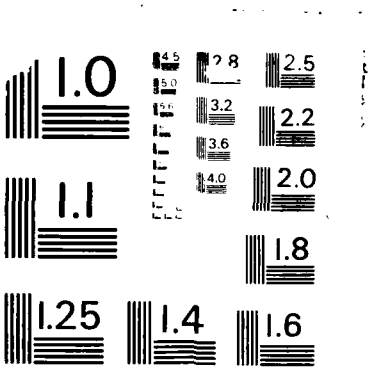
AD-A189 704 ON CATEGORIZING SOUNDS(U) DUKE UNIV DURHAM NC DEPT OF PSYCHOLOGY G R LOCKHEAD 17 NOV 87 AFOSR-TR-87-1912 AFOSR-85-0302 1/1

UNCLASSIFIED F/G 5/8 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963

T DOCUMENTATION PAGE

(2)

AD-A189 784

1a REPORT UN	1b RESTRICTIVE MARKINGS
2a SECURITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-1912
6a NAME OF PERFORMING ORGANIZATION Duke University	6b OFFICE SYMBOL (If applicable)
6c ADDRESS (City, State, and ZIP Code) Durham, North Carolina	7a NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research/NL
	7b ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332-6448
8a NAME OF FUNDING / SPONSORING ORGANIZATION AFOSR	8b OFFICE SYMBOL (If applicable) NL
	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-85-0302
8c ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61102F	2313	A6	

DTIC ELECTE JAN 05 1988 S D

11 TITLE (Include Security Classification)

On Categorizing Sounds

12 PERSONAL AUTHOR(S)
Gregory R. Lockhead

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM 1Aug85 TO 31Jul87	17 Nov 87	16

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19 ABSTRACT

Judgments of sounds depend on context. How a sound is labeled depends on the sounds that just occurred (sequence effects) and the sounds that might occur (set effects or range effects). These dependencies are sufficiently large that they sometimes predict performance better than the stimulus itself. This report summarizes studies of context conducted during two years of AFOSR support. These studies of sound classification evaluated features of a memory model constructed to account for univariate judgments. The data show how response variability depends on stimulus variability, and demonstrate the importance of experimental details such as whether feedback is given and whether an identification function is present. It is concluded that three variables are needed to describe the collection of results. These are the stimulus itself, the stimulus or response (depending on feedback) on the just prior trial, and an average (called a memory pool) of the stimuli on each of several earlier trials.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
John F. Tangney	(202) 767-5021	NL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Duke University
DURHAM
NORTH CAROLINA

DEPARTMENT OF PSYCHOLOGY

POSTAL CODE 27706
TELEPHONE 919—684-4321

November 17, 1987

Dr. John F. Tangney
Program Manager
Life Sciences Directorate
Air Force of Scientific Research
Bolling Air Force Base,
D.C. 20332-6448

Re: AFOSR-85-0302

Dear John:

I have enclosed a report on the progress made during the first two years of AFOSR support. I had not understood that a technical report was needed rather than a brief summary.

Please let me know if anything else will be needed. I am looking forward to next month and meeting the other people supported by your office.

Accession For
NTIS GRA&I [X]
DTIC TAB [ ]
Unannounced [ ]
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

Sincerely yours,

Gregory R. Lockhead
Professor

2 NOV 1987

AFOSR-TR- 87-1912

A Technical Report. Progress on AFOSR Research Grant No. AFOSR-85-0302.

