2.877	1.475	3	.3
JL	20	-.116	-.060	2.122	4	.2
	40	.980	1.124	2.500	4	.1
	60	1.847	2.051	1.096	4	.4
	80	2.430	2.616	.174	4	.9

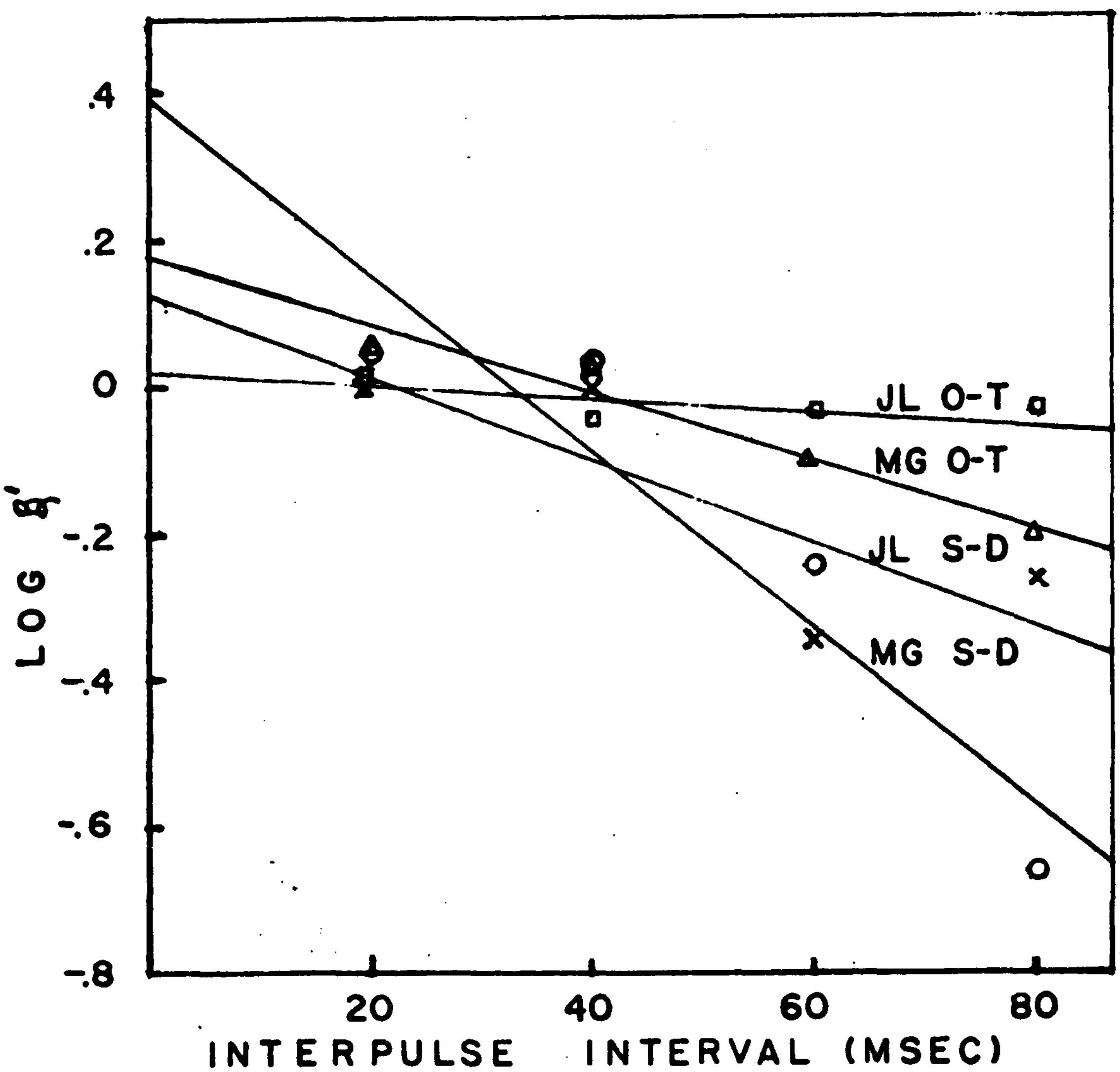
Table 4

Comparision of logδ'M for the "One-Two" and "Same-Different" Conditions and Obtained Values of t

		Logδ'M				
Observer	IPI	O-T	S-D	t^{a}	df	p
MG	20	-.032	-.003	.723	4	.5
	40	-.069	.104	1.061	4	.4
	60	-.033	-.152	1.330	4	.3
	80	.177	-.284	1.035	3	.3
JL	20	-.016	-.020	.066	4	.9
	40	-.006	.078	1.917	4	.2
	60	.122	-.042	3.405	4	.05
	80	-.022	-.150	1.310	4	.3

reported a signal for a fixed interpulse interval warranted further analysis. This point was the first point of transition from the report of no signal to the report of signal. For the "one-two" condition, this was the point where the observer reported high uncertainty that two flashes were seen ("Two, 3"). For the "same-different" condition, this was the point where the observer reported high uncertainty that the stimuliwere different (Different,3"). Figure 5 shows logβ' at these criterion points as a function of interpulse interval; straight lines were fit to the criterion points by the method of least squares. In general, as the interpulse interval was increased log β' decreased. However, the rate of change in criterion with respect to interpulse interval duration was dramatically different between the two-flash and two-pulse conditions. For the two-flash condition there was very little change in criterion as the interpulse interval was increased; in fact, for JL the slope of the line was nearly zero. For the two-pulse condition there was a steady decrease in log β' as the interpulse interval was increased. The greatest differences in criterion at this criterion setting between the two conditions occurred at the longest interpulse intervals. This result suggests that as the interpulse interval increased the observers maintained nearly a constant decision criterion for "twoness" in the "one-two" condition while they adopted an increasingly more lax criterion in the "same-different" condition.

Figure 5. Logβ' as a function of interpulse interval. Triangles and circles represent the data points for the "one-two" and "same-different" conditions, respectively, for MG. Squares and crosses represent the same respective data points for JL.



.4
.2
0
-.2
-.4
-.6
-.8
LOG β'
20
40
60
80
INTERPULSE INTERVAL (MSEC)
JL O-T
MG O-T
JL S-D
MG S-D

IV. Main Experiment: Discussion

The findings of the present investigation supported the implied contention of several investigators (Lewis, 1967; Lewis, 1968; Lewis & Mertens, 1971; Mertens and Lewis, 1972; Purcell & Stewart, 1971; Theodore, 1972) that the measures of two-resolution obtained from classical psychophysical and forced-choice pulse visual temporal procedures are measures of the same visual phenomenon. No differences in signal detection measures of sensitivity for the two response conditions were found. The trend of increasing sensitivity with increasing interpulse interval duration was consistent with all other studies that have reported increasing discrimination with increasing interpulse interval duration.

In this study differences in procedure were the basis for distinguishing the the "one-two" from the "same-different" condition. The most prominent feature of all classical psychophysical studies of two-pulse resolution is that the observer is presented a single stimulus and asked to report whether or not two flashes are seen. Another characteristic of these studies is that there is no way in which the experimenter has an indication the observer is indeed following the instructions, or what criterion the observer is adopting. In many studies (Mahneke, 1958; Stewart & Purcell, 1971; Kietzman, 1967) two-pulse stimuli of varying interpulse intervals were presented within a single block of trials and the experimenter had no objective indication of the correctness of the observer's response. The present investigation made no attempt to remove this characteristic for the "one-two" condition (it was considered an important characteristic of the two-pulse resolution literature.) Furthermore, it could not be determined what effect an experimental manipulation of this characteristic would have on the observer's criterion. In short, the present study attempted to study two-pulse resolution as it was studied in the past without changing or manipulating its most salient characteristics but to study it within the realm of signal detection theory.

