THE DETECTION OF VISUAL INTENSITY DIFFERENCES BY PIGEONS

WILLIAM HODOS AND JAMES C. BONBRIGHT, JR.¹

UNIVERSITY OF MARYLAND AND WALTER REED
ARMY INSTITUTE OF RESEARCH

Pigeons were trained in a conditional discrimination procedure to discriminate between visual stimuli that varied in intensity. The magnitude of the intensity difference ranged from 0.80 to 0.05 logarithmic units. Psychometric functions were calculated from the data and the mean difference threshold for the 11 subjects was approximately 0.12 logarithmic units. A signal-detection analysis of the data suggested that stimuli that were below the calculated threshold were detectable to the subjects.

The recent concern with problems of attention and/or stimulus control (Terrace, 1966; Gilbert and Sutherland, 1969) has pointed to the need for more information about the sensory capacities of experimental organisms. In such experiments, a crucial aspect is for the experimenter to minimize the possibility that extraneous stimuli or stimulus differences may be correlated with the occurrence of reinforcement or punishment. In visual experiments, the most common sources of extraneous stimuli are differences in intensity or brightness. Pigeons are often the subjects of such visual studies and yet only minimal data are available about the limits of pigeons' ability to detect differences in intensity.² Two previous studies of intensity difference thresh-

old have been carried out. The first by Mentzer (1966) is difficult to interpret because of the large variability across subjects. The second, by Shumake, Hatfield, and Smith (1966) presented actual data for only a single subject. They reported however that two additional subjects performed comparably.

An additional purpose of the present experiment was to compare the traditional psychometric function and the indicators of sensitivity derived from it, *e.g.*, the point of subjective equality, difference threshold, *etc.*, with the indicators of sensitivity derived from the so-called modern psychophysics or signal-detection analysis (Swets, 1964; Green and Swets, 1966). The latter approach attempts to separate sensory factors from motivational factors in the analysis of psychophysical judgments.

METHOD

Subjects

Eleven White Carneaux pigeons, obtained from Palmetto Pigeon Plant, Sumter, S.C., ranged in age from 2 to 8 yr at the start of the experiment. Both cocks and hens were used. Their free-feeding weights were in the range of 500 to 600 g. During the experiment, the subjects were maintained at 75 to 80% of their free-feeding weights.³

¹The authors wish to express their gratitude to Willard Larkin for his many valuable suggestions and to Mrs. Jeanne Lonsbury for typing the manuscript. Reprints may be obtained from William Hodos, Department of Psychology, University of Maryland, College Park, Maryland 20742.

²In the strict photometric sense, the term "intensity" is incorrectly used in this paper because "intensity" applies to light from a point source. Since the stimuli in these experiments were area sources, the term "intensity" is inappropriate. However, we will follow Cornsweet's (1970) admittedly improper use of the term "intensity" to convey stimulus magnitude. Such terms as "luminous flux" or "luminance" would be more appropriate designations of stimulus magnitude in the case of an area source of light; however, since we have no information on whether the subjects were discriminating the total amount of light leaving the key (flux) or the amount of light per unit area of the key surface (luminance), a neutral term such as "intensity" seems preferable, in spite of the fact that it refers to a different type of light source.

³In conducting this research, the investigators adhered to the "Guide for Laboratory Animal Facilities and Care" as promulgated by the Committee on Revision of the Guide for Laboratory Animal Care of the Institute of Laboratory Animal Resources, National Academy of Sciences—National Research Council.

Apparatus

A three-key Lehigh Valley Pigeon Test Chamber was modified so that light from an externally located slide projector would illuminate the center key. The diameter of each key was 2.5 cm. The keys had been fabricated from Polacoat rear projection screen. The interior of the subject's chamber, which had been painted flat white, was illuminated by two ceiling-mounted General Electric F6T5-CW-HH fluorescent lamps. The light from the lamps was diffused by a sheet of white, translucent Plexiglas.