On Categorizing Sounds

Gregory R. Lockhead

Duke University

Durham, NC, 27706

November, 1987

On Categorizing Sounds

Gregory R. Lockhead
Duke University
Durham, NC, 27706

Abstract

Judgments of sounds depend on context. How a sound is labeled depends on the sounds that just occurred (sequence effects) and the sounds that might occur (set effects or range effects). These dependencies are sufficiently large that they sometimes predict performance better than the stimulus itself. This report summarizes studies of context conducted during two years of AFOSR support. These studies of sound classification evaluated features of a memory model constructed to account for univariate judgments. The data show how response variability depends on stimulus variability, and demonstrate the importance of experimental details such as whether feedback is given and whether an identification function is present. It is concluded that three variables are needed to describe the collection of results. These are the stimulus itself, the stimulus or response (depending on feedback) on the just prior trial, and an average (called a memory pool) of the stimuli on each of several earlier trials. While a prediction equation thus must be relatively complex, the tasks do not require exceptionally complex cognitive processing by the subjects. The data from pigeons and people are essentially identical. For both classes of subjects, the position of the criterion for judgment as defined in statistical decision theory is not fixed during a condition. Rather, its position from trial to trial depends on sequence. It is concluded that judgment is a dynamic process that a successful prediction model must be based on a trial-to-trial analyses of data. Models based only on data averaged across sequences cannot account for a large proportion of the variability in the responses. Work is in progress to examine if these descriptions of performance in univariate tasks generalize to multidimensional situations.

On Categorizing Sounds

This research project began with the general goal of better understanding how complex sounds are identified. The specific and more immediate goals are to evaluate a model designed to account for sequence effects in judgment, to examine context effects when multidimensional sounds are identified, and to examine ways to improve judgment accuracy after the fact by adjusting responses according to what is known about context effects. This report summarizes the approach and some of the work completed during the first two years of AFOSR support.

Judgments of stimuli are markedly affected by context (what stimuli might occur) and by sequence (what stimuli recently did occur) in all data sets that have been examined. This means that at least some of the variability in judgments is not random error. Models of the form

$$R = f(S) \qquad (1)$$

where R is a response and S a stimulus, average across many of these effects and thus cannot account for this variability. Such models often describe averaged data very well but cannot predict individual responses with the precision possible using models that consider momentary or trial-by-trial behaviors.

A model commonly used to describe magnitude estimation (ME) data provides an example. Averaged MEs are well described by Stevens' power law,

$$R = kI^b \qquad (2)$$

where R is the mean magnitude estimation, I is stimulus intensity, and k and b are constants. However, ME data are often so variable that the response on any particular trial is poorly predicted by Equation 2. Sequential constraints sometimes bias the response by more than an order of magnitude (3).

This does not mean such models are wrong. They might only be incomplete. For example, one can maintain the spirit of Stevens' search for an underlying invariance (Equation 2) and also describe sequence effects. One such attempt is:

$$R_N = (I_N)^b + \underline{a}[(I_{N-1})^b - (I_N)^b] \qquad (3)$$

where $\underline{a}$ is a positive constant. According to this description, if the previous stimulus was larger than the current one R_N is biased upward, and if the previous stimulus was smaller the response is biased downward.

Equation 3 predicts individual judgments in ME tasks very much better than does Equation 2 which does not account for any of the trial-by-trial structure in ME and other data (5).

A different equation considers effects of sequence farther back than the one trial addressed in Equation 3. This description assigns adaptive-like changes to the memory scale and describes performance slightly better than do any of Equations 1-3:

$$R_n = S_n + \underline{a}(M_{n-1} - S_n) + \underline{b}(M - M_p) + e \qquad (4)$$

where M_{n-1} is the remembered value of the stimulus on the just previous trial, M is a constant representing the average of all memories over the course of the experiment, M_p is the average memory of several prior stimuli and called the memory pool, a and b are positive constants, and e is an error term (see 5, Eq. 8).

The studies summarized ahead were conducted to further examine how responses to a stimulus depend on context, on sequence, and on particulars of the experimental procedure.

Stimulus distribution affects response variability.

Stimulus Range. The variability of responses to a stimulus is larger when stimuli in the set differ more from one another. This is often called a range effect. A metaphor from optics describes the result and the assumption. In order for an optical lens to provide a wide field of view, its magnification power and thus its resolution must be low. If the range to be viewed is made more narrow, then a higher magnification can be used. This results in increased resolution.

Precision (response variability) and field of view (stimulus range) are inversely related in optics and in many other physical systems. Judgment data are analogous to this. If the physical range over which stimuli vary is small, response resolution is great compared to when the physical range is large (6 and references there).

Stimulus Sequence. A different interpretation than range of why response variability correlates with stimulus variability is proposed here. This concerns sequence. The suggestion is that responses are more variable on trials that successive stimuli are more different from one another.

This would occur if subjects judge a stimulus by estimating the distance between it and the previous stimulus, and if any error in these estimations is magnified by the distance involved. Then, response variability would be small when successive stimuli are very similar and would be large when successive stimuli are very different.