Incorporated within the present investigation were features which in a sense served to indicate whether or not the observer was following the "one-two" instructions. For example, in view of the fact that criterion was a measure sought, it was necessary to present only one value of the interpulse interval in a block of trials (Emmerich, 1968). Since in 50% of the trials "noise" stimuli (single pulses) were presented (instead of all double-pulses of varying interpulse interval durations), the experimenter had an objective indication of the observer's performance. Although this indication was helpful to the experimenter, and was advantageous relative to other studies in which such an indication was nonexistent, it still did not eliminate the fact that the observers could be scored as correct or incorrect independent of the "oneness" or "twoness" of the stimulus package. A more sensitive and perhaps a more important indication that the observer was following the "one-two" instructions was the change in the criterion values of the "Two, 3" ("Two flashes, very uncertain") setting as the interpulse interval was increased. Since the "Two, 3" point represented the minimal criterion point for the report of two flashes it was expected that the observer's criterion for a report of two flashes would not change appreciably as the interpulse interval was increased (Kietzman, Note 1). Considering that there was a response indicator for the correctness of the observers' responses and that the minimal criterion value for a report of two flashes did not change appreciably as the interpulse interval duration was increased, it may be stated with greater certainty that the observers in the present study did follow the "one-two" instructions.

The most prominent procedural feature of all forced-choice two-pulse resolution studies, which was retained in the "same-different" condition was that at least one comparision and one test stimulus was presented on a single trial and the observer was asked if a difference was observed. In most forced-choice two-pulse resolution studies e.g. (Kietzman, 1967; Lewis, 1968; Nilsson, 1969) the observer was permitted to use any and all cues to determine a "difference"; the observer need not have seen two flashes.

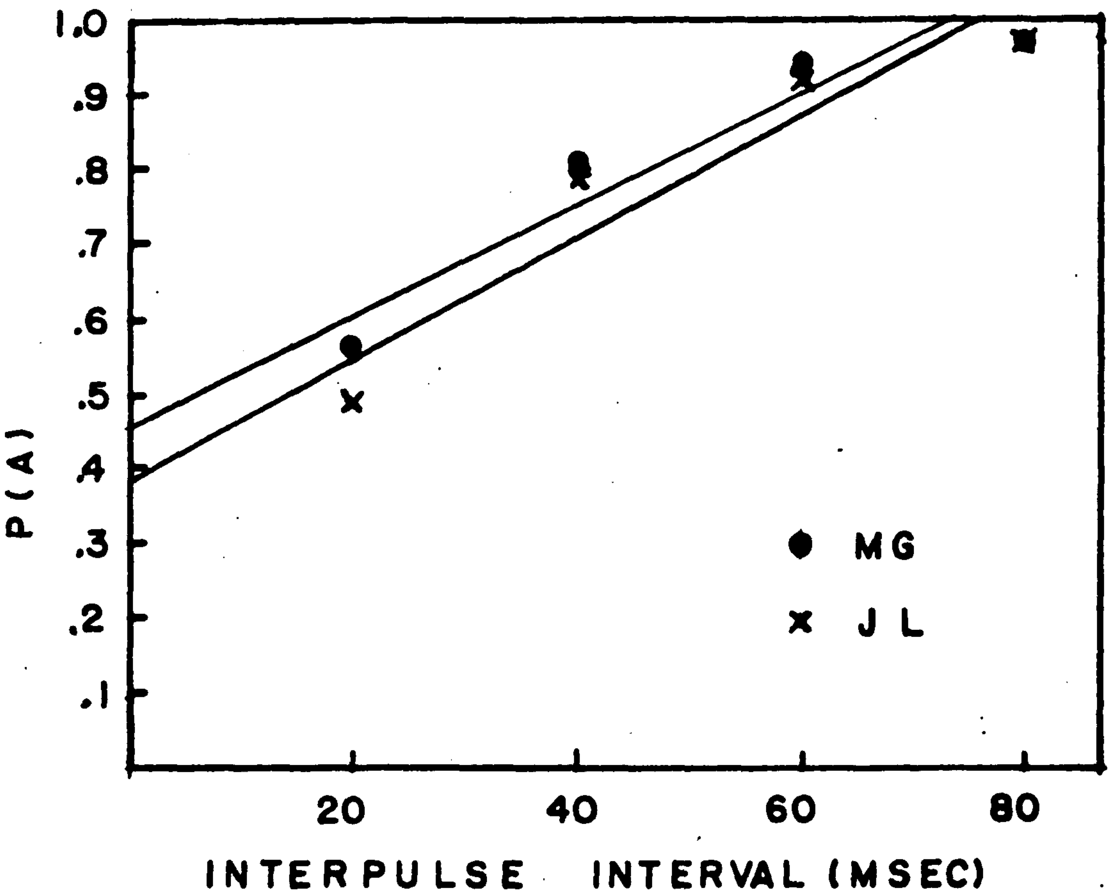
It has been found (Kietzman & Sutton, 1968) the various cues used by an observer to determine a "difference" (e.g., "shorter", flick", "color", and others) differed as the duration of the interpulse interval was increased. It was expected therefore, that the observer's criterion for "Different" would change as the duration of the interpulse interval was increased. The results confirmed this expectation in that the value of the minimal criterion point for reporting a difference, "Different, 3" (Different, very uncertain") did change as the interpulse interval was increased. Both observers became increasingly more lax as the interpulse interval was increased. Kietzman and Sutton had also found that at the longer interpulse intervals a greater variety of cues was used by the observers than at the shorter interpulse intervals. If this occurred in the present "same-different" procedure than the laxer criterion, i. e., increased likelihood that an observer will report "different", at the longer interpulse intervals can be explained by the increased probability of stimulus cues" occurrence.

The P(A) sensitivity values obtained from the "same-different" condition of the present rating scale study are comparable to the results obtained in studies that have employed the forced-choice method under similar experimental conditions (Kietzman, 1967; Kietzman & Sutton, 1968; Lewis, 1967; Lewis, 1968; Lewis & Mertens, 1971; Mertens & Lewis, 1972). In order to justify this conclusion, a brief discussion of how the comparison stimulus two-pluse threshold was determined is in order. Kietzman (1967) had first plotted his data to show accuracy of discrimination as a function of interpulse interval; lines of best fit for the data were determined by the method of least squares. He then specified the threshold as the interpulse interval that produced 50% correct judgements after correction for chance. Kietzman had used a three interval forced-choice procedure in which the proportion of correct discriminations expected by chance is .33; the adjusted "50% correct" (threshold) point is equal to an uncorrected proportion of correct responses, P(C) of .67. The same correction for chance calculations were made by Kietzman and Sutton (1968). Lewis (1967), Lewis (1968) and Lewis and Mertens (1971) had used an adaptive forced choice procedure (BUDTIF; Campbell, 1963) in which the "stepping rule" (Witherill & Levitt, 1965) had provided a corrected "50% correct" (threshold) estimation or P(C) equal to .75. Mertens and Lewis (1972) reported their forced choice data in uncorrected form and did not use a "threshold" analysis, although a threshold may be determined from their data.

Green and Swets (1966, pp. 113-114) have stated that the corrected P(C) in a forced choice procedure is equal to the area under the ROC curve, P(A), obtained in an equivalent rating scale (or yes - no) task. In view of this fact, and after considering that in many studies only the threshold value was reported (instead of the entire psychometric function) and that different "correction for chance" procedures were used, an attempt was made to compare P(C) values (equal in most cases to the corrected "50% correct" threshold point) with P(A) values in the present study. Figure 6 shows P(A) for the "same-different" condition as a function of interpulse interval; lines of best fit for the data were determined by the method of least squares for both observers. From Figure 6 the question was asked, "What interpulse interval value would estimate a P(A) equal to the P(C) reported in previous investigations?" Table 5 shows the P(C) and the corresponding interpulse interval at which the "threshold" was determined in previous studies (where the stimulus conditions approximated those used in the present study.) Table 5 also shows the interpulse interval in the present study that would estimate the P(A) equal to the P(C) by using the linear regression analysis. As shown in Table 5, these interpulse interval values are well within the range of the "threshold" values previously reported.