The overall chamber illumination was 247.6 lx. This was measured with an SEI photometer and a Leeds and Northrop calibrated test plate. The luminance of the walls varied from 57.0 to 59.5 cd/m². The ceiling luminance varied from 482 cd/m² directly below the lamps to 159.1 cd/m² in the darkest corner.

The stimulus system consisted of a Kodak Model 700 Carousel Projector that had been modified for remote operation and a series of Kodak No. 96 Wratten neutral density filters. The filters were calibrated for optical density with a Photovolt Model 520-M photometer. The following is a list of the nominal density values of the filters used; the values given in parentheses are the densities that were measured with the Photovolt photometer: 0.80 (0.805); 0.40 (0.440); 0.30 (0.325); 0.20 (0.210); 0.10 (0.105). In addition to the Kodak filters, a microscope slide cover glass was used as a nominal 0.05 neutral density filter. Its actual density was 0.053. The neutral density filters were inserted between sheets of Kodak B-351 Slide Cover Glass and mounted in 5 by 5 cm metal slide holders. The standard (0.0) was prepared as described for the variable series, except that no Wratten filters were inserted between the cover glasses. Two sets of filters were made, one for each of the two determinations of performance.

The light source for the center key was a Sylvania DEK 500 W projection lamp. The lamp was powered through a Sorensen Voltage Regulator and the lamp current was continuously maintained at 3.5 A. A 1.0 neutral density filter was permanently mounted in the path of the light beam. With no additional filters in the light path, the luminance of the center key was 318.5 cd/m². The side keys

were each illuminated by a Tung-Sol No. 1819 lamp powered from an unregulated 28 v dc power supply. The left side key luminance was 108.2 cd/m² and the right side key luminance was 135.9 cd/m².

Procedure

After training the pigeons to eat from the grain dispenser of the pigeon chamber, they were trained to peck the side keys when illuminated. The reinforcer was 3 to 5 sec of access to mixed grains in the grain dispenser. When this training was accomplished, the birds were trained to peck the center key when it was illuminated. A single peck on the center key darkened that key and illuminated the two side keys, which when pecked would produce access to the grain dispenser. The response requirement on the center key was gradually raised to 10 responses. When this preliminary training was completed, the pigeons were trained to peck the left-side key only when the 0.8 neutral density filter was in the chamber of the projector and thus attenuated the light on the center key by 0.8 logarithmic units. Right-side key pecking was reinforced only when the light on the center key was unattenuated. Thus, the basic procedure was a conditional discrimination: left-side key pecks produced grain when the center key had been "dim" and right-side key pecks produced grain when the center key had been "bright". In addition to the conditional discrimination procedure, a 50% reinforcement contingency was added to increase the number of trials per session. This contingency consisted of illumination of the light in the grain dispenser after each correct response, but only a random 50% of the correct responses were followed by the delivery of grain.

When the performance of the conditional discrimination exceeded 90% correct for three successive sessions of 240 trials each, the psychophysical training was begun. A psychophysical training session began with a 20 trial "warm-up" period. The procedure was the same as in the previous conditional discrimination sessions. The stimuli were 0.0 *vs* 0.8. The data of the warm-up period were discarded. Following the warm-up, a 20-trial "assessment" period was in effect, the purpose of which was to determine whether the subjects' performance level was high enough to war-