Stimulus range or stimulus sequence? Both above interpretations are consistent with the observed effects of range on response variability. Range models state the effects directly and are essentially restatements of the empirical observation. Sequence models predict them indirectly since it is only on average that successive stimulus differences are larger when the stimulus range is larger.

The studies summarized ahead were conducted to determine if response variability in judgment data is due to stimulus range or to successive stimulus differences. These studies were based in four observations:

(a) When stimulus range is large in absolute identification (AI) conditions, response variability is also large (1). This is a between-conditions effect of range.

(b) There is assimilation in all judgment data that have been examined for sequence effects. From trial-to-trial, the mean response to a stimulus tends toward the value of the stimulus or response (depending on particulars of the task; cf. 5) on the prior trial.

(c) For any particular stimulus range condition, the

magnitude of assimilation (the amount by which the mean response shifts toward the prior trial) tends to be greater on trials that successive stimuli are more different. This is a within-conditions range effect.

(d) Stimuli were selected randomly in these studies reporting assimilation. Thus, successive stimuli were more different on average in large stimulus range conditions.

These observations allow the suggestion that the response variability correlated with stimulus range is due, instead of to range, to stimulus sequence. To test this inference, identifications in an AI task were examined for individual stimuli as a function of what stimulus occurred on the prior trial, in conditions where the stimulus range was narrow and conditions where it was relatively wide (16).

One result is was that the mean variability in the responses to any particular stimulus was greater in conditions having greater stimulus range. This replicates the earlier similar finding between-conditions. Another result was that much of the variability associated with range is also correlated with sequence. In both range conditions the responses to any particular stimulus were more variable when just previous stimulus was more different from it.

These results occurred whether humans (16) or pigeons (12) were the subjects, and whether the stimulus range was small or large. This agrees with the view that response variability is associated with successive stimulus differences (sequence). Apparently response variability is associated with stimulus range only to the extend that changes in range produce changes in sequence.

Consistent with Equations 3 and 4, the identification of simple sounds is a complex and dynamic task. Subjects attempt to remember both recent and earlier events in order to maintain a veridical mapping between the response scale and their memories of the stimulus scale, and those memories are affected by sequence.

If these effects due to sequences only produced minor effects they might be ignored, at least for practical purposes. However, the effects are not minor. They often result in a response shift of half or more of the response scale.

An example of the ways responses depend on successive stimuli is useful. This example is from a tone identification task (16, Experiment 1). There were three tone intensities in a condition. Two of the conditions were a narrow range and a spread-high range. In the narrow range condition, the three 1,000 Hz sinewaves were 58, 60, and 62 dB intense. In the spread-high range condition, these intensities were 58, 60, and 66 dB.

Successive stimuli were selected randomly for each trial in each separately conducted condition. The subjects were asked to identify each stimulus by pressing the appropriate one of three buttons. There were 250 trials per subject for each condition. Because the conclusion was the same for all comparisons, only one is described here.

Two of the stimuli, the 58 and 60 dB tones, were identical in these two conditions. Performance in identifying these two tones was measured in each condition with an analysis based on statistical decision theory (SDT). Four results are of interest:

(a) The 58 and 60 stimuli were better identified (larger d') in the narrow range (containing a 62 dB third tone) than the spread-high (containing a 66 dB third tone) condition. Thus, the wider stimulus range resulted in increased response variability to the two unchanged stimuli.

(b) Successive responses were positively correlated. This shift in the average response toward the previous trial is assimilation.

(c) The magnitude of assimilation was greater on trials that the physical difference between successive stimuli was greater.

(d) The variability of responses to each stimulus was greater on trials that successive stimuli were more different.

In summary: the mean response and the variability of responses to a stimulus were greater on trials that successive stimuli were more different.

Concerning method. There is an implication in these results concerning the use of statistical decision theory. Ordinarily, variability (noise) in SDT studies is assigned to the signal or system, not to the criterion which is assumed to be fixed within conditions.

The assumption of a fixed criterion is not a requirement of SDT and is not consistent with these data. Here, the position of the criterion changes from trial-to-trial during the study (cf. Figures 3 and 5 in 16). Its position on any particular trial depends on stimulus sequence. The support for this conclusion is that hit-rates and false-alarm-rates increased or decreased together within a condition, depending on the particular sequence.

