

A PATIENT BY ANY OTHER NAME . . . : CLINICIAN GROUP DIFFERENCE IN LABELING BIAS¹

ELLEN J. LANGER² AND ROBERT P. ABELSON³

Yale University

The effect of labels on clinicians' judgments was assessed in a 2×2 factorial design. Clinicians representing two different schools of thought, behavioral and analytic, viewed a single videotaped interview between a man who had recently applied for a new job and one of the authors. Half of each group was told that the interviewee was a "job applicant," while the remaining half was told that he was a "patient." At the end of the videotape, all clinicians were asked to complete a questionnaire evaluating the interviewee. The interviewee was described as fairly well adjusted by the behavioral therapists regardless of the label supplied. This was not the case, however, for the more traditional therapists. When the interviewee was labeled "patient," he was described as significantly more disturbed than he was when he was labeled "job applicant."

The fact that labels create sets that influence subsequent perception has long been established. Researchers have generally studied these effects by providing different labels and observing the reactions they occasion in their subjects.

Kelley (1950), extending Asch's (1946) work, has shown that by assigning the label warm/cold to a lecturer, one could significantly affect another's perceptions of that person. A more recent study (Huguenard, Sager, & Ferguson, 1970) demonstrated the same result in simulated employment interviews. Along with varying the interviewer's initial set (warm/cold), they also varied the length of the interview (10, 20, or 30 minutes). While the interviewer's initial set significantly affected his after-interview ratings, the length of the interview did not. Thus, the effect of labels is pervasive and not readily overridden by the additional information that may be provided by a prolonged interaction. In another study of this kind (Rapp, 1965), the researcher had pairs of subjects describe a child's behavior. One member of

each pair was informed that the child was feeling "under par," while the other was given the opposite label. Their written descriptions of the child's behavior were significantly different in the predicted direction.

Because of its implication in the institutional and social ostracism of a large group of individuals, one of the labels most widely studied is that of the "mental patient" (Braginsky, Braginsky, & Ring, 1969). In a provocative study, Rosenhan (1972) and some of his colleagues entered psychiatric hospitals as pseudopatients. Upon admission under assumed names, they complained of hearing voices. All of the additional information they supplied was veridical. All but one of these pseudopatients was diagnosed as schizophrenic—the exception was labeled manic-depressive. Right after their admission their symptoms ceased, but they were not immediately discharged. Although trying to behave as sanely as possible in order to obtain release, the initial label was apparently still influential. When they were finally discharged, the diagnosis was schizophrenia "in remission."

Rothaus, Hanson, Cleveland, and Johnson (1963) asked employment placement interviewers to conduct a typical placement interview with a patient. Prior to the interview, they were given forms regarding the patient's background. Half of these forms were couched in problem-centered terms, and half were couched in mental illness terms. Those inter-

¹ This research was conducted while the first author was a National Institute of Mental Health predoctoral fellow (1 FO1 MH 54544-01).

² Requests for reprints should be sent to Ellen J. Langer, Department of Psychology, Yale University, New Haven, Connecticut 06510.

³ The authors are very grateful to those persons who made the necessary arrangements for us to conduct the present study. Our thanks are also extended to Barry Cook for his technical advice.

viewers who were set to view the patient in terms of interpersonal problems gave much more positive after-interview ratings. Temerlin (1968) had psychiatrists, clinical psychologists, and graduate students in clinical psychology diagnose a sound-recorded interview after hearing the interviewee described by a prestigious confederate to be "a very interesting man because he looks neurotic, but actually is quite psychotic." Among the different control groups, one diagnosed the tape without prior suggestions and one diagnosed it with the suggestion reversed. The experimental group rated the interviewee as mentally ill significantly more often than did the controls. In addition, 60% of the psychiatrists diagnosed psychosis.

In most of the above experiments, the investigators provided the different labels and observed the effects on the subjects' behavior, presumably taking precautions to assure that the groups differed only in the label that they were given. While individual differences in the utilization of labels were not examined, it is reasonable to assume that the prior beliefs or attitudes that one brings to the experimental situation will affect the use of labels. In addition, the explicit training that one has received in regard to the use of labels should also be relevant. Clinicians are supposed to be trained to withhold diagnosis until many cues are utilized. However, Temerlin (1968) has shown that clinicians' judgments are also susceptible to labeling bias effects. Since labels make information processing manageable, their use is certainly adaptive. However, there are times, as in the case of the therapeutic interview, where labels may have the deleterious effect of preventing a relatively objective evaluation. It is therefore of interest to know whether explicit training to avoid the use of such labels would be successful in overcoming this tendency.

In the present experiment, it was hypothesized that the therapeutic orientations of clinicians would influence the effect that labels had on their clinical judgments. In particular, a behavioral orientation toward clinical practice typically includes severe skepticism about the utility of diagnostic categories and labels. Yates (1970), for example,

stated:

One important consequence of the application of the medical model to abnormal behavior was the attempt to derive a diagnostic or classificatory system for the pigeonholing of patients. There are at least three serious objections which can be brought against any such system: it is unreliable; it is invalid; and even if it were both reliable and valid, it would serve no useful purpose [p. 5].