Why did Kietzman and Sutton (1968) find such large differences in two-pulse measures of temporal resolution obtained from classical psychophysical and forced-choice procedures while the present study which compared rating scale measures of sensitivity from procedures which were analogous toeach of these did not find differences? The most obvious answer is that Kietzman and Sutton compared a criterion-free measure derived from a forced-choice procedure with a noncriterion-free measure derived from high

Figure 6. P(A) for the "same-different" condition as a function of interpulse interval duration. Blackened circles indicate the obtained data points for MG; crosses for JL.



1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
P (A)
20
40
60
80
INTERPULSE INTERVAL (MSEC)
MG
JL

Table 5

Interpulse Interval Durations which Estimate P(A) to Equal P(C)

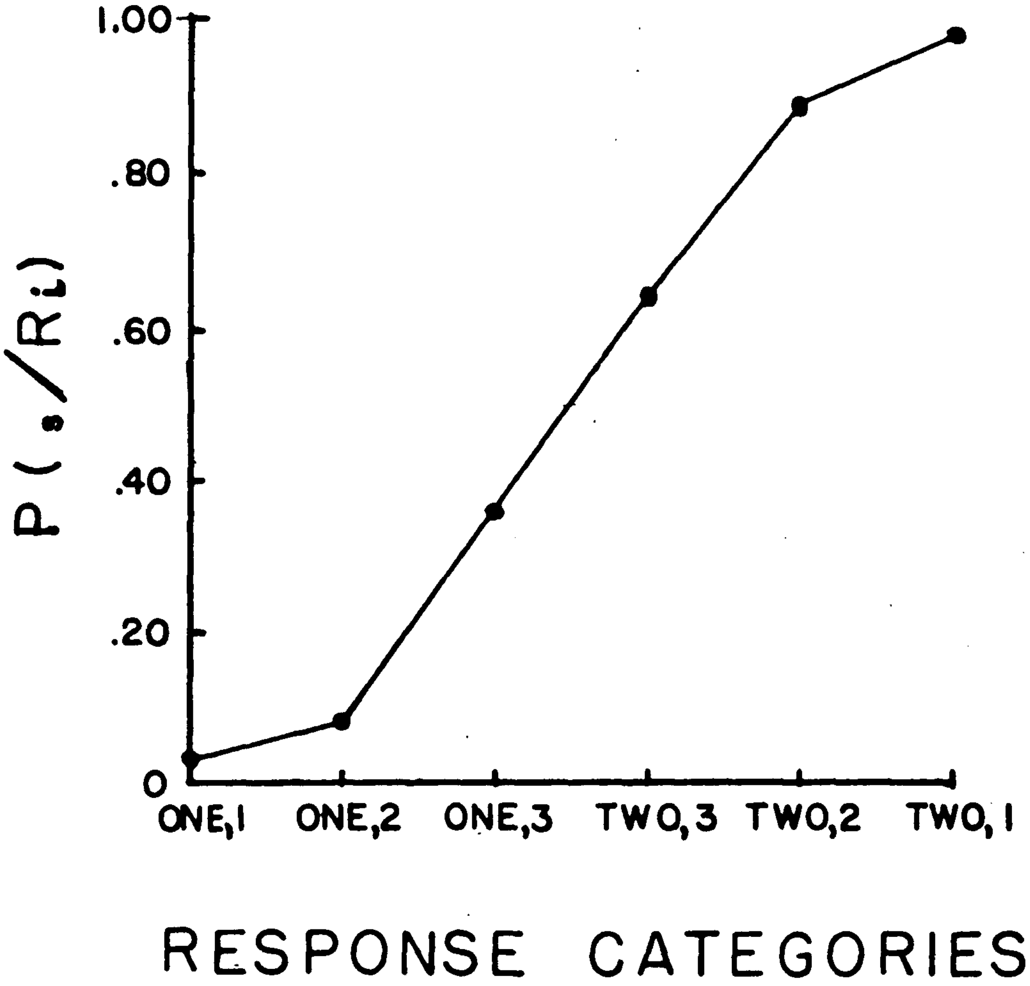
Study	P(C) at "threshold"	IPI at "threshold"	IPI to estimate P(A)= P(C)	
			MG	JL
Kietzman (1967)	.67	16-31	27.94	34.39
Kietzman & Sutton (1968)	.67	31	27.94	34.39
Lewis (1967)	.75	32-34	39.70	44.58
Lewis (1968)	.75	32-41	39.70	44.58
Lewis & Mertens (1971)	.75	41-47	39.70	44.58
Mertens & Lewis (1972)	.55[a]	15-21	10.29	19.10

[a]P(C) was reported, this does not correspond to a threshold value.

threshold theory (classical psychophysics). The present study compared measures that were assumed to be criterion-free and found no differences between their values. As noted above, the "same-different" results were similar to the forced-choice results reported by Kietzman and Sutton. Further discussion of the differences between high threshold theory and the rating scale task may help to clarify the differences in findings.

In the signal detection rating scale task it is expected that the observer will be able to report the a posteriori probability that a signal has occurred, i.e., that the observer is able to order events according to the likelihoods that they are signal. More concretely, it is expected that the observer is capable of reporting a subjective probability (Swets, Tanner, and Birdsall, 1961; McNicol, 1972, p. 170). If the observer is capable of performing the rating scale task, then P(s/Ri), the probability that a signal is presented and a response category is used, will increase monotonically as a function of the confidence that the event was a signal (a posteriori function; McNicol, 1972, p. 168). The a posteriori function for MG at an interpulse interval duration of 60 msec for the "one-two" condition is shown for illustrative purposes in Figure 7; similar functions would be found for the other interpulse interval durations, for the "same-different" condition for JL.

Figure 7. A posteriori function for MG at an interpulse interval duration of 60 msec for the "one-two" condition.


