Judging Psychopathology: Expectation and Contrast

MELVIN MANIS

University of Michigan and Ann Arbor VA Medical Center

AND

JOAN R. PASKEWITZ

Ann Arbor VA Medical Center

An experiment was performed to assess the role of expectations in psychosocial contrast effects. Subjects were initially presented with a series of word definitions that were either high or low pathology; their task was to indicate those definitions that had been produced by schizophrenic (vs nonschizophrenic) patients. After this "induction" series, subjects were presented with a common "test" series that included both definitions and handwriting samples; they evaluated the degree of psychopathology that was implied by each item. Some subjects were required to state their expectations in connection with the test items. First they attempted to guess the percentage of schizophrenics in the overall test series; second, before being presented with each test item, they guessed whether the patient in question was (or was not) schizophrenic. Subjects' expectations were consistent with the induction series to which they were assigned; those assigned to the high-pathology induction expected the test series to include more schizophrenic patients than those in the low-pathology group. Nonetheless, despite these expectations, subjects showed a contrast effect in their evaluations of the test definitions (but not in evaluating the handwriting samples). That is, subjects from the low-pathology induction group rated the test definitions as being more pathological than did those in the high-pathology condition. Within-group correlations indicated that the subjects' expectations were directly related to subsequent judgments. A two-path model is discussed, relating (a) the respondents' induction group assignment, (b) his/her expectations, and (c) subsequent test judgments.

The social judgment literature includes many demonstrations of the fact that our evaluations (interpretations) of incoming stimuli can be

363

systematically influenced by the context in which these stimuli are presented. The contrast effect is a prime example. Respondents who are asked to rate a series of stimuli along some specified dimension normally displace their judgments away from the bulk of their recent experience; for example, a respondent who has previously evaluated the psychopathology implied in a series of confused word definitions will normally rate "midscale" definitions as implying less psychopathology than someone who has previously rated low-pathology definitions (Manis & Paskewitz, 1984). This type of effect has been repeatedly demonstrated in a wide range of psychophysical and psychosocial domains, including judgments of weight (Helson, 1964), judgments of affect (Manis, 1967; Manis & Armstrong, 1971), and judgments pertaining to the severity of different crimes (Pepitone, 1976).

Despite the robust replicability of the contrast phenomenon, several theoretical approaches suggest that biased judgmental experiences might, in some circumstances, contribute to a very different, assimilative-like pattern. Consider an experiment in which a respondent has just evaluated several high-pathology definitions that have been "randomly selected" from the files of a hospital. This experience might plausibly lead to an expectation that another sample of definitions from the same population (hospital) would also be relatively high in psychopathology. Such an expectation, in turn, might increase the likelihood that our hypothetical judge would regard the individual definitions in this second sample as being indicative of severe psychopathology. This account focuses on the possibility that (a) the judge's expectations for the future will normally be congruent with his/her recent experience and (b) these expectations might exert a direct influence on subsequent judgments.

The link between expectation and judgment has been given a more formal expression in Bruner's model of perception (1957) and in perceptual theories that posit a Bayesian-like information processing system. Bruner emphasized the importance of "perceptual readiness," a principle that has received recent emphasis in studies of priming (Higgins, Rholes, & Jones, 1977), which show that the accessibility of a perceptual category can be enhanced (at least temporarily) by frequent evocations.

Bayesian analyses of social judgment emphasize a related idea; the Bayesian view suggests that a judge's categorization of a given vocabulary definition (e.g., "schizophrenic") may depend not only on the definition then being judged, but may also be affected by the judge's base-rate expectations. Thus, according to Bayes' theorem, a judge who anticipated that the next patient s/he was to encounter would probably be schizophrenic, should be *more* likely than other judges to produce a diagnosis of "schizophrenia" after studying a concrete sample of that patient's behavior. Kahneman and Tversky (1973) have challenged the adequacy of the Bayesian model as a valid description of human judgment, showing that people are normally less than optimally affected by base-rate con-

siderations. Nonetheless, these investigators and others (e.g., Manis, Dovalina, Avis, & Cardoze, 1980), have reported clearcut evidence that base-rate information often exerts an effective influence on social judgments.

Previous research in this program (Manis & Paskewitz, 1984), has demonstrated the replicability of the contrast phenomenon in judgments of psychopathology; some aspects of this work, however, offer suggestive evidence for an expectation model. For example, in one study, subjects who had initially evaluated a series of high-pathology definitions (and who had subsequently rated a series of "midscale" test definitions) anticipated that an additional sample of patients from that same hospital would probably include a relatively high proportion of schizophrenics. This suggests that expectations are *directly* affected by the respondent's previous experience with "biased" stimulus arrays.