Since the position of the criterion varies from trial to trial, d's calculated from averaged data (sequence ignored) must be smaller than d's calculated separately for each stimulus sequence. This is because averaging treats sequential dependencies as noise, which they are not, resulting in a deflated estimate of discriminability.

It is not necessary in statistical decision theory to assume the noise is in the system and the criterion is invariant during an experimental condition. It is equally correct to assume, instead, that the noise is in the criterion and the system and signal are error free. Assigning the noise to the system is only a convention and SDT conclusions are identical under either assumption. Too, it is not novel to suggest the criterion is noisy. Green (2) concluded that response biases and internal variability both affect performance in a proper SDT experiment, i.e., with one signal to be detected rather than three to be identified as here.

Since the criterion is variable from trial to trial, d's measured for each sequence provide, compared to averaging across trials, a more precise measure of discriminability and a more complete description of the data. Specifically, for the study just summarized consider responses to the 58 dB tone. They were shifted most on average and were most variable on trials following the 66 dB tone (wide range condition), and these effects were regularly reduced for successively quieter prior tones.

These results are not because of only 3 stimuli in the

study. They also occur using 10 stimuli (6). In general, the mean response to a stimulus is biased toward the previous trial, and the variability of the responses to a stimulus is greater when the successive stimulus difference is greater.

Accordingly, when discriminability is calculated based on averaged data, that estimate will be smaller than the average of the estimates calculated for each sequence. Which analysis is correct may be arguable and may depend on the purposes of the study, but it is always an advantage to remove variability from data.

The presence of an identification function affects performance.

When there are sequence effects in identification tasks there must also be errors. What particular errors occur depends on specifics of the task. The concern here is with how errors relate to an identification function (IDF).

An IDF is present whenever the experimenter has specified a label for each stimulus and the subject is to assign each the correct label.

Absolute identification (AI) is such a task. If the experimenter does not assign labels and the subject is allowed to believe there are many, many stimuli, there is assumed to be no IDF. Magnitude estimation is such a task.

It is known that the mean response to a stimulus depends on sequence and on the presence or absence of an IDF (6). This study compared the variability of responses to each stimulus when there was and was not an IDF (17).

Consider an absolute identification study in which the experimenter provides the subject 10 response buttons and feedback in the form of the numerals 1 - 10 after each response. The subject will appropriately assume there are 10 stimuli and nothing smaller than value 1 nor larger than value 10 will occur. Thus, there is an IDF.

The assumption is different in magnitude estimation (ME) tasks. Subjects are neither told nor led to assume that only a few stimuli (commonly six) will be presented and they have no reason to constrain their responses to any range. There is no IDF.

It has frequently been noted (10) that AI data are concave down in comparison to ME data. The reason is not known. This study examined if the reason is because there is an IDF in AI tasks but not in ME tasks.

To explain the idea it is necessary to note there is assimilation in both AI and ME data. The response to a stimulus tends to be small when the prior stimulus was small, and large when the prior stimulus was large. However, in AI tasks (IDF present) but not in ME tasks (no IDF), manifest assimilation to stimuli at the ends of the range in constrained. As described next, this accounts for the one data set being concave with respect to the other. It also accounts for observed variability differences in the two classes of data.

Consider this example situation in an AI task with subjects provided the responses 1 - 10. The last stimulus was called #1 and the next stimulus is judged as less than the memory of this #1. Due to assimilation, this is not a rare event (5). Since

there is an identification function a response less than 1 is not available. Hence, #1 is given.

The identical situation occurs with intense stimuli. Whenever a stimulus is called #10 and the next stimulus is judged larger than the memory of this one, this perception cannot be reported and the stimulus is labeled #10.

There is no such constraint with intermediate stimuli in AI tasks. If the last stimulus was called #5, whether the next one is judged as smaller or larger than the memory of this #5, many acceptable responses are available.

This analysis indicates that the mean responses to end stimuli tend toward the middle of the scale and the variability of those responses (but not necessarily of the perceptions) is less than that to central stimuli.

The situation is different when there is no IDF, as in magnitude estimation (ME) tasks. There is then no restriction on responses. If a stimulus is judged as greater or less than the prior stimulus, a response is always available to reflect this. Therefore, there is no reason for response variability to depend on scale position.