1.00
.80
.60
.40
.20
0
P (s / Ri)
ONE,1
ONE,2
ONE,3
TWO,3
TWO,2
TWO,1
RESPONSE CATEGORIES

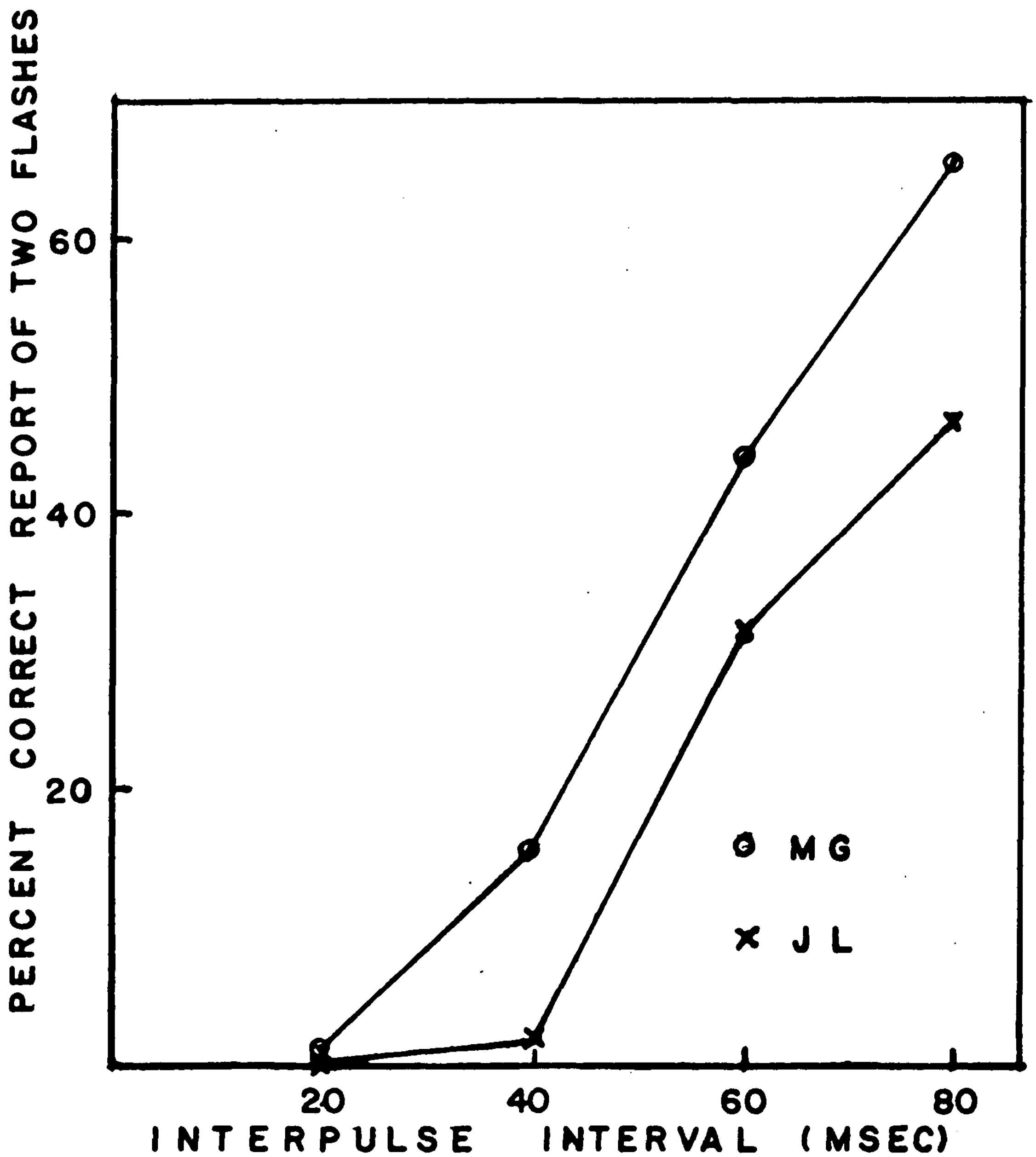
Signal detection theory assumes that there is no signal too weak to be detected while absolute threshold theory assumes that there fixed cutoffs, i.e., a "true" value of detection (the threshold) below which the observer is incapable of reporting a signal (usually 50% of the time). Hence, the a posteriori function for high threshold theory is very different from that in the rating scale task. In high threshold theory, any signal above threshold will be reported as a "Yes" (or as in this case, "Two") response allocated to the highest category of confidence (strictest category). Any signal below threshold would be reported as "No" ("One") and any method of assigning subthreshold events to confidence categories would be a random one (McNicol,1972, p. 172).

In order to make the "one-two" data of the present investigation comparable to high threshold data the cummulative probabilities of responses to signal and noise trials of the strictest response category, "Two, 1, ("Two, very certain) were used. They were adjusted using the following formula:

$$\text{Percent correct report} = \frac{P(S/s) - P(S/n)}{1 - P(S/n)} \times 100$$

This was done to provide "true" detection probabilities. Figure 8 shows the percent correct report of two flashes as a function of interpulse interval duration for both observers. The functions were very similar to the function reported by Kietzman and Sutton (1968); in fact, the function for MG was nearly identical. Moreover, the interpulse interval duration at 50% correct report of "two" would be more than

Figure 8. Percent correct report of two flashes as a function of interpulse interval duration for the strictest criterion point.


PERCENT CORRECT REPORT OF TWO FLASHES
20
40
60
20
40
60
80
INTERPULSE INTERVAL (MSEC)
MG
JL

double the interpulse interval duration to produce a P(A) of .67 for the "same-different" condition. Hence, when the observer is limited to use only the strictest criterion to report two flashes, as in classical psychophysical procedures, differences in the two-flash and the Comparison stimulus two-pulse "thresholds" may be found.

Both Class A and Class B observations (Brindley, 1960, pp. 144-150) were obtained in the present investigation. Class B observations are those in which the observer must subjectively report on some quality or intensity of sensation. The "one-two" condition provided Class B observations because the observers reported whether or not the stimulus presented produced the quality of "oneness" or "twoness". Class A observations are those in which the observer must report under certain experimental conditions whether or not a particular stimulus produced the same sensation or a different sensation. The "same-different" condition provided Class A observations because the observers were permitted to use all possible cues to report "different" or "same". The results of the present study tend to support the notion that both classes of observations are useful provided that certain criteria which were met in the present study are met. The Class B observations should be obtained under the same viewing conditions as Class A observations (Brindley). The Class B observations should produce the same results as Class A observations (Brindley). When feasible, feedback should be given for both classes of observations (Cornsweet & Pinsker, 1965) if the results obtained from these observations are to be compared. In addition, some valid interpretation of the observer's decision making process should be derived from the observer's subjective report.

Control Experiment

An attempt was made in the present study to keep the "same-different" rating scale procedure as comparable as possible to the forced-choice procedures that have been used in other studies of two-pulse temporal resolution (Kietzman, 1967; Kietzman & Sutton, 1968; Lewis, 1967; Lewis, 1968; Lewis & Mertens, 1971; Mertens & Lewis, 1972). however, several aspects of the "same-different" and forced-choice procedures were different. Although in both procedures Class A observations (Brindley, 1960, pp. 144-150) were obtained, how these observations were reported by the observer was somewhat different. In the same-different" procedure the observer was asked to report whether the second stimulus appeared to be "same" or "different" from the first (comparison) stimulus and also to report the certainty of his response. In the forced-choice procedures the observer was asked to report in which interval the " stimulus was presented. The stimuli were also presented somewhat differently. In the "same-different" procedure the observer had to make a decision after a single observation interval following the comparison stimulus. In the forced-choice procedures the observer was forced to make a decision after two or more observation intervals following the comparison stimulus. In view of these differences, there was some question as to whether the "same-different" and forced-choice procedures would give comparable results, even though the results of the main experiment had suggested that they would (see Table 5).

Taking advantage of the fact that the uncorrected percentage of correct responses, P(C), in a two interval forced-choice task is equal to the distribution-free sensitivity measure, P(A), in an equivalent rating scale task, a control experiment was specifically designed to compare these measures. Therefore, a rating scale procedure and a two interval forced-choice procedure provided the experimental conditions for the control experiment. The rating scale procedure provided signal detection measures for two-pulse discrimination as "same-different" responses measured with reference to a single pulse comparison stimulus. This was the identical "same-different" procedure that was used in the main experiment. The two interval forced-choice procedure provided signal detection measures for two-pulse discrimination as "different" responses from single pulse comparison stimuli. The forced-choice procedure was similar to that used in previous studies of two-pulse temporal resolution. The signal detection measures of sensitivity obtained from the two procedures were compared to determine if they would produce similar results.

V. Control Experiment: Method

Observer

Observer JL, the author, also participated in the main experiment. Although JL was thoroughly familiar with the design and purpose of the study, he was naive as to the order of presentation of stimulus conditions and to the results of any one testing session; data analysis was not begun until the last testing session was completed.