Our earlier work also provided suggestive evidence that the respondent's expectations might have a direct impact on his/her subsequent judgments. Most importantly, there was normally a positive correlation between the respondents' reactions to the definitions shown in the initial ("induction") series, and their subsequent evaluations of the "test" definitions; within each experimental condition, respondents who had interpreted many of the induction items as being "schizophrenic" tended to show a similar reaction to the test definitions. Moreover, this correlation was not attributable to a simple response bias (i.e., to individual differences in the respondents' readiness to use the label "schizophrenic"). Since the respondents in this experiment had all been told that the test definitions would come from the same population (hospital) as the previously presented induction series, we may reasonably interpret their reactions to the induction items (i.e., the number of patients they thought to be schizophrenic) as an indirect measure of what they expected to find in the test series. Results from the present experiment support this reasoning.

Summing up, then, there are both theoretical and empirical reasons to believe that (a) the respondent's experience in a judgmental situation may directly affect his/her expectations concerning subsequent (to-be-evaluated) stimuli. Moreover, (b) these expectations may directly influence subsequent evaluations. The present experiment was designed to provide additional evidence regarding this formulation in a setting that allowed us simultaneously to examine the presumably *assimilative* impact of the respondents' expectations, and the *contrastive* effects that normally result from exposure to a biased stimulus array. As we shall see, our results support a model in which biased exposure affects subsequent judgments through two relatively independent "paths." One path, which operates through the respondent's expectations promotes an *assimilation* effect; the other more direct path induces a *contrastive* judgmental pattern.

The design of our experiment was simple. After receiving either a high-
or a low-pathology series of vocabulary definitions, some subjects were
informed that they were going to evaluate behavior samples from an
additional set of patients, selected from the same pool (population) of
hospital records as the first sample. Before receiving the "new series,"
subjects in this condition were instructed to guess how many schizophrenics
were likely to be included in this second (test) sample. Moreover, for
these subjects the test series was administered in a somewhat unusual
fashion; prior to the presentation of each test item the respondents
"guessed" whether the patient in question was (or was not) schizophrenic.
A sample of the patient's behavior (either a word definition or a handwriting
sample) was then presented, and the respondent indicated his or her
"diagnosis" (schizophrenic or not schizophrenic) based on the available
information, and rated the patient's degree of "thought distortion." Fol-
lowing this, the respondents guessed whether the "next patient" in the
series would be schizophrenic or not, etc. The sequence of guesses
followed by evaluations continued throughout the test series.

Since it seemed possible that the subjects' evaluations of the test items
might be affected by the questions we posed regarding expectations (see
above), a parallel set of judgments was collected from people who were
not asked about their expectations. There were, in fact, two control
groups in which respondents evaluated the test materials without stating
their expectations. One group received the same information as the "ex-
pectation" subjects; they were told that the test series included material
that was drawn from a second sample of patients from the same hospital
as the patients in the first (induction) sample. A second group was told
that the test series included additional behavior samples from the same
patients whose definitions they had just rated. These contrasting instructions
were designed to explore the respondents' reactions to "new patients,"
as compared with their reactions to "new behavior samples."

A respondent who thought that the test series was comprised of behavior
samples from the same patients as the induction series might, we believed,
assimilate the test items to the previously presented induction definitions
(high versus low pathology), instead of showing the contrast effects that
we had so regularly obtained in previous studies. We felt, for example,
that a respondent who had been assigned to the high-pathology condition,
and who believed that the test items derived from the "same patients"
as those in the preceding induction series might perhaps take account
of the pathological impression she/he had initially formed of these patients
when evaluating their subsequent test responses. An approach of this
sort would yield an assimilative pattern, rather than the contrast effects
we had so regularly obtained in the past. This prediction was not fulfilled,
however.

METHOD

Experimental Design

The considerations that are outlined above led to an experimental design that involved two main "between subject" variables based on differences between the stimuli presented during the induction series (high- vs low-pathology definitions), and on differences in the instructions that were presented just prior to the test series (there was one "expectation" condition, and two "no-expectation" control conditions). In addition, the experiment involved two "within-subject" variables. One of these variables was devised by considering the subject's performance as she/he proceeded through the test series, which was divided into three trial blocks. The other within-subject variable reflected the fact that there were two types of test items (word definitions and handwriting samples) that were arranged in a single unpredictable sequence. This last variable (item types) was introduced to explore the possibility that the effects produced by a biased induction series might depend on the similarity between the materials presented in the test and the induction phase of the experiment.

Previous studies in this program have shown that contrast effects are quite robust on test items that resemble the items in the induction series (e.g., when both series are composed of definitions); on the other hand, when the test items are "different" from those in the induction series contrast effects are weak and unreliable (Manis & Paskewitz, 1984). We hoped to replicate this pattern of results in the present study, using a more carefully selected set of test items, and more importantly, to see how our respondents' expectations would affect their judgments of the definitions and the handwriting samples.

Subjects

The subjects in this study were undergraduates at the University of Michigan who were paid for their participation. The study involved a total of 108 respondents; each experimental condition included half males and half females. Subjects were normally run in small groups (2–7), with some people being tested individually.

Induction Booklets

Subjects were told that the study concerned "social judgments." After explaining that vocabulary definitions are often used to help in clinical evaluations, we presented the respondents with a series of 26 definitions in booklet form (1 to a page). These definitions were to be used in deciding if the individual patients who had produced them were schizophrenic or not. Subjects were told that the definitions which they would see came from a representative cross-section of the patients in a state hospital. They were given 10 s to react to each definition.