If clinicians adopt the ideology associated with their training, it would seem likely that behavior therapists would tend not to display labeling effects characterizing the judgments of clinicians who had received more traditional training.

Two groups of therapists with appropriately contrasting training were shown a videotaped interview with a man who had recently applied for a new job. Before viewing the tape, half of the subjects were told that the interviewee was a job applicant, and the other half were told that he was a patient. It was predicted that (a) when the interviewee was labeled "patient" he would in general be perceived as a more disturbed individual than when he was labeled "job applicant," and (b) this labeling bias would be less for the behavior therapists than for the traditional therapists.

METHOD

Subjects

Forty clinicians associated with university departments known to be either behaviorally or psychodynamically oriented served as subjects. These clinicians were either graduate or postdoctoral clinical students, residents, or faculty members. Twenty-one clinicians from the State University of New York at Stony Brook represented the behavioral bent, while 9 clinicians affiliated with New York University and 10 clinicians affiliated with the School of Psychiatry at Yale University represented the more traditional view. The institutional differences in orientation are clear. The descriptive printed handout given to all Stony Brook applicants states that

the program . . . finds expression in the behavioral point of view toward Clinical Psychology, in general, and the behavior modification approach to therapy in particular. Students seeking training in other orientations would be ill-advised to enter our program.

By contrast, the comparable New York University document makes no mention whatever of the terms "behavioral" or "behavior modification" and gives

as a prime objective of the training program, "To familiarize the student with the theories and practice of dynamic psychotherapy." The *Yale Psychiatry* (not Psychology) *Course Bulletin* likewise makes no mention of behavior therapy, and in rather eclectic terms speaks of "personality dynamics" and the "treatment of [the] ill person." The question of the conformance of individual subjects to the dominant orientations of their institutions is addressed later. Subjects were randomly assigned to one of two conditions. Ten therapists from Stony Brook, 5 from New York University, and 5 from Yale were in the "patient" condition. The remainder were in the "job applicant" condition. The mean age was 28.47 years for the behavior therapists and 29.21 years for the analytic therapists.

Materials

The subjects were recruited for a study designed to evaluate a videotaped interview. The videotape was of an interview by a bearded professor (the second author) of a younger man of about 26 years. The young man was one of several individuals recruited through a newspaper ad offering \$10 to someone who had recently applied for a new job and was willing to be interviewed. The interview itself was unstructured, but it centered around the interviewees' feelings and experience relating to his past work. Since clinical interviews very often concern the patient's reactions to or interpretations of his life situation, his occupation in particular (Howard, Orlinsky, & Hill, 1969), it seemed likely that a work-oriented interview between a patient and therapist would not be seen as unusual.

Our stated interest was in the evaluation of the young man, and not in the nature of the interview probes. Therefore, the interviewer's voice was eliminated from the videotape as much as possible. The tape was cut down to the most interesting 15-minute segment, and the result was an authentic, rambling, and autobiographical monologue by the young man describing a number of jobs he had held and dwelling particularly on his conflicts with bureaucratic authorities in his job as a youth worker and during an abortive business enterprise. The life situations described were complex and ambiguous, and the man's style intense, but uncertain, so that he could easily be seen either as sincere and struggling or as confused and troubled.

Procedure

The subjects entered a room in the school with which they were affiliated, were seated, and were read one of the following sets of instructions:

[*Job applicant condition.*] Thank you for coming. Lately I've been studying job interviewing. All that I'm going to ask you to do is to view part of a videotaped interview with a man who has recently applied for a new job and then fill out a short questionnaire evaluating the job applicant. The interviewer's voice has been elimi-

nated as much as possible so as not to distract you from focusing on the job applicant.

[*Patient condition.*] Thank you for coming. Lately I've been studying patient interviewing. All that I'm going to ask you to do is to view part of a videotaped interview with a patient and then fill out a short questionnaire evaluating the patient. The interviewer's voice has been eliminated as much as possible so as not to distract you from focusing on the patient.

The label assigned to the interviewee constituted the independent variable. The experimenter made the label salient once again at the end of the tape: "Here is the job applicant [patient] evaluation form."

Dependent Measure

The evaluation form was a questionnaire asking for a brief free-response description of the interviewee, his gestures, attitudes, and the factors that probably explained his outlook on life. It also asked what kind of job they would recommend for him. Open-ended rather than multiple-choice questions were used so as to make the task most natural and congenial to the clinicians.

These descriptive replies were later quantified by having five graduate student raters, blind to the experimental hypotheses and conditions, rate each of the 40 randomly ordered clinicians' questionnaires on a scale from 1 (very disturbed) to 10 (very well-adjusted) for the clinician's beliefs about the interviewee. The mean interrater correlation over the 40 judgments was .76. The five ratings were averaged to yield a mean adjustment rating for each clinician questionnaire. By Spearman-Brown formula, the reliability of this mean adjustment rating is .94.

