A combination of information integration methodology and measures of retention was used to investigate how subjects differentially attend to and weight information in judgmental tasks. Subjects were shown sets of test scores for hypothetical students and were asked to rate the performance of each student or predict each student's performance on a comprehensive final exam. Subjects given the prediction task showed a recency effect—their predictions were weighted more heavily by the second half of the scores than by the first—and tended to discount a deviant score. Subjects given the rating task showed a recency effect only when explicitly told to take into account trends in scores. The prediction group that showed the most discounting of deviant scores was also the group that had the highest rate of recall of deviant scores on tests of retention. Discounting thus appears to be an active process, rather than a lowered level of attention. (Author)
Information Integration, Retention, and Levels of Information Processing

Irwin P. Levin
University of Iowa

A combination of information integration methodology and measures of retention was used to investigate how subjects differentially attend to and weight information in judgmental tasks. Subjects were shown sets of test scores for hypothetical students and were asked to rate the performance of each student or predict each student's performance on a comprehensive final exam. Subjects given the prediction task showed a recency effect—their predictions were weighted more heavily by the second half of the scores than by the first—and tended to discount a deviant score. Subjects given the rating task showed a recency effect only when explicitly told to take into account trends in the scores. The prediction group that showed the most discounting of deviant scores was also the group that had the highest rate of recall of deviant scores on tests of retention. Discounting thus appears to be an active process, rather than a lowered level of attention.

The information integration approach (Anderson, 1974) has proved quite useful in describing complex cognitive processes with simple algebraic models. In many instances, the descriptive models of information integration theory have been shown to be superior to normative models in explaining and interpreting performance in human judgment tasks (Anderson, 1974; Levin, 1975; Shanteau, 1972). That is, the descriptive models provide a more fruitful framework for operationalizing and quantifying the effects of variables that affect judgmental and decision processes.

While providing an adequate quantified description is an important step in understanding human judgments, it is also important to investigate the underlying structure of the process captured by the descriptive model. Thus, for example, it would be of interest to relate the differential weighting of different components of information to the information integrator's tendency to attend differentially to these different components. This is analogous to the old problem of identifying functional, as contrasted with nominal stimuli in learning tasks.

A recent study by Levin (1976) will help to formalize this problem and lay the groundwork for the new study to be reported in this paper. In this study subjects were asked to make judgments about each of a series of sets of numerical scores. This follows earlier work in "intuitive statistics" (Anderson, 1968; Levin, 1975; Peterson & Beach, 1967). Some of these sets contained an "outlier" score that was deviant from the rest. In Experiment 1 the scores in a set were displayed simultaneously, and in Experiment 2 the scores were displayed serially. Each experiment included a "descriptive condition" where subjects were asked to estimate (subjectively) the mean or average value of the scores presented in each set and an "inference condition" where subjects were asked to infer the mean of the population from which
the set of scores represented a random sample. In Experiment 1—simultaneous presentation—the deviant or outlier scores were weighted at least as much as other scores in forming a descriptive judgment. But deviant scores were weighted less or discounted in forming an inference. This represents an experimental demonstration of the "intuitive" analog of descriptive and inferential statistics. However, in Experiment 2—serial presentation—the discounting effect was reduced (when subjects responded at the end of each sequence) or eliminated completely (when subjects were asked to adjust their responses following each new piece of information).

Levin concluded that subjects will discount a deviant score in making an inference from a sample to a population when they recognize such a score to be unrepresentative of the population. However, the ability to detect an unrepresentative score depends upon the way the information is presented.

The present study represents an attempt to replicate and extend the findings of the Levin (1970) study with regard to the distinction between intuitive descriptive and inferential statistical judgments. In addition, the present study represents an initial attempt to develop a method for measuring how subjects attend to the various components of an information integration task. The method involves a combination of information integration methodology and measures of retention. Anderson and Hubert (1963) also used a recall task in conjunction with an information integration task, but, in contrast to the present methods, their subjects were forewarned that they would have to recall the material presented.