The first six definitions in the induction series had all been rated as roughly midscale with respect to psychopathology (on an 11-point continuum) by a norm group. These items were used to quantify individual differences in response bias, as indicated by the number of "schizophrenics" that each respondent identified. Respondents who rated *many* of these definitions as being indicative of schizophrenia were regarded as having a relatively low threshold for the schizophrenic response label.

The next 20 definitions in the induction series constituted our main induction manipulation; half of the subjects received low-pathology items, with norm ratings between 1.50 and 3.95 (on an 11-point scale), while the other half was presented with high-pathology items (norm ratings between 8.00 and 10.68).

Test Items

When the induction booklets had been completed, subjects received a second series of behavior samples that was to be evaluated for psychopathology. There were two types of

test items: handwriting samples (each patient had presumably written the words "father" and "mother" on an admission form) and additional vocabulary definitions. Each test item was presented by means of a slide projector, with a new slide appearing every 10–15 s; before presenting the next item in the series, the experimenter waited until each respondent had completed his/her reaction to the preceding slide.

For each test item the respondents indicated (a) the amount of thought distortion that seemed to characterize the patient, using a 7-point scale that ranged from 1 (normal) to 7 (highly distorted/disturbed); (b) whether the patient seemed schizophrenic or not; and (c) their confidence with respect to the "diagnosis" provided in (b), using a 7-point response scale that ranged from "guess" (1) to "very certain" (7). The definitions had norm ratings that ranged between 4.1 and 7.8 on an 11-point scale; the handwriting samples received norm ratings between 4.1 and 6.6. The test series included 32 different items (16 definitions and 16 handwriting samples); the first 2 items (1 handwriting sample, 1 definition) served as warm-ups, with the remaining 30 items constituting the main source of judgment data for this experiment. Definitions and handwriting samples were presented in a single un-predictable order to all respondents.

Expectations

One third of the subjects assigned to the low-pathology condition and one third of those assigned to the high-pathology condition provided "expectation" measures in addition to the judgments that were obtained from all respondents. Following the induction series, these subjects first attempted to "guess" the number of schizophrenic patients in the upcoming test series, which (they were told) had been drawn from the same hospital as the first (induction) sample. In addition to this global expectation statement, these respondents also attempted to guess, on a trial-by-trial basis, whether the "next patient" in the series was or was not schizophrenic. Following each of these guesses, the respondent was presented with the next item in the test series (a definition or handwriting sample), and evaluated that patient's level of psychopathology using the procedures outlined previously. This cycle (a guess, followed by "informed" judgments based on either a vocabulary definition or a handwriting sample) was repeated for each of the 32 test trials.

To guard against the possibility that the introduction of expectation measures might in some way interfere with the context effects that had previously been obtained, "control" subjects were not required to provide global expectation measures (at the start of the test series), nor were they required to guess as to the schizophrenic versus nonschizophrenic character of the individual "patients" in the test series. Subjects assigned to the expectation condition were compared with these "control" subjects to see if the expectation measures had affected the respondents' evaluations of the individual "patients" presented in the test series.

"Control" subjects were divided into two groups, depending upon their understanding as to the origins of the patients in the test series. One group (one third of all subjects) was told that the test sample was drawn from the *same hospital* as the induction sample; this was the same understanding as was provided to the subjects from whom we collected "expectation" measures. Another group (the final one third of all subjects) was provided with a rather different understanding; these subjects were led to believe that the test items were obtained from the *same patients* who had provided the definitions for the induction series.

RESULTS

Global Expectations

The respondents' expectations for the overall test series were roughly congruent with the sorts of definitions they had seen during induction.

That is, subjects generally anticipated that the test series was likely to include a continuation of what had come before ("more of the same"). Respondents who had been assigned to the high-pathology induction series anticipated that 60% of the behavior samples shown in the test series would be drawn from the records of schizophrenic patients; by comparison, respondents assigned to the low-pathology induction expected that only 38% of the patients in the test series would be schizophrenic. The difference between these estimates is highly significant [$t(34) = 3.83$, $p < .001$].

Trial-by-Trial Expectations

Figure 1 presents mean values that reflect the respondents' individual (trial-by-trial) "guesses" as they proceeded through the test series; these data are arranged in three blocks, each of which included 10 different trials. The first thing to notice about these data is the difference between the high- and low-pathology groups. Respondents who had previously seen predominantly high-pathology definitions generally anticipated that the test series would continue in a similar vein; in each trial block, the high-pathology group expected to encounter more schizophrenics than did subjects assigned to the low-pathology condition. To assess the reliability of this pattern the data were subjected to an analysis of variance (group assignment × respondent's sex) in which the dependent variable was the number of schizophrenic guesses observed in each trial block. This analysis yielded a significant "group" effect [$F(1, 32) = 4.44$; $p < .05$] for block 1, which attenuated over the succeeding trial blocks.