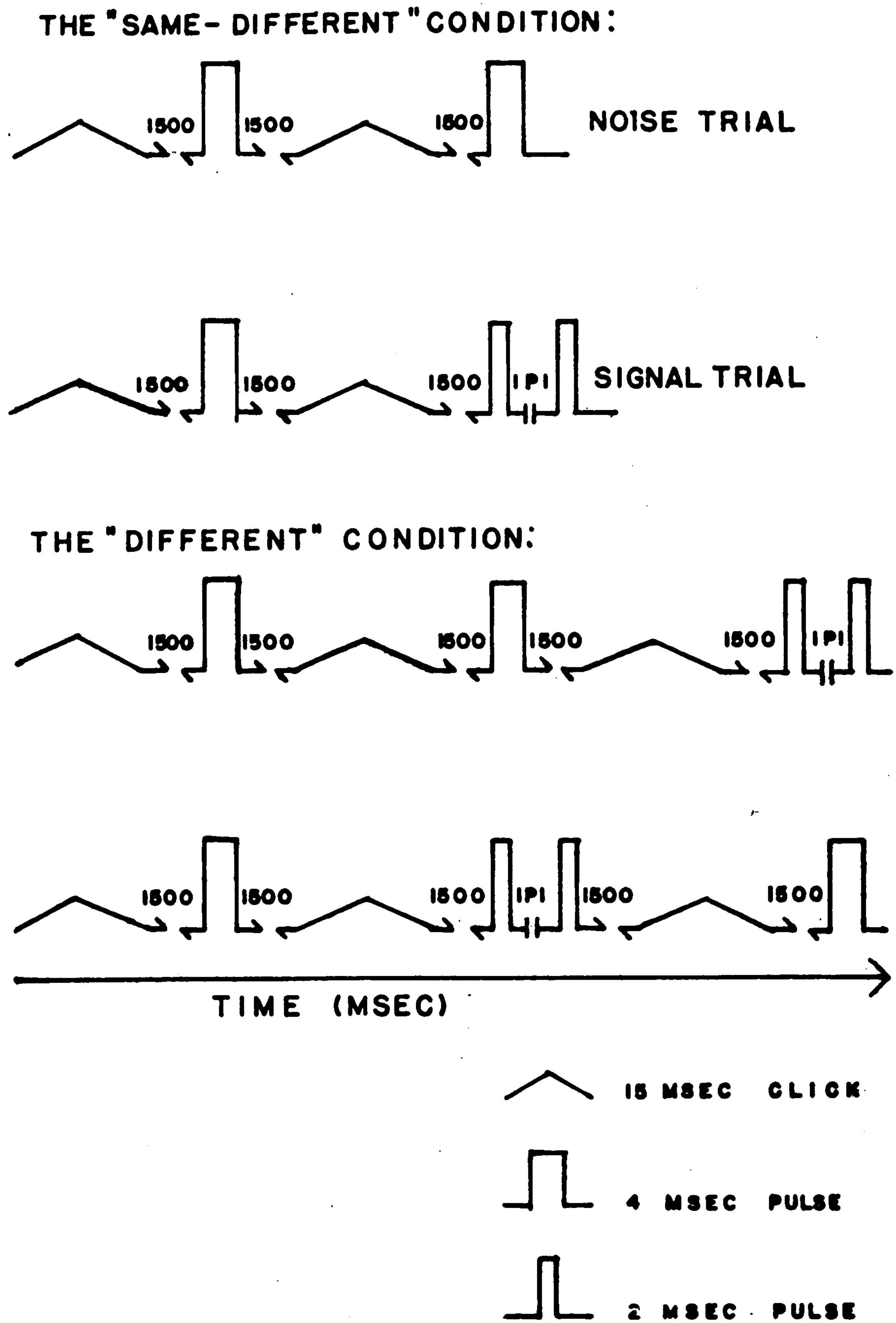
Apparatus

The instruments used in the main experiment for monitoring light intensity and for presenting stimuli were again used in the present study. Pulse luminance was the same as the main experiment .43 mL. (1.37 cd/m^2).

Procedure

Two signal detection procedures provided the experimental conditions for the present study. One procedure was designed to measure two-pulse discrimination as "same-different" responses measured with reference to a single pulse comparison stimulus using a rating scale procedure (the "same-different" condition). The other procedure was designed to measure two-pulse discrimination as "different" responses from single pulse conparison stimuli in a two interval forced-choice procedure (the "different" condition). The sequence of events for both conditions is shown in Figure 9. A noise trial for the "same-different" condition consisted of a 15 msec warning click followed by a 1.5 sec foreperiod and then a 4 msec light pulse (the comparison stimulus); after 1.5 sec this sequence was repeated with the second 4 msec light pulse being the noise stimulus. A signal trial for the

Figure 9. Trial sequences for the "same-different" and different" conditions. IPI indicates an interpulse interval duration of 20 or 40 msec. See text for details.


THE "SAME- DIFFERENT" CONDITION:
1500
1500
1500
NOISE TRIAL
1500
1500
1500
IPI
SIGNAL TRIAL
THE "DIFFERENT" CONDITION:
1500
1500
1500
1500
1500
IPI
1500
1500
1500
IPI
1500
1500
TIME (MSEC)
15 MSEC CLICK
4 MSEC PULSE
2 MSEC PULSE

"same-different" condition consisted of a 15 msec warning click followed by a 1.5 sec foreperiod and then two equal energy 2 msec light pulses separated by an interpulse interval (the signal stimulus). The trial sequence for the "different" condition consisted of three successive observation intervals; each observation interval was preceded by a 15 msec click and a 1.5 sec foreperiod. A 1.5 sec dark interval followed the observation intervals. In two of these intervals a 4 msec light pulse was presented. In the other interval two equal-energy 2 msec light pulses separated by an interpulse interval (the test stimulus) were presented. A 4 msec light pulse (the comparison stimulus) was always presented in the first observation interval, whereas the two equal-energy 2 msec pulses and the other 4 msec pulse were presented randomly from trial to trial in the second or third observation intervals. Two interpulse interval durations were used for both conditions:20 and 40 msec.

During each session JL was seated and positioned on a chin rest and given 8 min to dark adapt. The fixation lines were then adjusted so that they were just visible. For the "same-different" condition he was to report whether the stimulus following the second click was "same" or "different" from the stimulus following the first click. He was also to report the certainty of his response by using the following code: "1" signified "very certain", "2" signified "moderately certain", and "3" signified "very uncertain". He was permitted to use any and all possible cues to determine a "difference". For the "different" condition, JL was to report in which observation interval, the second or the third, the "different" stimulus most likely occurred. Again, he was permitted to use any and all possible cues to determine a "difference". For both conditions, JL was given feedback, i.e., when he correctly reported "same" or "different" or the interval in which the "different" stimulus was presented he was told "Yes" by the experimenter.

The "same-different" and "different" conditions were counterbalanced throughout a session (ABBA). Each component of the counterbalancing, e.g., A, contained four blocks of 16 trials. One interpulse interval was used throughout a block and was randomly selected in both conditions. For the "same-different" condition there were approximately 50% signal and 50% noise trials. There were 256 trials per session; 128 per condition ("same-different" or "different"). Four experimental sessions were conducted.

VI. Control Experiment: Results

The results of the control experiment indicated that there were no differences in sensitivity between the "same-different" and "different" conditions; they produced similar results.

The uncorrected percentage of correct responses, P(C), in the present two interval forced-choice task was compared to the distribution-free measure, P(A), in the rating scale task; these values may be compared directly (Green and Swets, 1966, P. 47), and are shown in Table 6. In as much as both P(C) and P(A) are proportions, Gibson (Note 3) suggested that a z-test for independent proportions be used to test the significance of the difference between P(C) and P(A) at each interpulse interval. For P(C), N = 256, the total number of trials. For P(A), N cannot exceed 256, the total number of trials. If the null hypothesis cannot be rejected with the N value of P(A) equal to its maximum (256), it cannot be rejected for any value of N less than that. Table 6 also shows obtained values of z at each interpulse interval . For all values of z the null hypothesis was not rejected ($p > .01$).

Sensitivity measures d' and d'M were also compared at each interpulse interval for both conditions, and this comparison is shown in Table 6. For the "different" condition d' was determined using the following formula:

$$d'_{\text{Forced-choice}} = \sqrt{2} \times z(C)$$

where z(C) is the z-score of P(C). For the "same-different" condition d'M was determined in a similar manner to the way it was determined in the main experiment. It is interesting that d' for the forced-choice ("different") procedure is larger than for the rating scale ("same-different") procedure by a factor of

Table 6

Comparison of d'M and P(A) for the "Same-Different" Condition with d' and P(C) for the "Different" Condition

IPI	d'M	d'	P(A)	P(C)	z[a]	p
20	.357	.544	.636	.652	.382	.76
40	.990	1.294	.792	.816	.662	.54

about $\sqrt{2}$. This result would be predicted by signal detection theory.