In the present study subjects were not forewarned that they would be tested for recall. Subjects were given a series of sets of eight numbers each, where the numbers in a set were said to be a succession of test scores for a student. Some subjects were told to rate the performance of each
student in relation to the other students (analogous to the descriptive condition of the previous study), and some subjects were asked to predict the performance of each student on a subsequent comprehensive final exam (analogous to the inference condition of the previous study and the guess condition of Anderson, 1968). Following a number of such rating trials or prediction trials, subjects were given additional sets of scores and were then asked (without pre-cuing) to recall the scores they had just seen or estimate the trend present in the scores. The retention measures should then provide evidence on how the subjects were using information in the judgmental (integration) tasks.

METHOD

1. Construction of stimulus sets

   a. 9 sets were formed by factorially manipulating the mean of the first 4 scores (M1 = 50, 60, or 70) and the mean of the second 4 scores (M2 = 50, 60, or 70). All scores in these sets were within 10 points of the mean of the set.

   b. 4 sets were formed which contained deviant scores. These sets had a mean of 60 and the deviant score was approximately 25 points above or below the mean and appeared in the first half or second half of the scores.

   c. End anchors, practice trials, and filler sets were included.

2. Instructional conditions

   a. Group 1 (n = 17 students from introductory psychology classes at the University of Iowa): Subjects were asked to rate the performance of each student using a 20-cm. graphic (line-mark) rating scale (converted to a 20-point scale) labeled "very good" and "very bad" at the extremes. These subjects were told that each score should be of equal importance in rating performance.
b. Group 2 (n = 19): Subjects were asked to rate the performance of
each student using the above scale. In addition, they were told to
take into consideration whether or not the student showed improvement
over the course of the semester.

c. Group 3 (n = 20): Subjects were asked to predict the performance
of each student on a comprehensive final exam using the same graphic
scale as the other groups.

d. Group 4 (n = 17): Subjects were asked to predict final exam
performance as in Group 3, but, in addition, they were told that
occasionally a student will have an unusually "good" or "bad" day
that may not necessarily be a good predictor of subsequent
performance. These additional instructions were designed to promote
discounting of deviant scores.

3. **Retention questions** (following 25 trials of the information integration
task)

a. Trend question: After making their rating or prediction response
to a set of scores where M1 = 74 and M2 = 61, subjects were asked
to indicate whether the first half or the second half of the scores
was higher.

b. Uncued recall: After making their rating or prediction response to
a set of scores containing an "outlier," subjects were asked to
write down as many of the numbers in the set as they could recall.
(The order of presentation of the trend question and the uncued
recall was counterbalanced across subjects.)

c. Cued recall: On the final trial of the experimental session, the
subjects were told ahead of time that they would be asked to recall
the scores in the next set. This was included to provide a baseline
for examining the uncued recall and to provide a basis for analyzing
the comparability of the various groups.

4. Procedures and time intervals

The eight numbers in a set were printed on 7 x 22" cards which were held
up by the experimenter for a period of 8 sec. Subjects were given
approximately 5 sec. to respond in Part 1 (information integration task)
and 30 sec. to respond in Part 2 (retention task). Subjects were tested
in groups of 8 to 10.

RESULTS

1. Information integration analysis of responses to sets without deviant
scores (see Figure 1)

a. Group 1: Both M1 and M2 were significant sources of variance. The
M1 x M2 interaction did not reach significance at the .05 level.
There was no significant order effect. Results are consistent with
a simple averaging model.

b. Group 2: The M2 effect was large and statistically significant,
but M1 did not approach statistical significance. The M1 x M2
interaction was significant. There was a large and statistically
significant recency effect. Results are consistent with the notion
that subjects (as instructed) took into account the trend in the
scores, even to the extent of sometimes "penalizing" students who
scored higher in the first half.

c. Groups 3 and 4: M1 and M2, and M1 x M2 were significant sources of
variance for each group. There was a statistically significant
recency effect for each group (but not as large as in Group 2).
Results are consistent with the notion that the most recent events
(scores) are weighted more heavily in predicting future performance.
2. **Information integration analysis of responses to sets with deviant scores** (see Figure 2) Note: If responses to sets with + deviant score are lower than responses to sets with - deviant score, then this is evidence for discounting.

a. Group 1: Deviant scores are weighted more than other scores (opposite of discounting) and this difference is statistically significant.

b. Group 2: Deviant scores are weighted less than other scores, but this difference did not reach statistical significance.

c. Group 3: Deviant scores are weighted less than other scores, but this difference reached statistical significance only when the deviant score occurred in the second half.

d. Group 4: Deviant scores are weighted less than other scores, and this difference was statistically significant. Results are consistent with the notion that subjects tend to discount a deviant score in an inference condition when it is perceived to be unrepresentative of the population.