Judgments of Psychopathology

Before exploring the relationship between our respondents' stated expectations and their subsequent judgments, it seemed prudent to determine whether the introduction of expectation questions had affected the judgments associated with the various test items, and whether we had succeeded in replicating the general pattern of results we had observed in previous
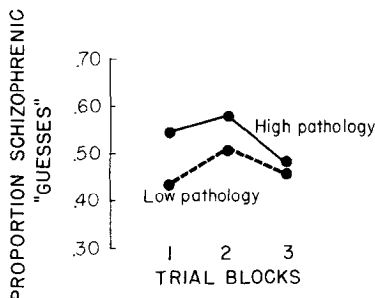

FIG. 1.   Mean trial-by-trial "guesses" for each test series trial block.

experiments (Manis & Paskewitz, 1984). We therefore subjected the judgment data to an analysis of variance, using a dependent variable that combined the respondents' ratings of thought pathology (1–7) and their judgments regarding the patient's diagnosis as schizophrenic or not (weighted by their confidence in that assessment). These two variables were highly correlated ($r = .79$) and were simply added together, after standardization, to ensure that they would receive equal weight in the final index.

The results of this analysis indicated that the respondents' judgments were not reliably affected by the instructions presented to the different groups. The main effects and the interactions involving this instruction variable were uniformly nonsignificant. Pairwise comparisons between the "same hospital" and "same patient" groups, and between the "same hospital" and the "expectation" groups yielded similarly nonsignificant results. This means that the respondents in our "expectation" group, who were required to make explicit guesses before evaluating the various test items, did not differ from the other respondents in evaluating the behavior samples that were presented in the test series.

Figure 2 presents the main results deriving from this analysis. The vertical axis is a difference score that was obtained by subtracting the judgments of those assigned to the high-pathology condition from the judgments of those in the low-pathology group. Positive difference scores thus indicate the presence of a contrast effect (i.e., respondents assigned to the low-pathology induction series rated the test items as implying more psychopathology than did those in the high-pathology group). Results are presented separately for the vocabulary items and for the handwriting samples.

These results successfully replicate earlier studies in this program of research. The respondents who initially read high-pathology definitions (in induction) rated the definitions of the test series as being significantly less pathological than the respondents who had initially rated low-pathology
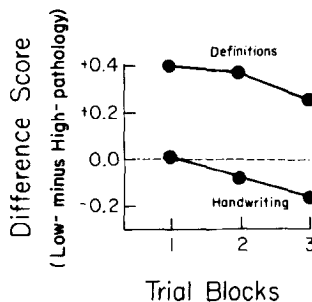

FIG. 2. Combined index difference scores (low- minus high-pathology induction) × trial block for test definitions and handwriting samples.

definitions—contrast. On the other hand, the induction experience (high vs low pathology) did not appreciably affect their reactions to the hand-writing samples.

The reliability of these effects was established through analysis of variance. First, in an overall analysis that involved all of the test items, there was a significant interaction between the respondents' induction condition (high vs low pathology) and the type of behavior sample presented for judgment (definition vs handwriting sample; $F(1, 96) = 31.27$; $p < .001$). Subsequent analysis indicated (as suggested in Fig. 2) that the difference between the high- and low-pathology groups was significant when the respondents were evaluating definitions, but not when they were evaluating handwriting samples. The F ratios for these latter comparisons were $F(1, 96) = 23.05$, $p < .001$, and $F(1, 96) = 1.42$, $p > .20$, respectively.[1] If we restrict our attention just to the respondent's *diagnoses* (schizophrenic vs not schizophrenic), we find that subjects in the low-pathology group labeled 54% of the test definitions as schizophrenic, compared with 42% from subjects in the high-pathology group.

One of the most interesting aspects of these data is the difference between the results that are summarized in Figs. 1 and 2. While the high-pathology respondents anticipated that the various behavior samples in the test series had a relatively high likelihood of coming from schizophrenic patients (compared to the low-pathology group, see Fig. 1), when subsequently presented with concrete vocabulary definitions they rated them as *less* pathological than did respondents assigned to the low-pathology condition (see Fig. 2); i.e., in response to the vocabulary items, the judgments of the two groups were the reverse of the expectations they had just expressed. For the handwriting samples, on the other hand, there was no appreciable difference between the judgments of the high- and low-pathology groups.

Correlations between Expectations and Judgments

The results that are presented above indicate that the respondents' reactions to the test items did not directly reflect the sorts of expectations they had formed. Indeed, it seemed conceivable that their evaluations of the vocabulary definitions may have been *inversely* related to their expectations. Some theories might suggest, for example, that a respondent who anticipated that the next test trial would present a behavior sample taken from a schizophrenic patient might, upon examining one of our "midscale" test definitions, conclude that the item in question was sub-

---

[1] Both curves in Fig. 2 have a general downward trend. For the *vocabulary definitions*, this signifies a steady diminution in the magnitude of the contrast effect, a pattern that has appeared quite regularly in our earlier work. However, while the handwriting data show a similar downward thrust (toward assimilation), this pattern is of uncertain replicability, for it has *not* been observed in earlier studies of this type (see Manis & Paskewitz, 1984).

stantially less confused than s/he had expected, leading to the judgment that the patient in question was *not schizophrenic*. A scenario of this sort would lead to a *negative* relationship between the respondents' expectations and their subsequent judgments. To examine this possibility the respondents' guesses on each trial (schizophrenic vs not schizophrenic) were correlated with their judgments for that trial. The results are shown in Table 1, which summarizes the correlations separately for the two types of test items. The Table presents average correlation coefficients (Pearson r's) for the test items at various points in the test series. For example, the top left entry in the Table (.10) indicates that for the first five definitions of the test series, there was (on the average) a positive relationship between the respondents' "guesses" and their judgments.