If this interpretation of why the form of AI data differs from that of ME data is correct, then there are two summary conclusions. (1) Assimilation in AI affects intermediate responses symmetrically but end responses asymmetrically. Since there is no such tendency in ME data, mean AI responses are concave with respect to ME responses. (2) Responses to end stimuli should be less variable than those to intermediate stimuli in AI data but, again, no such difference is predicted for ME data.

To examine this inference, MEs were collected when feedback was and was not provided after each response. The stimuli were 30 1,000 Hz tones that varied in loudness in 1 dB intervals. The no feedback condition was a standard ME task with free modulus and no explicit IDF. The feedback condition was the identical task and stimulus sequence but with feedback given after each response. This feedback was the average response that had been given by other subjects in the ME task.

The results are clear. The mean and variability of responses across stimuli were different in the two conditions. Particularly, response variability was less at the extremes than at the center of the scale (an inverted U) in the feedback data but no such effect is suggested in the ME data.

A prior attempt to account for the "U" has taken this non-linear relation between response variability and stimulus intensity as evidence that subjects use the ends of the stimulus range as subjective standards for judgments, and concludes that performance is less variable when stimuli are more similar to a standard than when stimuli are far from a standard.

Another suggestion that has been given to account for the same result is the exact opposite. This assumes that intermediate stimuli serve as the standard (some authors) or as the adaptation level (other authors) and performance on the end stimuli is less variable because, being far from the standard, the end stimuli are more distinctive.

In the study reviewed above (17), the stimuli were identical in both tasks but the response means and variabilities over

stimuli were different. Therefore, arguments based on intensities like those just summarized must predict the same function in the two data sets. Since the functions are different, such intensity-based interpretations are not sufficient.

One obvious difference between the two conditions is whether feedback is given. Thus, feedback might considered as the reason for the differences in the data. However, the "U" occurs in AI tasks whether or not feedback is given. Thus, feedback is not the source of the difference.

The only suggested interpretation consistent with all of the data is the importance of identification function. When an IDF is present, responses cannot be smaller than the smallest assigned number or larger than the largest assigned number. This constrains the mean and variability of responses to those stimuli. Here, if the current stimulus is judged as louder than the one just called #10, the subject can do more than hit key 10, hard, which often happens.

There are no such conflicts in an ME task. The second #10 is simply assigned a larger response than was the first #10 and the subject has no awareness of giving different responses to identical stimuli. This happens often (5, Fig. 3).

Conclusion Sequence effects imply variability in the responses to individual stimuli. When an IDF is present, this variability is less to stimuli at the ends of the scale than to intermediate stimuli. In disagreement with models of reference points, this does not necessarily mean that extreme stimuli are better identified than internal stimuli. Rather, some errors simply cannot be observed when the response scale is restricted.

Concerning sources of sequence effects.

Sequence effects (SEs) might be associated with stimulus intensities, sensory processes, encoding processes, perceptual mechanisms, memory systems, or response processes. Each possibility has been proposed in the literature and the answer is not known.

If SEs are due to a single mechanism, it is not stimulus intensity, a sensory response, encoding processes, or a perceptual mechanism. This is because the same SEs occur when there are no intensities to respond to. In a demonstration of this, the signal generator was turned off in an AI experiment and subjects were asked to guess the stimulus selected randomly for each trial. Feedback was given. Each guess was overly similar to the prior feedback (6). This is assimilation. Since assimilation occurs whether or not there are stimuli, it cannot be due to just stimuli or stimulus effects.

The neural attention-band and assimilation in memory. Although sequence effects do not require intensities, it is logically possible they are due to intensities when intensities are present. A neural account that makes this assumption has been offered (7, 8). That model was used to account for the observations that successive ME responses are positively and highly correlated when successive stimuli are similar, and that the magnitudes of these correlations decrease with increases in the difference between successive stimuli.

The auditory attention-band model equates this result with a differential ability of the system to sample incoming neural fibers. According to that model, aspects of two signals that fall within the hypothesized band are less variably estimated than are those same aspects when they fall outside the neural band. This is because when successive signals are within the band observers are presumed able to preserve the ratio of the successive intensities in their responses. Thus, the responses are correlated. If one of the signals is outside the band, as when the successive stimuli are very different, the sample is then not sufficient to allow the ratio of the intensities to be incorporated in the response, eliminating the correlation. Separation in dB is involved in the predictions because the attention band is assumed to be narrow and centered on the past stimulus intensity.