rant psychophysical testing or whether additional training on the basic 0.0 *vs* 0.8 problem was necessary. In the assessment period, the stimuli again were 0.0 *vs* 0.8. If the subjects' performance during this period was less than 90% correct (*i.e.*, more than two errors), the stimuli remained the same during the balance of the 240 trials of the session. Such a session was regarded as a training session. If two or fewer errors were made during the assessment (*i.e.*, 90% correct or better), then the psychophysical procedure was carried out and the data collected during the assessment period were used as the first determination of performance on 0.0 *vs* 0.8 for purposes of calculating a psychometric function. The assessment period was followed by five, 20-trial periods in which the stimuli were 0.0 *vs* 0.4, 0.3, 0.2, 0.1, and 0.05 respectively, in that order. Within each stimulus pair, the sequence of stimulus presentation was determined by a quasi-random discrimination sequence. After the 20 trials at 0.0 *vs* 0.05, another warm-up period (0.0 *vs* 0.8) was given, from which the data were discarded. Then, the entire procedure and stimulus sequence were repeated, except that less than 90% correct responses during the second determination of performance with 0.0 *vs* 0.8 did not influence the subsequent presentation of stimuli; *i.e.*, since the decision had already been made that a given session was to be a psychophysical testing session, no further assessment contingency was required in the second series of stimuli. To summarize the stimulus sequence, 14 stimulus pairs were presented during a testing session. These were 0.0 *vs* 0.8 (warm-up; data discarded), 0.0 *vs* 0.8 (assessment), 0.0 *vs* 0.4, 0.0 *vs* 0.3, 0.0 *vs* 0.2, 0.0 *vs* 0.1, 0.0 *vs* 0.05, 0.0 *vs* 0.8 (warm-up; data discarded), 0.0 *vs* 0.8, 0.0 *vs* 0.4, 0.0 *vs* 0.3, 0.0 *vs* 0.2, 0.0 *vs* 0.1, 0.0 *vs* 0.05. The warm-up period between the first and second descending series of stimuli served to dissipate extreme response biases that developed in some subjects during the later, more difficult discriminations of the first sequence and would otherwise be carried over into the earlier discriminations of the second sequence. The result of this mid-session warm-up was to make the data of the second stimulus sequence more consistent with those of the first.

The data collected during the first and second series of stimuli were combined to form

a daily psychometric function with per cent correct plotted as a function of the density difference between each stimulus pair. The "point of subjective equality" (PSE) was determined by a graphic procedure in which a perpendicular line was drawn from the point at which the psychometric function crossed 75% correct (halfway between chance and perfect discrimination) to an intensity value on the abscissa. The point on the psychometric function that intersected with the 75% correct line was taken to be the PSE. The difference between the standard (0.0) and the PSE was taken as the difference threshold. If the psychometric function crossed the 75% correct line more than once, the data of that session were discarded in order to avoid the interpretive ambiguities that such a situation creates. However, this occurred in less than 5% of the sessions and usually in the earlier sessions.

Data were collected for each subject until its performance satisfied the criterion of stability, which was less than $\pm 25\%$ deviation of the individual thresholds from the mean during five successive sessions. For example, if the mean threshold for five successive sessions were 0.10, then all of the individual thresholds would have to be in the range of 0.125-0.075 during those five sessions, in order for the stability criterion to be satisfied.

Pigeon C-430 posed special problems that required some departure from the procedure used with the other birds. This particular pigeon had the lowest thresholds of any of the birds studied. In a number of sessions, it performed at better than 75% correct on all stimulus pairs, including 0.0 *vs* 0.05. To determine whether this high performance represented high sensitivity or some artifact such as an extraneous cue or memorization of segments of the discrimination sequence, in several sessions the sixth stimulus pair in the sequence was 0.0 *vs* 0.0; *i.e.*, there was no stimulus difference. The bird's performance on the "no difference" problem ranged from 45 to 55% correct. The mean performance was 50%. Replacing the original 0.05 neutral density filters with new filters again resulted in performance of 75% correct or better. Since performance on "no difference" tests was chance, in those sessions in which performance on the 0.0 *vs* 0.05 problem was greater than 75% correct, the psychometric function

Fig. 1. Difference thresholds for each subject as a function of the number of sessions. The ordinate on the left of each graph expresses the threshold in terms of the optical density of the neutral density filters. The ordinate on the right of each graph expresses the threshold in terms of the percentage difference in luminance between the standard and variable stimuli. The shaded region of each graph indicates the five sessions in which the subject satisfied the criterion of stability. The dotted line indicates the mean threshold during the last five sessions. The height of the shading indicates the maximum permissible range of thresholds during those five sessions in order for the stability criterion to be satisfied.

was extrapolated to 0.0 (no difference) and a per cent correct value of 50% was assigned to that point. As may be seen in Figure 1, this extrapolation was necessary in only two of the last five sessions.