The most striking thing about Table 1 is the absence of any negative entries; that is, regardless of the type of test item they were evaluating (definition or handwriting sample), and regardless of the positioning of the item to be evaluated within the test series (early or late), there was a positive relationship between the respondents' guesses and their immediately subsequent judgments. People who expected that the upcoming behavior sample was likely to be taken from the records of a schizophrenic patient were more likely to label that test item as being schizophrenic (after it was shown) than were those who expected the next patient to be nonschizophrenic.

Table 1 summarizes the results of 30 different test trials. Twenty-five of these correlations were positive, and only five were negative, an outcome that has a p value that is significant beyond the .001 level when evaluated by the binomial test (Siegel, 1956). This pattern appeared in a very similar form whether the test item in question was a word definition or a handwriting sample; thus, 13 of the 15 test definitions showed positive correlations between the subjects' expectations and their diagnoses, while the handwriting samples yielded twelve positive correlations and three negative ones. Finally, these results were essentially unaffected when the data for each test trial were recomputed as partial correlations to

TABLE 1
CORRELATIONS BETWEEN "GUESSES" AND
SUCCEEDING SCHIZOPHRENIC JUDGMENTS

	Trial block			
Test item	1	2	3	Total
Definitions	.10	.09	.17	.12
Handwriting	.07	.13	.18	.13
All items	.08	.11	.18	.12

Note. Trial-by-trial correlations are averaged within trial blocks.

take account of (a) the induction group (high versus low pathology) to which the individual respondents had been assigned and (b) individual differences in overall readiness to assign the label "schizophrenic" (as measured by the number of midscale definitions the respondent had classified as schizophrenic at the start of the induction series).

These data indicate that there was a consistent positive relationship between the subjects' expectations for a given "patient" and their subsequent evaluation of his/her behavior sample. Nonetheless, the magnitude of these correlations was quite low, averaging only .12 for both the partialed and the unpartialed correlations. It seems likely, however, that the magnitude of these correlations was adversely affected by the unreliability of the two measures that were being correlated. Since each variable was measured by a single item (i.e., a "guess" in one case, a judgment concerning schizophrenia in the other), it is reasonable to assume that there was substantial measurement error in both cases, although limitations in the data set made it impossible to establish this in quantitative terms and to correct for attenuation.

Global Prediction and Induction Performance as Predictors of Test Performance

The correlational results that are presented above describe the "molecular" structure of our respondents' reactions, relating their expectations to their subsequent judgments on a trial-by-trial basis. We now turn to a more molar presentation in which we relate the respondents' *general expectations* (at the start of the test series) to the psychopathology that they reported in evaluating the two types of test items. While this approach is insensitive to the fine structure that is reported in Table 1, it enables us to consider a closely related question using a more reliable, multi-item "criterion" variable (the respondent's average response to the definitions and to the handwriting samples).

As noted previously, upon completing their assessments of the definitions in the induction series, all respondents in the expectation condition were asked to indicate the percentage of the to-be-judged test items that they anticipated from schizophrenic patients. To determine the relationship between these "global expectations" and the respondents' subsequent reactions to the test items we created multitrial indices, based on each respondent's average reactions to (a) the 15 vocabulary definitions, (b) the 15 handwriting samples, and (c) the total series of 30 different behavior samples (a plus b). We then correlated our "global expectations" measure with each of these criteria, while partialing out effects that were attributable to the respondent's group assignment (high vs low pathology) and to individual differences in their readiness to use the label "schizophrenic" (as determined at the beginning of the induction series). Table 2 presents these partial correlations for two related criterion measures: one measure

was solely based on the respondents' diagnoses (schizophrenic vs not schizophrenic), while a second measure included not only their diagnostic judgments, but also their confidence in each "diagnosis" and their ratings of thought disorder. While these results were promising in showing positive relationships between our respondents' global expectations and their subsequent judgments, only the results for the handwriting samples were significant at the .05 level. Reactions to the vocabulary definitions, on the other hand, were not significantly related to the global expectation measure for reasons that remain obscure.

Induction Performance and Test Performance

The results displayed in Table 2 suggested that when evaluating the handwriting samples, the respondents' global expectations had a direct but modest effect on their diagnostic judgments. However, since the predictor variable in these analyses consisted of a single item (i.e., the respondents' stated expectation regarding the percentage of schizophrenics in the upcoming test series), it seemed likely that this measure of expectation had but limited reliability. If so, this would inevitably dilute its relationship to the respondents' subsequent evaluations. We therefore sought a more reliable measure of "expectations," and were led to examine the judgments our subjects had produced in response to the induction series.