The attractions of this attention-band hypothesis are several. A neural mechanism is implicated, the idea is easily communicated, and the theory describes a complex data set.

There are also reasons not to be attracted to this hypothesis. One already observed is there are SEs even when there are no intensities to involve a neural band. Another, described ahead, is that while the attention-band describes one aspect of the data it cannot account for other aspects.

The attention-band model is based on positive correlations between successive responses when successive stimuli are similar. This is assimilation. The suggestion here is that this sequence effect is associated with response processes rather than a neural attention band. To test this, two people who knew the IDF and two others who did not gave MEs to 30 tones. The results and implications of this pilot study are described below.

Standard Magnitude Estimation. This was a standard ME study. The stimuli were 30 1,000 Hz sinewaves that differed in 1 dB steps from 51 to 80 dBA. The tones had a 15 ms. rise time and were turned off abruptly after 500 ms.

Using the analysis method reported by Luce and colleagues, correlations between successive responses were calculated as a function of the difference between successive stimuli. These correlations are high and positive (about 0.8) for the most similar successive stimuli and become smaller as the intensity difference between successive stimuli increases. This replicates the results reported by Luce, Green, and Weber.

Another feature in these data is that these positive correlations do not simply drop to zero as the successive stimuli become more different. Rather, they pass smoothly through zero and steadily become more negative (to about $r = -0.7$) when successive stimuli are very different.

Negative correlations are also seen in published figures although the magnitudes there are smaller. This is probably because the published data are averaged over groups of stimulus differences. Since there are relatively few occasions of large stimulus differences (in a 30 stimulus study, there are 29 occasions of a 1 step difference for every 1 occasion of a 29 step difference) such averaging may be appropriate to provide stability in the results. However, such averaging masks the individual correlations for extreme stimulus differences, which are large and negative.

The observed positive correlations mean that each response tends toward the previous response. If this assimilation were the only process involved, this response compression would continue until every stimulus is assigned the same response.

This does not happen. Instead, the response compression becomes noticeable to the subjects, particularly when the successive stimuli are very different, and they take action to correct the response scale. The action needed is to expand the response range. If such -corrections occur primarily on trials that successive stimuli are very different (3), successive responses on those trials would be negatively correlated, which they are.

Magnitude Estimations with feedback. To examine the suggestion that sequential effects reflect both assimilation and subsequent corrections to maintain a veridical response scale, MEs were collected with feedback given after each response. Since feedback provides for response correction on each trial, there should be relatively little cumulated compression of the response scale to be corrected. If this is the case, then the negative correlations observed at large stimulus differences in ordinary ME data might not occur when feedback is given.

Subjects gave MEs to 30 tones. Feedback was given after each response. This feedback was the average response to each stimulus that had been given by subjects in the ME study just described.

To analyze the results, the average correlations between successive responses were calculated as a function of the difference in dB between successive stimuli. There is no consistent pattern in the data. Thirty-eight of the 60 correlations are positive; the largest of these positive correlations were scattered at -29, -1, 11 and 25 dB differences. None of the correlations approach the magnitudes observed in the ME data, +0.8 and -0.7.

While the overall average correlation in these feedback data is positive (assimilation), there is no observed tendency for successive responses to be more positively correlated when successive stimulus differences are smaller. This is in marked contrast with the findings in ME data.

Also, there is no trend toward negative correlations at large stimulus differences in these feedback data. This is again in marked contrast with the ME data. The distribution of correlations over stimulus differences is generally positive but otherwise erratic.

Memory or response processes. The attention-band model is based on neural activities in the ear. Since the stimuli and the stimulus sequences were identical the ME and ME-with-feedback studies, sensory stimulations were the same. Thus, the attention-band model as presented cannot predict different correlations in the two data sets. Since the correlations are different, the attention-band model is not supported.

This collection of results is consistent with a model in which successive events assimilate in memory and subjects attempt, nonetheless, to maintain a reliable mapping of stimuli onto their response scale. In its development, Equation 4 was expressly different for feedback than no-feedback data. The emphasis of the model is that there is always assimilation and

subjects use whatever is available to help them maintain a veridical response scale.