RESULTS

The development of stable intensity difference thresholds for the 11 pigeons is shown in Figure 1. For each pigeon, the difference

Table 1
Visual Intensity Difference Thresholds (Density) for Each Subject

Subject	Mean Difference Threshold of the Last Five Sessions	Difference Threshold Range During the Last Five Sessions	Difference Threshold Derived from Mean Psychometric Function of Last Five Sessions
C-430	0.05	0.045-0.050	0.05
C-429	0.08	0.07 -0.10	0.07
C-431	0.10	0.10 -0.11	0.08
C-180	0.11	0.08 -0.13	0.09
C-139	0.08	0.07 -0.10	0.09
C-158	0.09	0.08 -0.10	0.09
C-340	0.10	0.08 -0.12	0.10
C-344	0.16	0.11 -0.19	0.15
C-096	0.20	0.18 -0.25	0.19
C-339	0.20	0.15 -0.25	0.19
C-322	0.20	0.15 -0.24	0.20
Grand Mean = 0.125		Grand Mean = 0.118	

threshold is plotted as a function of the number of sessions. The right tail of each curve is enclosed in a shaded region. This region indicates the five successive sessions in which the pigeons satisfied the criterion of stability. The broken line through the shaded region represents the mean threshold for the five sessions and the vertical extent of the shaded region indicates the limits of the $\pm 25\%$ range. At the right side of each graph, the optical density differences have been converted to per cent luminance difference. As may be seen in the figure and also in Table 1, the range of intensity difference thresholds was 0.25 to 0.045 density. In terms of per cent luminance differences, the range is from approximately 40% to approximately 10%. The mean difference threshold (density) for all subjects was 0.125 or approximately 25% difference in luminance.

Figure 2 depicts the mean psychometric function for each subject. This function represents all data collected during the five stability sessions. For each subject, the mean per cent correct has been plotted as a function of the density difference in the stimulus pairs. On each graph, the horizontal broken line indicates the locus of 75% correct. The vertical broken line indicates the density difference associated with 75% correct on the psychometric function. Each point on the function represents the mean of 200 trials. As seen in Table 1, the difference thresholds calculated from the mean psychometric functions in nearly half the cases were slightly smaller

than the difference thresholds calculated from the mean of the individual difference thresholds. The remaining cases were identical. The differences between these methods of calculating almost certainly represent artifacts of the averaging process. Since either method seems to be an appropriate way to calculate the mean threshold for individual subjects, the table presents both sets of data so that the reader may take his or her choice. The grand mean based on the mean psychometric functions was 0.118. The difference between the two methods of calculating the grand mean is therefore 0.007 in optical density or about 2% in luminance.

In addition to the determination of psychometric functions, a signal-detection analysis of the data was also carried out. Since we had no basis for making many of the assumptions about the statistical nature of the distributions of events in the nervous system that are the substrate of sensation (Green and Swets, 1966), we used two non-parametric indicators that make no such assumptions. These indicators are based solely on the geometry of the unit square of signal-detection theory. The first of these (Pollack and Norman, 1964), defines a family of isosensitivity contours (receiver operating characteristics or ROC curves) in which the parameter A' represents the sensitivity of the subject. A' may vary from 0.50 (no detection of the stimulus or stimulus difference) to 1.00 (perfect detection). The second method (Hodos, 1970) defines a family

Fig. 2. Psychometric functions for each subject based on the mean performance during the last five (stable) sessions. The horizontal broken line represents 75% correct. The intersection of the 75% correct line with the psychometric function indicates the point of subjective equality. The difference between the point of subjective equality and the standard (zero) is the mean difference threshold.

of isobias contours that represent the subject's criterion or bias towards one or the other response alternatives. The bias index ranges from +100% (complete bias towards one response alternative; very high criterion; very conservative observer) through zero bias (no bias towards either alternative; unbiased observer) to -100% bias (complete bias towards the other alternative; very low criterion; very liberal observer).