Since our instructions to all subjects suggested that the two patient samples would be rather similar on the whole (indeed, one third of the respondents was told that the test series included behavior samples from the "same patients" as the induction series), we reasoned that the number of schizophrenic "diagnoses" that each respondent produced during the induction series (his/her induction performance) might plausibly be regarded as an implicit indicator of what she/he expected to find in the forthcoming test series. This assumption is consistent with the fact that for the 36

TABLE 2
PARTIAL CORRELATIONS BETWEEN GLOBAL
PREDICTION AND TEST JUDGMENTS

Test item	Criterion	
	Schizophrenic judgment	Combined index
Definitions	.15	− .02
Handwriting	.36*	.35*
All items	.29	.18

Note. These correlations control for (a) the psychopathology level of the induction series and (b) response bias (see text).
 * $p < .05$, two tailed.

respondents for whom expectation measures *were* available, there was a substantial correlation ($r = .52$, $p < .01$) between the number of "schizophrenic diagnoses" in the induction series and these subjects' stated expectations for the upcoming test series.

To assess the relationship between induction performance (as a predictor) and our respondents' subsequent test judgments, we calculated a series of partial correlations like those presented in Table 2, the main difference being that our "expectation" measure was now based on the respondent's summed reactions to the individual induction items, rather than his/her verbalized expectations for the forthcoming test series. The results are shown in Table 3. Our judgmental "criterion" measure is a standardized index that includes the respondent's diagnosis (schizophrenic or not schizophrenic), weighted by his/her confidence in this diagnosis, plus a separate judgment of thought pathology. Similar results were obtained when this analysis was repeated using just the respondents' diagnoses as the criterion measure. The entries in Table 3 have been statistically corrected (through partial correlation) to neutralize effects attributable to (a) the context condition to which the individual respondents were assigned, and (b) individual differences in willingness to use the label "schizophrenic," as measured at the very beginning of the experimental session.

The first thing to notice in Table 3 is the fact that all of the entries are positive; while some of these results fail to achieve conventional significance levels when considered separately, within each row, no two correlations are reliably different from one another. Following conventional statistical logic we therefore averaged the obtained correlations across samples, as shown in the last column of the table, emerging with a set of results that were uniformly significant at the .01 level. This analysis provides substantial evidence that judgmental performance in this task is sensitive to a form of base-rate effect; respondents who indicated that

TABLE 3
PARTIAL CORRELATIONS BETWEEN INDUCTION PERFORMANCE AND TEST JUDGMENTS

	Condition			
Test item	Expectation (same hospital)	Same hospital	Same patients	Mean
Definitions	.20	.16	.53*	.31**
Handwriting	.36*	.16	.41*	.31**
All items	.32	.20	.59*	.41**

Note. These correlations control for (a) content of the induction series, and (b) response bias (see text).

  * $p < .05$, two tailed.

  ** $p < .01$, two tailed.

an initial set of patients (the induction set) seemed to include a relatively large proportion of schizophrenics, exhibited a similar judgmental pattern when evaluating the patients in the succeeding test series.

DISCUSSION

Before considering the theoretical significance of these results, let us summarize the main findings of this study, focusing on the relationship between the respondents' induction experiences, their expectations, and their judgments.

Expectations and Prior (Induction) Experience

Subjects who were initially presented with a "high-pathology" series of definitions generally expected that a second (test) sample of patients, drawn from that same hospital, would include more schizophrenics than did subjects who were assigned to a low-pathology induction series. This difference in expectations appeared both in our respondents' "global expectations" for the upcoming test series, and in their trial-by-trial "guesses" regarding the schizophrenic vs nonschizophrenic character of individual patients (although this latter effect proved to be relatively short-lived).

Judgment and Prior (Induction) Experience

Respondents who had initially been presented with an induction series consisting of high-pathology definitions rated subsequent (test) definitions as implying less psychopathology than did those in the low-pathology group—contrast. The high- and low-pathology groups did not, however, differ in their evaluations of handwriting samples. These results constitute a replication of our earlier research (Manis & Paskewitz, 1984).

Expectation and Judgments

Expectations were directly related to subsequent judgments. A variety of analyses revealed a positive relationship between the respondents' expectations and their subsequent judgments. For example, respondents who expected to find substantial psychopathology in the overall test series, or who expected that the next patient in the series would be "schizophrenic," were subsequently more likely to "see" psychopathology than respondents who did not expect psychopathology.

In evaluating the results that relate expectation to subsequent judgments, it is important to recognize that they are manifested in a simple correlational form. This, in turn, raises concern as to the possible influence of artifacts such as response biases, that might somehow have induced the observed correlations. We have approached this issue in two ways.