VII. Control Experiment: Discussion

The results of the control experiment provide evidence that the "same-different" (rating scale) procedure, that was also used in the main experiment, produces equivalent measures of sensitivity to the "different" (forced-choice) procedure, which was designed similarly to other forced-choice studies of Lewis, 1968, Lewis and Mertens, 1971, and in particular, Mertens and Lewis, 1972, where almost the identical procedure was used. Regardless of whether the observer was to report "same" or "different" (as in the rating scale procedure) or the interval in which the "different" stimulus appeared (as in the forced-choice procedure), the results are the same. The results are also the same independent of whether a single observation interval with a comparison stimulus or two observation intervals with a comparison stimulus was presented.

For observer JL, the P(A) value at an interpulse interval duration of 20 msec in the control experiment (.64) is somewhat different from the value obtained in the main experiment (.49) at the same interpulse interval duration. The P(A) value at an interpulse interval of 40 msec is the same for both experiments (.79). One reason for this difference across the two experiments may be due to some effect of practice; JL performed better at the more difficult level of discrimination (the 20 msec. interpulse interval) due to practice effects (Kietzman, Note 4).

Mertens and Lewis (1972) compared a two-interval forced-choice procedure with an equivalent yes-no procedure for two-pulse temporal resolution. They reported the P(C) and P(A) sensitivity measures to be nearly identical, as in the present study. If the results of the present study can be conpared and generalized with those of Mertens and Lewis, then it may be said that both the rating scale and yes-no procedures give similar results to the two interval forced-choice procedure for measures of visual two-pulse temporal resolution. These results would be predicted from signal detection theory and point to the internal consistency of the theory.

VIII. General Discussion

In the main experiment of the present study an attempt was made to use the most important features of classical psychophysical and forced-choice studies of two-pulse visual temporal resolution studies in analogous signal detection rating scale procedures. The most salient characteristics of the classical psychophysical procedures are a single observation interval, the absence of a comparison stimulus, and "one-two" flash responses which are Class B observations. The classical psychophysical procedures do not permit feedback to be given to the observer because all of the stimuli are two pulse stimuli. The experimenter must therefore rely upon the observer's subjective report of "two flashes". In addition, there is no objective way of determining whether the instructions to the observer are followed. Furthermore, the measures of two-pulse discrimination (i.e., the two-flash threshold) obtained from these procedures are contaminated with criterion. In the "one-two" procedure of the present study were a single observation interval, with no comparison stimulus, and "one-two" flash responses which were Class B observations; all of these features were analogous to the classical psychophysical procedures. However, feedback was possible because both one and two pulse stimuli were presented. Feedback provided some indication that the instructions were followed by the observer. The fact that the minimum criterion setting for reporting two flashes did not change as the interpulse interval was increased was also an indication that the instructions were followed by the observer.

Furthermore, the sensitivity measures obtained from this signal detection rating scale procedure are criterion-free.

The most prominent features of the forced-choice studies of two-pulse visual temporal resolution which were retained in the "same-different" condition of the present study were that at least one comparison and one test stimulus were presented on a single trial, Class A observations, and discrimination measures which are criterion-free. Feedback has been given in many forced-choice two-pulse resolution studies (e.g., Lewis, 1968; Lewis & Mertens, 1971; Mertens & Lewis, 1972) and it was also given in the present investigation in the "same-different" procedure. There are differences between the forced-choice design and the rating scale design so therefore, there were differences between the previous forced-choice procedures and the "same-different" (rating scale) procedure. Although in both procedures Class A observations were obtained, how these observations were reported was somewhat different. In the forced-choice procedures the observer was asked to report in which interval the "different" stimulus with reference to the comparison stimulus was presented. In the "same-different" procedure the observer was asked to report whether the second stimulus appeared to be "same" or "different" from the first stimulus (i.e., the comparison stimulus). In the forced-choice procedures the observer was forced to make a decision after two or more observation intervals following the comparison stimulus. In the "same-different" procedure the observer had to make a decision after a single observation interval following the comparison stimulus. These differences do not appear to be important because the "same-different" procedure had produced results comparable to the previous forced-choice studies (Table 5). Moreover, the results of the control experiment supported the hypothesis of no difference.

It was thought that the extra information given by the comparison stimulus in the "same-different" condition would have improved sensitivity and that sensitivity differences would have been obtained between the "same-different" and "one-two" conditions. However, the "one-two" condition was not completely free from a comparison stimulus. On 50% of the trials a single pulse (the noise stimulus) was presented and there was nothing in the experimental design to prevent the observer from using the memory of this stimulus and using it as the "comparison stimulus".

The Class B observations of the "one-two" procedure conform to the most stringent criteria for their conversion to Class A observations. The Class B observations produced from the "one-two" procedure were obtained under the same experimental conditions as the Class A observations produced from the "same-different" procedure. According to Brindley (1960, pp. 144-150) when Class B observations are not produced in this manner they are more likely to be confounded with response bias and therefore are less likely to provide observations which represent fundamental relations between stimuli and sensation. The fact that feedback was given in the "one-two" condition may make this data more reliably correlated with the effect of manipulating the interpulse interval than subjective aspects of "twoness" (Cornsweet & Pinsker, 1965).[11] In studies that have found large differences between the two-pulse discrimination measures obtained from classical psychophysical and forced-choice procedures (i.e., Kietzman, 1967; Kietzman & Sutton, 1968) the comparisons were made from the results of separate studies under comparable, but not identical viewing conditions. Feedback could not be given in the classical psychophysical procedure because a two-pulse stimulus was presented on every trial.

Using Brindley's and Cornsweet and Pinsker's line of reasoning, the Class B observations produced from the classical psychophysical studies are more likely to be confounded with the effect of observer biases regarding "twoness" than the Class B observations produced from the "one-two" condition of the present study. It is suggested here that the Class B observations produced in the present study may be of greater value with regard to the visual system's ability to resolve two pulses of light. It is also suggested that the classical psychophysical

---

[11]Certainly this is only one point of view regarding the effect of feedback in psychophysical research. Indeed, the effect feedback may have on psychophysical measures is a matter of controversy (see Carterette, Friedman & Wyman, 1966).

studies are of value with regard to information about subjective aspects of "twoness" discrimination.

Implicit within the distinction of Class A and Class B observations is the issue of criterion-free and noncriterion-free psychophysical measures. The present study compared criterion-free discrimination (sensitivity) measures between the "one-two" and "same-different" conditions and found no differences between their values. Kietzman (1967) and Kietzman and Sutton (1968) compared a criterion-free threshold measure derived from the forced-choice procedure with a noncriterion-free threshold measure derived from classical psychophysics and found large differences in their values. It was shown in the present study that when the "one-two" data for the strictest response category were adjusted to provide measures of discrimination comparable to the threshold measures of classical psychophysics that large differences in "threshold" (see Figure 9 and Table 5) would be found. This result suggests that when the observer adopts a strict criterion for reporting "two flashes" large differences in the classical psychophysical and forced-choice threshold values will be obtained. Support for this contention comes from the results of a series of studies by Tong (1972), who had reported a significantly lower two-flash threshold for observers that were instructed to adopt a lax criterion than for observers who were instructed to adopt a strict criterion. Further support also comes directly from the studies of Kietzman and Sutton (1968). They had reported that one observer with the forced-choice method at an interpulse interval of 46 msec was correctly discriminating the two pulses nearly 100% of the time with "flick" being the predominant cue at this interpulse interval. The "flick" cue was the closest cue to "twoness" in view of the fact that "twoness" was not an available cue report category for the observer to use. The same observer had a constant stimulys threshold (50% report of two flashes) value of 73 msec interpulse interval duration. Kietzman and Sutton suggested that one of the reasons why the "flick" cue was not used at 46 msec interpulse interval duration with the constant stimulus method was that the observer was instructed to report "twoness" and ignored "flick". This suggests a strict criterion in that the observer was not willing to report "twoness" to a "flick" until he was "very certain" of "two". This example also suggests an obvious relation between the problem of stimulus attributes (multiple cues in discrimination) and criterion, and how these factors may interact in this case to increase threshold values.