In Figure 3, the data from each subject have been plotted in the form of the conditional probability of a correct positive response *versus* the conditional probability of an incorrect positive response (Green and Swets, 1966). Each point within each unit square represents the mean performance on a particular stimulus pair during the five stable sessions. Thus, each point represents 200 trials at a particular intensity difference. The key within the figure indicates the density difference that each point represents. In all of the unit squares shown in Figure 3, the points representing the largest density differences

are clustered in the upper left corner, indicating that these stimuli were highly detectable. The points representing the smaller density differences progressively approach the main diagonal, which indicates that these stimuli were more poorly detectable. In each case, the ordinal relationship between the index of detectability (A') and the magnitude of the intensity difference is preserved. The Spearman rank difference correlation (ρ) ranged from +0.99 to +1.00, which is highly significant.

The response bias analysis of the data in Figure 3 indicates that except for the smallest intensity difference, most of the subjects maintained a fairly consistent response criterion. In almost every case in which the bias remained fairly constant (isobias) over the range of stimulus intensity differences, an abrupt shift in response bias may be seen at the smallest stimulus difference (0.05). Some of the birds maintained a moderate positive bias (e.g., C-096 and C-139) which indicates a moderate bias to peck the side key that pro-

Fig. 3. A signal-detection analysis of the data. The curved lines arching over the main diagonal of each unit square represent non-parametric isosensitivity curves (operating characteristics) (Pollack and Norman, 1964). The numbers on each curve are the index of detectability (A') in which 1.0 = perfect detection; 0.50 = no detection. The curved lines originating in the upper-left square and ending in the lower-left and upper-right corners of the square are non-parametric isobias curves (Hodos, 1970). The numbers on each curve indicate the percentage of bias. Positive bias scores indicate a bias toward the standard intensity response. Negative bias scores indicate a bias toward the variable intensity response. The symbols within each square represent the conditional probability of an incorrect positive response for each of the six stimulus comparisons used.

duced grain in the presence of the standard (0.0 density) stimulus. Several others, such as C-430 and C-431, showed no consistent response bias towards either key. Several other cases, notably C-344, C-322, and C-339 showed a systematic change in bias throughout the entire series of intensity differences. These three birds showed a progressively increasing tendency to peck the left key, which produced grain in the presence of the variable (attenuated) stimulus.

DISCUSSION

In general, the results indicate an orderly relationship between the magnitude of the stimulus difference and the empirical proba-

bility of a correct response. The difference threshold for the average subject was approximately 0.12 density, which represents a luminance difference of about 25% or about 80 cd/m^2 . In the best case, the difference threshold was 0.05 density or about 11% luminance difference. Due to the relationship between stimulus energy and differential sensitivity as expressed by Weber's Law, the difference thresholds reported here are meaningful only for intensity stimuli in the vicinity of 318 cd/m^2 . For less intense stimuli, according to Weber's Law, the difference threshold should be proportionately smaller. To compare the present data with previous investigations, the Weber fraction was calculated according to the following formula:

$$\frac{\Delta I}{I} = k \quad (1)$$

in which ΔI = the difference threshold, I = the intensity of the standard and k = the constant of proportionality. For the average subject reported here, $k = 0.23$. In the case of C-430, the best subject, $k = 0.11$. Mentzer (1966), using an intensity of 171 cd/m², reported values of k in the range of 3.69 to 0.20. The mean value of k for his subjects was approximately 0.54. Data reported by Shumake, Hatfield, and Smith (1966) indicated that the Weber fraction was constant at about 0.30 in the range of 5-50 cd/m².