1. Each respondent's readiness to use the label "schizophrenic" was assessed by simply counting the number of schizophrenic "diagnoses"

she/he emitted in response to a series of midscale definitions, presented at the very beginning of the experimental session. We reasoned that individual differences in the respondents' use of the schizophrenia label might directly affect both (a) their measured expectations, and (b) their subsequent judgments, ultimately producing a positive (but spurious) correlation between these two variables. We therefore resorted to partial correlations to control for individual differences in the schizophrenic bias, and repeatedly found that the linkage between expectation and judgment could *not* be explained by reference to this variable.

2. Another type of artifact that we considered focuses on individual differences in willingness to use the *extremes* of the response scale. Differences in this regard might well affect the extremity of the respondents' judgments (pathological–nonpathological). However, they would seem irrelevant to at least two of our three expectation measures: (a) the number of patients labeled as schizophrenic during the *induction* series and (b) the respondent's successive guesses (schizophrenic versus not schizophrenic) prior to each item of the *test* series. Both of these variables ultimately derive from a simple "yes–no" choice, and hence would be unaffected by extremity biases. Nonetheless, these "two-state" expectation measures showed clearcut relationships between our respondents' expectations and their subsequent judgments, as summarized in Tables 1 and 3. We conclude, therefore, that the linkage between expectation and judgment does not derive (artifactually) from individual differences in the respondents' willingness to use extreme ratings.

Might the correlation between expectation and judgment reflect a simple *consistency* reaction? A respondent who has just guessed that the next patient in the test series is likely to be schizophrenic might be inclined to label that patient's handwriting or definition as "schizophrenic" in the interests of consistency. While this is surely conceivable, we should note that from the respondent's point of view, it is doubtless most important (in an attempt to "look good") that the *appropriate* diagnoses get applied to the various patients, even if this requires a response that is inconsistent with a previously verbalized expectation. Hence we do not believe that consistency motivation, designed to enhance one's status in the eyes of the experimenter, played much of a role here, if any.

CONCLUSION

The results that are summarized above suggest the operation of a two-factor process linking the respondent's assigned induction group (high-vs low-pathology definitions) and his or her evaluation of subsequent definitions. This account is presented schematically in Fig. 3, using path analytic conventions. In essence, our figure suggests that there are two causal paths that derive from the respondent's assigned induction experience.

EXPECTATION

Global Expectation
Trial by Trial Expectation
Induction Performance

+

+

INDUCTION   JUDGEMENT

High vs. Low Pathology   Schizophrenic vs.
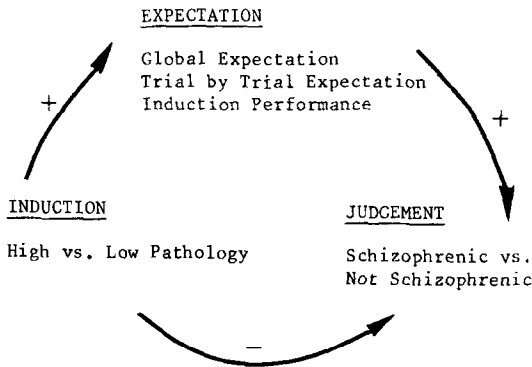   Not Schizophrenic

−

Fig. 3. Schematic representation of two causal paths deriving from respondents' assigned induction experience.

1. One path, which produces its effects independent of expectations, is responsible for the frequently observed negative relationship between prior experience and subsequent judgment. This is the path that mediates the classic contrast effect. We suspect that in many cases the contrast path results from the respondent's spontaneous recognition that the stimulus then being judged is rather different (e.g., more pathological) than other stimuli of that type that she/he has recently evaluated. Thus, when a vocabulary definition is perceived to be more confused than the definitions previously rated in the induction series, respondents may draw the unwarranted conclusion that this most recent (test) definition is pathological in some "absolute" sense (i.e., schizophrenic). We suspect that this type of logical leap is common; respondents may find it quite natural to label a certain definition as extreme ("schizophrenic"), based on the recognition that it is more pathological than the definitions that preceded it. In essence, this inference derives from a confusion between ordinal and absolute judgments, for the fact that a given definition is more pathological than those preceding it does not logically require the inference that the definition in question is pathological in an absolute, "schizophrenic," sense.

2. The second path in Fig. 3 reflects the way in which expectations mediate judgment.[2] This path works in "opposition" to the first path, for it produces an assimilative (or base-rate) pattern of response rather than a contrastive one. More concretely, as shown in Fig. 3, there is

---

[2] Figure 3 implies that (a) global expectations, (b) trial-by-trial expectations (summed over trials), and (c) induction performance provide alternative measures of what our respondents expected to encounter in the test series. Happily, the correlations between these measures were uniformly positive. The product–moment correlation (r) between measures (a) and (b) above was .62; r between measures (b) and (c) was .22; finally, the correlation between measures (a) and (c) yielded an r of .52.

normally a positive relationship between the respondent's initial experience (in induction) and the expectations that she/he harbors with respect to the future; these expectations, in turn, are directly (positively) related to subsequent judgments, as documented in Tables 1–3.