According to this latter view, feedback is a reliable indicator of the value of the stimulus just judged and subjects should regularly adjust the response scale in terms of this. If they do, then successive responses should not correlate. They do not.

It is concluded that the correlations between successive responses in ME are not due to the sensory system. There is assimilation with or without stimulus involvement. And when there are stimuli, sequential contingencies are associated with the prior responses or the prior feedback or whatever other information is available to help the observers maintain a reliable mapping of the stimulus and response scales.

Briefly noted.

What determines that something will be classified as a chair or a cup or loudness #5 is not fully understood. This is partly because the stimuli examined are usually complex and, probably, because classification processes are also complex. To simplify the category problem, we have been examining how people classify very simple things. The goal is to uncover classification rules, if there are, that might also apply when more complex stimuli or tasks are involved.

This work is based in the observation that psychophysical judgments are usually more variable when stimuli in the set are more variable. For example, increasing the physical range over which stimuli vary in an identification task results in increased variability in the responses (1). This is for when the experimenter determines what stimuli belong to which category.

The studies summarized here examined what occurs when the subject rather than the experimenter decides category memberships. We asked people to categorize some simple stimuli. The simplest situation we have thought of is to give people two dots on an otherwise blank piece of paper.

Using the method of adjustments, we asked people to place a third dot between these two stimulus dots, such that the two dots on the left were just grouped, leaving the dot on the right by itself. The average of these dot positions was 64% of the distance between the stimulus dots, with the response placed closer to the left than the right stimulus element. The standard deviation of these 25 judgments is 4.6%.

Using the same stimulus dots, we asked another 25 people to place a dot that was just not grouped with the dot on the left. The average response dot was again left of halfway between the stimulus dots. This mean placement is 57% of the distance between the dots. The standard deviation is 4.4%.

The mean of these just and just not grouped values is taken as an estimate of a subjective boundary separating the three dots into two groups. This boundary position is 60.5% of the distance between the stimulus dots. Since the method to estimate this

boundary position is identical to Fechner's method of adjustment to determine a JND, it is referred to here as a JNB (Just Noticeable category Boundary).

This value, 60.5, is numerically close to 61.8. Called the golden section, 61.8 is the unique result when the ratio of the smaller inter-dot distance (response to the grouped dot) to the larger inter-dot distance (response to the not grouped dot) equals the ratio of this larger inter-dot distance to the total range (distance between the stimulus dots). In that case, the three distance measures would be interrelated.

To learn if this result is a proportion or a fixed distance, the study was repeated with a different dot separation. To learn if the result is at all general, the study was conducted with various stimuli. These are tone-beeps separated in time, lines that varied in height, other lines that varied in length, tones that differed in duration, and tones that varied in loudness. In each case, subjects were given a joystick control and asked to select the computer produced value that was just not grouped with one or the other of the two stimulus elements. The results are summarized in Table 1.

Table 1

The just not grouped position in 9 three-element studies.

Task	Mean just not grouped proportion
Dots separated in space	0.64
" " " " replication	0.62
Tone beeps separated in time	0.58
Heights of adjacent, parallel lines	0.67
Successive line lengths (eyemovement needed)	0.73
Durations of successive tones:	
250 msec. intertone intervals	0.63
500 " " "	0.67
1,000 " " "	0.62
Loudnesses of successive tones	0.64
Mean Ratio =	0.64

Each result in Table 1 is is the average of two conditions. These are the just not grouped with the first or leftmost element position, the just not grouped with second (or rightmost) element. This averaging masks order effects that are prominent in the tone-duration study.

Table 1 shows just not grouped values. While we have only collected a few just still grouped estimates, each of these is smaller than the comparable just not grouped numbers.

According to these results, the JNB task is easy for people to perform, the data are reliable between and within subjects, and the results expressed as proportions of stimulus range are nearly identical across stimulus sets. These results allow the suggestion that there is a general grouping or categorizing process.

We have only pursued this finding a little. People were asked to produce the stimulus value that appeared just grouped, or just not grouped, with various sets of 2 to 6 vertical lines having various heights, and with sets of 2 to 6 dots separated

variously across space.