The present data on sensitivity to intensity differences are relevant to research in the area of stimulus control and attention. In such experiments, an important aspect is to relate an animal's performance to a particular feature of the stimulus and not to other features. Frequently employed stimulus qualities are color differences, pattern differences, or different orientations of the same pattern. Often, the control of intensity differences is casual or inferred, as in the case of assuming that the intensity will be the same when the same pattern is presented in a different orientation. However, unless the stimuli are measured with a reliable photometer and some standards of acceptability established, a disconcertingly large intensity difference can easily be introduced as a confounding variable. For example, the lamp intensities in the in-line digital display projectors that are frequently used as a stimulus source in operant studies, often vary in intensity by as much as 0.7 to 1.0 logarithmic units, depending on filament resistance, filament orientation, and hours of operation. In addition, the optical density of the transparencies containing the images to be projected may vary. Thus, a horizontal bar and a vertical bar in such a unit, derived as they are from different light sources and different transparencies may be highly distinguishable on a basis of intensity as well as orientation. Such a confounding variable could produce serious interpretive difficulties. The present data indicate that most of the subjects could easily discriminate a difference as small as 0.2 logarithmic units, with successive comparisons. This difference might not be readily apparent to human observers, inexperienced in such psychophysical

judgments and making the comparison on a successive basis. Our recommendation would be that experimenters interested in controlling intensity differences between stimuli should match their stimuli to within 0.05 logarithmic units. This matching should be done by means of initial selection of lamps according to intensity and finer matching by neutral density filters, rather than by gross adjustment of lamp voltage, which could introduce unwanted color differences.

The signal-detection analysis of the data presented in Figure 3 indicates that in each subject's data, the ordinal relationship between the stimulus energy difference and the index of detectability (A') was maintained. A finding that is of importance for psychophysical theory is the fact that this ordinal relationship was preserved even below "threshold." The PSE would be represented in the unit squares in Figure 3 by a line running parallel to the main diagonal from 0.5 on the ordinate on the left to 0.5 on the abscissa at the top. This line would represent the locus of all combinations of correct positives and incorrect positives that would result in a total percentage correct of 75. The observation that the detectabilities of the stimulus differences are correctly ordered below this line supports the "no threshold" interpretation of psychophysical data (Swets, 1961; Green and Swets, 1966). If the intensity difference associated with performance that is halfway between chance and perfect detection represents a sharp cut-off below which detection is not possible, then one would expect to find a random relationship between the stimulus energy difference and the index of detectability. Since the relationship is not random, the "no threshold" position is supported. This observation has implications for the interpretation of animal psychophysical data of various sorts. Although such terms as "difference threshold" and "JND" have a certain amount of practical utility, one should recognize that these terms may have little relevance to the underlying psychophysical mechanism.

REFERENCES

- Cornsweet, T. N. *Visual perception*. New York: Academic Press, 1970.
- Gilbert, R. M. and Sutherland, N. S. (Eds.), *Animal discrimination learning*. New York: Academic Press, 1969.

- Green, D. M. and Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- Hodos, W. A non-parametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 1970, **74**, 351-354.
- Mentzer, T. L. Comparison of three methods for obtaining psychophysical thresholds from the pigeon. *Journal of Comparative and Physiological Psychology*, 1966, **61**, 96-101.
- Pollack, I. and Norman, D. A. A non-parametric analysis of recognition experiments. *Psychonomic Science*, 1964, **1**, 125-126.
- Shumake, S., Hatfield, C. A., and Smith, J. C. Brightness difference thresholds in the pigeon using the conditioned suppression technique. *Psychonomic Science*, 1966, **6**, 313-314.
- Swets, J. A. Is there a sensory threshold? *Science*, 1961, **134**, 168-177.
- Swets, J. A. (Ed.), *Signal detection and recognition in human observers*. New York: Wiley, 1964.
- Terrace, H. S. Stimulus control. In, W. K. Honig (Ed.), *Operant behavior: areas of research and application*. New York: Appleton-Century-Crofts, 1966. Pp. 271-334.

Received 30 September 1971.

(Final acceptance 24 April 1972.)