3. **Retention task** (see Table 1)

a. Trend question: All groups performed at a relatively high level. There were no significant group differences. Conclusion: Since scores are probably read in left-to-right order and are said to represent performance on successive exams, subjects notice trends even when that information is irrelevant to their judgments (Group 1).

b. Uncued recall: Overall performance (% correct) does not differ significantly or systematically across groups. The probability of correctly recalling the deviant score tended to be higher for the groups who showed the most discounting, but the differences were not significant.
Table 1
Summary of Responses on Retention Task

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Trend question:</td>
<td></td>
</tr>
<tr>
<td>% correct</td>
<td>76.5</td>
</tr>
<tr>
<td>Uncued recall:</td>
<td></td>
</tr>
<tr>
<td>Mean correct (out of 8)</td>
<td>3.29</td>
</tr>
<tr>
<td>Uncued recall:</td>
<td></td>
</tr>
<tr>
<td>% correct on deviant score</td>
<td>52.9</td>
</tr>
<tr>
<td>Cued recall:</td>
<td></td>
</tr>
<tr>
<td>Mean correct (out of 8)</td>
<td>4.65</td>
</tr>
<tr>
<td>Cued recall:</td>
<td></td>
</tr>
<tr>
<td>% correct on deviant score</td>
<td>41.2</td>
</tr>
</tbody>
</table>
c. Cued recall: All groups were comparable on overall performance and did approximately 50% better on cued than on uncued recall. The group with the most discounting (Group 4) had the highest rate of recall for deviant scores. (For cued and uncued recall combined, Group 4 had a significantly higher rate of recall for deviant scores than did other groups.)

d. Auxilliary finding: On uncued recall subjects recalled deviant scores at a significantly higher rate than other scores, but the reverse was true for cued recall. This is consistent with the notion that uncued recall taps the fact that deviant scores stand out and are noticed more, but in cued recall subjects attempt to maximize their probability of being correct by guessing items from the most frequent range of scores.

DISCUSSION

The results for the information integration task of the present study tend to support the findings of Anderson (1968) and Levin (1976b) who found differences between rating (or descriptive) tasks and prediction (or inference) tasks. Statistically significant discounting effects were found only in the prediction task of the present study. Another interesting finding was that recency effects were found in the prediction task.

There was no systematic relationship between the degree to which the various experimental groups displayed recency effects and their ability to identify trends on the retention task. Thus, the attempt to define a measure of retention which would show the extent to which subjects were attending to trends in the scores failed to account for performance differences in the integration task. However, as indicated, subjects may have developed a "set" to notice trends even when this information was irrelevant to their judgments.
The retention measure dealing with deviant scores was somewhat more informative. Group 4 had both the highest rate of discounting of deviant scores in the integration task and the highest rate of recall of deviant scores in the retention task. Thus, "discounting" can be described as an active process of attending to and systematically eliminating the deviant score—i.e., assigning it a zero weight—rather than a lack of attention to the deviant score.

While the retention measures did not discriminate greatly between groups showing different integration rules, the present method of combining information integration methodology and retention measures seems promising. A number of earlier studies of information integration have shown different integration rules depending on whether information was presented simultaneously or serially (Anderson, 1973), the nature of instructions regarding the processing of deviant information (Anderson & Jacobson, 1965; Levin, 1976b), and the use to be made of the information presented (Levin, 1976a). In these cases, it would be particularly interesting to detect differences in the way subjects attend to (and retain) different aspects of the information presented. Such studies would serve to tie together the following research topics in human information processing: levels of information processing (Craik & Lockhart, 1972), information integration, and retention.
REFERENCES


Levin, I. P. Comparing different models and response transformations in an information integration task. *Bulletin of the Psychonomic Society, 1976, 7, 78-80.* (a)

Levin, I. P. Processing of deviant information in inference and descriptive tasks with simultaneous and serial presentation. *Organizational Behavior and Human Performance, 1976, 15, 195-211.* (b)