Three results from these studies may be of interest to the question: What is involved when people categorize collections of elements? First, the subjects had no difficulty understanding the "just grouped" and the "just not grouped" instructions. No one said the task was not sensible. Second, response variability both within and between subjects was surprisingly small. Third, the selected boundaries were more different, more variable, and required more time to produce when the stimulus elements were more variable. In general, the response values are directly proportional to the variability of the stimulus set.

No regular result like the 61.8 when using three elements was detected in these conditions using several elements. Here, the selected boundaries depend, or additionally depend, on specifics of the task. For example, the boundary position depends on configural relations between the stimulus members. Still, for each condition examined, the mean responses are directly proportional to stimulus variability, with configural factors described as an additive constant (15).

In AI and ME tasks, it is known that response variability increases with stimulus variability or stimulus range. The similar result occur here when the subject, rather than the experimenter, determines the category membership of an element. For each class of data examined, responses are regularly more variable when the stimulus set is more variable. If this is generally the case, it would support the suggestion that a fundamental nature of the categorization process is one of proportionalities.

References:

1. Gravetter, F. & Lockhead, G.R. Criterial range as a frame of reference for stimulus judgment. Psychological Review, 1973, 80, 203-216.

2. Green, D. M. Consistency of auditory detection judgments. Psychological Review, 1964, 71, 392-407.

3. King, M.C. Sequential effects in the variability of magnitude estimates of loudness. , 1980, Ph.D. dissertation, Duke University.

4. King, M.C. & Lockhead, G. R. Response scales and sequential effects in judgment. Perception & Psychophysics, 30, 1981, 599-603.

5. Lockhead, G. R. & King, M. A memory model of sequential effects in scaling tasks, Journal of Experimental Psychology: Human Perception and Performance, 1983, 9, 461-473.

6. Lockhead, G. R. Sequential predictors of choice. In, Preparatory States and Processes, S. Kornblum and J. Requin (Eds.), 1984, N.J., Erlbaum, 27-48.

7. Luce, R. D., & Green, D. M. The response ratio hypothesis for magnitude estimation, Journal of Mathematical Psychology, 1974, 11, 1-14.

8. Luce, R. D., Green, D. M., & Weber, D. L. Attention bands in absolute identification. Perception & Psychophysics, 1976, 20, 49-54.

9. Staddon, J., King, M., & Lockhead, G. Sequential effects in absolute judgments. Journal of Experimental Psychology: Human Perception and Performance, 1980, 6, 290-301.

10. Stevens, S.S. Psychophysics: Introductino to its perceptual, neural, and social prospects. G. Stevens (Ed.), 1975, NY, Wiley.

11. Ward, L. M. & Lockhead, G. R. Category judgments of loudness in the absence of an experimenter-induced identification function: Sequential effects and power-function fit. Journal of Experimental Psychology, 1972, 94, 179-184.

Work supported in whole or in part by AFOSR-85-0302, 1985-7.

12. Hinson, J. & Lockhead, G.R., Range effects in successive discrimination procedures. Journal of Experimental Psychology: Animal Behavior Processes, 1986, 12, 270-276.

13. Hinson, J. & Lockhead, G.R., Momentary and Global Maximizing, Behavioral Processes, 1987, 14, 21-34.

14. Lockhead, G. R., Category bounds and variability. In Object Perception: Structure and Process, B. Shepp and S. Ballesteros (Eds.), Erlbaum, NJ, in press. Also see Percepcion del Objeto: Estructura y Procesos, S. Ballesteros and B. Shepp (Eds.), Universidad Nacional de Educacion a Distancia, Madrid, Spain, in press.

15. Lockhead, G. R. Category Thresholds. Fechner Day. International Society for Psychophysics, 1987, Northampton, MA., 23-28.

16. Lockhead, G. R. & Hinson, J. Range and Sequence Effects in Judgment, Perception & Psychophysics, 1986, 40, 53-61.

17. Lockhead, G. R. & King, M. C. Variance in magnitude estimations, Fechner Day '86, B. Berglund, U. Berglund, & R. Teghtsoonian (Eds.). International Society for Psychophysics, Stockholm, 1986, 59-64.

END

DATE

FILMED

APRIL

1988

DTIC