Kietzman and Sutton suggested that another reason why this observer did not use the "fkick" cue at the 46 msec interpulse interval was that in the constant stimulus method the observer did not have other "nonflick" stimuli presented close enough in time for comparison. In the constant stimulus method used the observer was always presented two pulses on every trial and no feedback was given or even could be given. There was no indicator of response accuracy, i.e., a single pulse. According to Cornsweet and Pinsker (1965) the use of feedback may help to provide observations that are more likely to be free from response bias. However, feedback could not be given in this case without a single pulse stimulus. There is the possibility that without feedback and without the opportunity to observe "one" the observer was more reluctant to report "two". That is, in this situation the observer may have adopted a decision stragedy to report "two" only when he was ver certain of "two", i.e., a strict criterion. The total effect would be to increase threshold value. As mentioned above, the "one-two" condition of the present study was not completely free of a comparison stimulus and feedback was given. These factors, as well as the use of the rating scale task did not foster the adoption of a strict criterion in the present study. Indeed, these factors may account for the differences in results between that of the present study and that of Kietzman and Sutton.

A relation between criterion and multiple cues in discrimination is suggested from the present "same-different" data and the cue reports obtained from the Kietzman and Sutton (1968) study. It was shown in the present study that the minimum criterion setting for reporting "different" changed as the interpulse interval was increased; both observers became increasingly more lax as the interpulse interval was increased. Kietzman and Sutton reported that observers used a greater variety of cues at the longer interpulse intervals that was not present at the shorter interpulse intervals.

At the longer interpulse intervals the probability of "stimulus cues occurence" is increased and therefore an increase in the likelihood of reporting "different" (i.e., a more lax criterion) would be expected.

The results of the control experiment indicate that the "same-different" (rating scale) procedure produces a sensitivity measure, P(A), equal to the P(C) sensitivity measure produced from an equivalent two interval forced-choice procedure (the "different" procedure). The d' sensitivity measure derived from the forced-choice procedure was larger than the d'M sensitivity measure derived from the rating scale procedure by a factor of about the $\sqrt{2}$. The relationship that usually been reported for d' obtained from a two interval forced-choice procedure and d' obtained from an equivalent yes-no procedure is that the forced-choice d' is larger than the yes-no d' by a factor of $\sqrt{2}$. (McNicol, 1972, pp. 67-69). Since equivalent sensitivity measures have been obtained from comparable rating scale and yes-no procedures (e.g., Egan, Schulman & Greenberg, 1959,[12] d'M appears to be a good overall sensitivity measure for the rating scale task and comparable to the d' measure obtained from the yes-no task. Mertens and Lewis (1972) compared a two interval forced-choice procedure with an equivalent

---

[12]There are exceptions to this general finding. (See footnote 1

yes-no procedure for two-pluse temporal resolution. They reported the P(C) and the P(A) sensitivity measures obtained from these procedures to be equal. When the results of the control experiment of the present study are compared and generalized with those of Mertens and Lewis, it may be said that the rating scale and yes-no procedures produce similar results to the two interval forced-choice procedure for two-pulse visual temporal resolution. All of the above results regarding the control experiment would be predicted from signal detection theory and point to the internal consistency of the theory. These results should encourage additional efforts to apply signal detection procedures in similar studies of temporal resolution. In particular, a signal detection analysis may provide a definitive statement as to how pulse luminance and duration changes effect two-pulse measures of visual temporal resolution.

In summary, the present study provided a signal detection analysis of two-pulse visual temporal resolution thresholds using two rating scale procedures. One procedure, the "one-two" classical psychophysical and forced-choice "threshold" data the same large differences in "threshold magnitude previously reported were obtained when only the strictest criterion setting of the analogous rating scale-classical psychophysical procedure (the "one-two" procedure) was used. Based on the conditions of the present study it is concluded that there is no sensitivity difference between these measures and that the large differences in threshold magnitude previously reported were due to an effect of criterion.

Appendix

Use of terms and symbols in binary decision rating scale procedures. In a signal setection rating scale procedure when a binary response is given by the observer (i.e., the observer can respond "signal" (S) or "noise" (N) as well as express the certainty of his response) it is usually inferred that a response of very high certainty that noise was presented is equal to a response of very low certainty that a signal was presented. In this system two conditional probabilities are determined the probability of a hit, P(S/s) and the probability of a false alarm, P(S/n). Using 1 to signify "very certain", 2 to signify "moderately certain", 3 to signify "very uncertain", and the binary response S or N, the following six point scale can be constructed from high certainty that a signal was presented to high certainty that a noise was presented: S,1 - S,2, - S,3 - N,3 - N,2 - N,1. Table 7 shows how the conditional probabilities are determined using this system. Table 7 shows for example, that a response of N,3 when a signal was presented is considered a hit. This is because it is inferred that a response of N,3 is equal to a hypothetical response of S,6. A response of N,3 when noise is presented is considered a false alarm because it is inferred that the observer is using the hypothetical S,6 category.

Another system, suggested by Kietzman (Note 1) and adopted in the present study gives justice to the actual responses of the observer and not to inferred responses, it

Table 7

Traditional Usage of Binary Decision Rating Scale Terms and Symbols of Conditional Probabilities

Stimulus event	Response event	
	S 1, 2, or 3	N 1, 2, or 3
S	Hit	Hit
	P(S/s)	P(S/s)
n	False alarm	False alarm
	P(S/n)	P(S/n)

is presented in Table 8. Using the same example given above, a response of N,3 when a signal was presented is considered a miss, which it actually is. A response of N,3 when noise is presented is considered a correct rejection, which it actually is. This system implies that the binary decision ROC curve should not be plotted with the cummulative probability of a hit along the ordinate (because all of the responses are not actually hits) and the cummulative probability of a false alarm along the abscissa (because all of the responses are not actually false alarms. It does imply however that the binary decision ROC curves should be plotted with the cummulative probability of responses to signal trials, P(S/s) or P(N/s), along the ordinate and the cummulative probability of responses to noise trials, P(S/n) or (N/n), along the abscissa.

Table 8

Usage of Binary Decision Rating Scale Terms and Symbols of Conditional Probabilities Adopted in the Present Study

Stimulus event	Response event	
	S 1, 2, 3	N 1, 2, 3
s	Hit	Miss
	P(S/s)	P(N/s)
n	False alarm	Correct rejection
	P(S/n)	P(N/n)

Reference Notes

1. Kietzman, M.L. Personal communication, June 1977
2. Gibson, W. A. Personal communication, June 1977
3. Gibson, W. A. Personal communication, March 1978
4. Kietzman, M.L. Personal communication, April 1978

Selected Bibliography

Bartley, S.H. Principles of perception. New York: Harper and Brothers, 1958.

Bartley, S. H., & Wilkinson, F.W. Certain factors in producing complexity of response to a single pulse of light. Journal of Psychology, 1953, 9, 299-306.

Blackwell, H.R. Psychophysical thresholds: experimental studies of methods of measurement. Bulletin of the Engineering Institute of the University of Michigan, Number 36, 1953.

Boring, E. G. The physical dimensions of consciousness. New York: Dover Publications, 1963.

Boynton, R. M. Discrimination of homogeneous double pulses of light. In D. Jameson and L.M. Hurvich (Eds.) Handbook of psychophysiology, Vol. 7: Visual psychophysics, Chap 9. Berlin: Springer - Verlag, 1972.