Figure 4 presents this "two-path" model in a somewhat different form, emphasizing the well-known fact that within-group correlations may exhibit a very different pattern than between-group correlations (Messick & van de Geer, 1981). Figure 4 links induction experience, expectations, and judgments for the case in which test items and induction items come from the *same domain;* i.e., both definitions. The first thing to note is that the respondents who were initially assigned to a low-pathology induction experience typically evaluated the test definitions as signifying more psychopathology than did respondents assigned to the high-pathology group. This is the well-known contrast effect, reflected in the *negative* relationship between initial exposure (induction) and subsequent judgment. Note, however, that the respondents' expectations are *positively* related to their prior experience of psychopathology (in induction); i.e., subjects assigned to the high-pathology condition expected to encounter more schizophrenics than those in the low-pathology group. Lastly, there is a positive relationship between our respondents' expectations and their judgments, as reflected in the rising regression lines for both the high-pathology and the low-pathology groups.

In part then, these results testify to the direct influence of expectations upon subsequent judgments. When the impact of "other influences" (like response biases and prior experience) are controlled statistically, judgments roughly parallel expectations. In addition, however, there is clear evidence that prior experiences (in induction) generate contrast effects through a route that is independent of the respondent's stated expectations. The independent existence of these two causal paths suggests that their combined influence in any particular situation will depend upon their relative strengths. It is possible, for example, that the time course of the two paths operates rather differently. Suppose the impact of the contrast
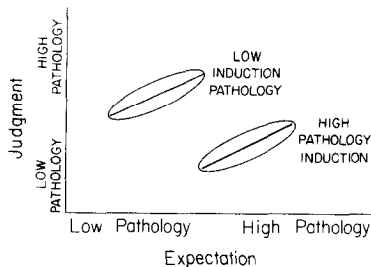

FIG. 4. Schematic representation of the "two-path" model showing that within-group correlations may exhibit a different pattern than between-group correlations.

route declined more rapidly than that of the expectation route. This might lead to a situation in which the net effect of the two paths ultimately yielded assimilation, rather than contrast. Results from previous research in this laboratory (Manis & Blake, 1963; Manis & Moore, 1978) have followed precisely this course. That is, while biased induction experiences initially produced clear evidence of contrast, the introduction of a delay period following induction was repeatedly associated with assimilation.[3] There are, of course, many variables other than the passage of time that might affect the relative potency of these two routes and (as a consequence) the resultant manner in which prior experience affects psychosocial judgments. Theoretical and empirical inquiries might profitably address this issue.

REFERENCES

Bruner, J. (1957). On perceptual readiness. *Psychological Review*, **64**, 123–152.

Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of script pointer and tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, **18**, 319–332.

Graesser, A. C., Woll, S. B., Kowalski, D. J., & Smith, D. A. (1980). Memory of typical and atypical actions in scripted activities. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 503–515.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, **13**, 141–154.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, **80**, 237–251.

Manis, M. (1967). Context effects in communication. *Journal of Personality and Social Psychology*, **5**, 326–334.

Manis, M., & Armstrong, G. W. (1971). Contrast effects in verbal output. *Journal of Experimental Social Psychology*, **7**, 381–388.

Manis, M., & Blake, J. B. (1963). Interpretation of persuasive messages as a function of prior immunization. *Journal of Abnormal and Social Psychology*, **66**, 225–235.

Manis, M., Dovalina, I., Avis, N. E., & Cardoze, S. (1980). Base rates can affect individual predictions. *Journal of Personality and Social Psychology*, **38**, 231–248.

Manis, M., & Moore, J. C. (1978). Summarizing controversial messages: Retroactive effects due to subsequent information. *Social Psychology*, **41**, 62–68.

Manis, M., & Paskewitz, J. R. (1984). Specificity in contrast effects: Judgments of psychopathology. *Journal of Experimental Social Psychology*, **20**, 217–230.

Manis, M., & Paskewitz, J. R. (in press). Judging psychopathology. In R. P. McGlynn, J. E. Maddox, C. D. Stoltenberg, & J. H. Harvey. (Eds.), *Texas Tech Symposium*

---

[3] Graesser and his colleagues (Graesser, Gordon, & Sawyer, 1979; Graesser, Woll, Kowalski, & Smith, 1980) have reported what may be a related phenomenon in an important series of memory studies. Briefly, he has repeatedly found that people initially show superior recall for behavior descriptions that are *incongruent* with their expectations. With the passage of time, however, they show better recall for behaviors that are *congruent* with expectations. In essence, these results suggest the initial dominance of a contrast-like mechanism (memory for the incongruent), followed by an assimilative-like process (superior memory for the congruent).

on *Interfaces in Psychology,* Vol. 2, *Social perception in clinical and counseling psychology.* Lubbock, TX: Texas Tech Univ. Press.

Messick, D. M., & van de Geer, J. P. (1981). A reversal paradox. *Psychological Bulletin,* **90,** 582–593.

Pepitone, A., & DiNubile, M. (1976). Contrast effects in judgments of crime severity and the punishment of criminal violators. *Journal of Personality and Social Psychology,* **33,** 448–459.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw–Hill.