Brindley, G.S. Physiology of the retina and the visual pathway. London: Edward Arnold, 1960.

Bruder, G.E., & Kietzman, M. L. Visual temporal integration for threshold, signal detectability, and reaction time. Perception and Psychophysics, 1973, 13, 293-300.

Burros, R. H. & Gibson, W. A. A solution of case III of the Law of Comparative Judgement. Psychometrika, 1954 19, 57-64.

Campbell, R. A. Detection of a noise signal of varying duration. Journal of the Acoustical Society of America, 1963, 35, 1732-1737.

Carterette, E. C., Friedman, M.P. & Wyman, M.J. Feedback and psychophysical variables in signal detection. Journal of the Accoustical Society of America, 1966, 39, 1051-1055.

Clark, W.C. Relations between the discrimination of single and multiple (CFF) dark intervals. Unpublished manuscript, New York State Psychiatric Institute and College of Physicians and Surgeons, Columbia University, 1963.

Clark, W. C., & Dillon, D. J. SDT analysis of binary decisions and sensory intensity ratings to noxious thermal stimuli. Perception and Psychophysics, 1973, 13, 491-493.

Clark, W. C. & Mehl, L. Values of d' obtained by various SDT procedures are not equal when thermal stimuli are judged. Proceedings, 79th Annual Convention of the American Psychological Association, 1971, 9-10.

Clark, W. C. & Pickett, R. Luminance required to detect the presence of a dark interval between two pulses of light. American Psychologist, 1959, 14, 412.

Cornsweet, T. N., & Pinsker, H. M. Luminance discrimination of brief flashes under various conditions of adaption. Journal of Physiology, 1965, 176, 294-310.

Davy, E. The intensity - time relation for multiple flashes of light in the peripheral retina. Journal of the Optical Society of America, 1952, 42, 937-941.

Dunlap, K. The shortest perceptible time-interval between two flashes of light. <u>Psychological Review</u>, 1915, <u>22</u>, 226-250.

Egan, J.P., & Clarke, F.R. Psychophysics and signal detection. In j.B. Sidowsky (Ed.) <u>Experimental methods and instrumentation in psychology</u>, New York: McGraw - Hill, 1966.

Egan, J.P., Schulman, A.I., & Greenberg, G.Z. Operating characteristics determined by binary decisions and by ratings. <u>Journal of the Acoustical Society of America</u>, 1959, <u>31</u>, 768-773.

Emmerich, D.S. ROCs obtained with two signal intensities presented in random order, and a comparison between rating and rating ROCs. <u>Perception and Psychophysics</u>, 1968, <u>3</u>, 35-40.

Geldard, F.A. <u>The human senses</u>. New York: John Wiley & Sons, 1953.

Graham, C.H. Some fundemental data. In C.H. Graham (Ed.) <u>Vision and visual perception</u>. New York: John Wiley & Sons, 1965.

Graham, C.H., & Margaria, R. Area and the intensity-time relation in the peripheral retina. <u>American Journal of Physiology</u>, 1931, <u>48</u>, 644-653.

Green, D. M. & Swets, J. A. Signal detection theory and psychophysics. New York: John Wiley & Sons, 1966.

Grice, R. G. Dependence of empirical laws upon the source of experimental variation. Psychological Bulletin, 1966, 66, 488-498.

Gruzelier, J.A., & Venables, P.H. The two-flash threshold, sensitivity, and in normal subjects and schizophrenics. Quarterly Journal of Experimental Psychology, 1974, 26, 1-11.

Hochaus, L. A table for the calculation of d' and Psychological Bulletin, 1972, 77, 375-376.

Hodos, W. Nonparametric index of response bias for use in detection and recognition experiments. Psychological Bulletin, 1970, 74, 351-354.

Ireland, F. H. An experimental investigation of the monocular and binocular critical two pulse interval. Unpublished doctoral dissertation, Fordham University, 1954.

Kietzman, M. L. Two-pulse measures of temporal resolution as a function of stimulus energy. Journal of the Optical Society of America, 1967, 57, 809-813.

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

Kietzman, M.L., & Smith, V.C. Discrimination of single from double light pulses as influenced by the location and duration of the dark period between the double pulses. Paper presented at the Eastern Psychological Assoc. meetings. Atlantic City, New Jersey, April, 1965.

Kietzman, M.L. & Sutton, S. The interpretation of two-pulse measures of temporal resolution in vision. Vision Research, 1968, 8, 287-302.

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

Kramer, R. The effect of energy manipulations on two-pulse measures of temporal resolution. Unpublished master's thesis, Hofstra University, 1969.

Levitt, H., & Treisman, M. Control charts for sequential testing. Psychometrika, 1969, 34, 509-518.

Lewis, M.F. Two-flash thresholds as a function of luminance in the dark-adapted eye. Journal of the Optical Society of America, 1967, 57, 814-815.

Lewis, M.F. Two-flash thresholds as a function of flash luminance and area. Perception and Psychophysics, 1968, 4, 241-244.

Lewis, M.F.,& Mertens, H.W. Two-flash thresholds as a function of comparison stimulus duration. Areospace Medicine, 1971, 42, 393-395.

Mahneke, A. Foveal discrimination measured with two successive light flashes. Acta Ophthalmologica, 1958, 36, 3-11.

Marascuilo, L.A. Extensions of the significance test for one-parameter signal detection hypotheses. Psychometrica, 1970, 35, 237-243.

McNicol, D. A primer of signal detection theory. London: George Allen & Unwin Ltd., 1972.

Mertens, h.W., & Lewis, M.F. Temporal discrimination of two-pulse stimuli as a function of psychophysical method and luminance. Journal of the Optical Society of America, 1972, 62, 823-826.

Nilsson, T.H. Two-pulse difference thresholds and their implications about sensory processes which limit temporal acuity. Unpublished master's thesis, University of Alberta, 1968.

Nilsson, T.H. Two-pulse-interval vision thresholds. Journal of the Optical Society of America, 1969, 59, 753-756.

Pearson, L.A., & Tong, J.E. Two-flash fusion threshold: the influence of age, psychophysical method, instructions, viewing conditions, sex and subject variability. British Journal of Psychology, 1968, 59, 407-414.

Purcell, D.G., & Stewart, A.L. The two-flash threshold: an evaluation of critical-duration and visual persistence hypotheses. Perception and Psychophysics, 1971, 9, 61-64.

Simpson, A.T., & Fitter, M.J. What is the best index of ability? Psychological Bulletin, 1973, 80, 481-486.

Stevens, S.S. The attributes of tones. Proceedings of the National Academy of Science, 1934, 20, 457-459.

Swets, J.A. Indices of signal detectability obtained with various psychophysical procedures. Journal of the Acoustical Society of America, 1959, 31, 511-513.

Swets, J.A. The relative operating characteristic in psychology. Science, 1973, 182, 990-1000.

Swets, J.A., Tanner, W.P., Jr. & Birdsall, T.G. Decision processes in perception. Psychological Review, 1961, 68, 301-340.

Theodor, L.H. The detectability of a brief gap in a pulse of light as a function of its temporal location within the pulse. Perception and Psychophysics, 1972, 12, 168-170.

Tong, J.E. Alcohol, arousal and temporal functions: signal detection experiments with two flash fusion and studies of time judgement and the effects of alchol. Unpublished manuscript, University of Guelph, Ontario, 1972.

Treisman, M. Relation between signal detectability theory and the traditional procedures for measuring thresholds: an addendum. Psychological Bulletin, 1973, 79, 45-47.

Treisman, M., & Watts, T.R. Relation between signal detectability theory and the traditional procedures for measuring thresholds: estimating d' from results given by the method of constant stimuli. Psychological Bulletin, 1966, 66, 438-454.

Wetherill, G.G., & Levitt, H. Sequential estimation of points on a psychometric function. <u>British Journal of Mathematical and Statistical Psychology</u>, 1965, <u>18</u>, 1-10.