Clinician Groupings

To make sure that the therapists affiliated with the different schools indeed held different theoretical orientations, a biographical questionnaire was administered after the completion of the session. Along with questions about the amount of clinical experience they had had to date, the clinicians were asked the label that they would give to the kind of psychotherapy they themselves practiced. All 21 Stony Brook clinicians called themselves behavior therapists, and 17 of the 19 subjects from traditional programs chose an analytic label. (Two from Yale said they practiced behavior therapy as well.)

In addition, the clinicians were asked how strongly they agreed or disagreed with four statements that touch issues of disagreement between schools of therapy. It was presumed that the majority of behavior therapists would agree with the first statement and disagree with the rest,⁴ while the reverse

⁴ As it turned out, Item 2 on childhood experiences was by far the most discriminating of the four items, although all yielded results in the anticipated direction.

would tend to be true for the analytic therapists: (a) If you have cured the symptom you have usually solved the problem; (b) the examination of childhood experience is essential to effective psychotherapy; (c) The official APA Diagnostic Nomenclature for Psychiatric Disorders is helpful to both patient and clinician; (d) Most people need some kind of psychotherapeutic help.

By scoring each item on a scale from 0 to 4 (with Item 1 reversed), a four-item Likert scale score from 0 to 16 was assigned to each clinician. The median score for the 40 subjects was 4.5. Sixteen of the 21 Stony Brook subjects had scores below the median, whereas 15 of the 19 presumed traditionalists had scores above the median. The four exceptions were all in the Yale group, which had a score distribution intermediate between the "classical" New York University group and the "radical" Stony Brook group.

These large group differences in individual attitudes and self-designations (albeit with a few stray cases), combined with the clear institutional differences in theoretical orientation, provide assurance that a group comparison in assessment of the interviewee is a meaningful test of our second hypothesis.

RESULTS

The mean adjustment ratings for all cells are shown in Table 1. The New York University and Yale subgroups of traditional therapists are shown separately since it is of interest to know to what extent they were differentially susceptible to a labeling effect.

One evident tendency in the table is that when the interviewee is labeled job applicant, there is not much difference in mean adjustment ratings by clinicians at the three schools. However, the patient label seems to produce sharp differential effects. For both traditional groups, the means are on the "disturbed" side of the midpoint of the 1-10 adjustment scale: the Yale mean of 4.80 slightly so and the New York University mean of 2.40 very much so.

The complete analysis of variance of the data for Table 1 is shown in Table 2. The 5 degrees of freedom between cells are decomposed into five single degrees of freedom contrasts—one for label, one for the main effect of training (behavioral vs. traditional), one for subgroup within traditional (Yale vs. New York University), one for the Label \times Training interaction, and one for the Label \times Subgroup interaction. From the standpoint of our hypotheses, the crucial contrast is the Label \times Training interaction: Do the tradi-

TABLE 1
MEAN ADJUSTMENT RATING BY INTERVIEWEE LABEL
AND CLINICIAN GROUP

Clinical group	Interviewee label	
	Job applicant	Patient
Behavior therapist		
Stony Brook	6.26	5.98
<i>n</i>	10	11
Traditional		
Yale	6.52	4.80
<i>n</i>	5	4
NYU	5.88	2.40
<i>n</i>	5	5

tional clinicians generate a significantly bigger adjustment difference between job applicant and patient than do the behavioral clinicians? The answer is yes ($F = 4.75$, $p < .05$). The differential in labeling effect between Yale and New York University subgroups of traditionalists is, on the other hand, not significant ($F = 1.29$, $p > .25$).

A slightly different way of looking at labeling bias is to take the difference in mean adjustment ratings for job applicant and patient in each of the three schools in the study (Stony Brook, .28; Yale, 1.72; New York University, 3.48). It is possible, taking appropriate account of the different *N*s in the six cells, to apply a Tukey multiple-comparisons test (Winer, 1962) to these differences. The outcome of this test is that the New York University subgroup was significantly more susceptible to labeling bias than the Stony Brook group ($q = 3.51$, $df =$

TABLE 2
ANALYSIS OF VARIANCE OF MEAN
ADJUSTMENT RATINGS

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Label (A)	1	28.85	10.26***
Clinicians (B)			
Behavioral vs. traditional (B ₁)	1	14.76	5.25**
Subgroup of traditionals (B ₂)	1	10.87	3.87*
A \times B			
A \times B ₁	1	13.34	4.75**
A \times B ₂	1	3.64	1.29
Within cells	34	2.81	

* $p < .10$.
** $p < .05$.
*** $p < .01$.

3/34, $p < .05$). The Yale subgroup, intermediate to the other two, cannot be declared significantly different from either on these small sample sizes.

DISCUSSION

There is far too much information surrounding any situation for any individual to process more than a small fraction of it. Labels provide one vehicle through which the input may be organized. They serve as categories or sets that, in addition to structuring the previous input, determine what further information is attended to. Thus, by assigning different labels (in this case "job applicant" or "patient"), different people may be led to view this same event in vastly disparate ways.

In the study just described, all of the subjects saw the same videotaped interview. Yet when asked to describe the interviewee, the behavior therapists said he was "realistic"; "unassertive"; "fairly sincere, enthusiastic, attractive appearance"; "pleasant, easy manner of speaking"; "relatively bright, but unable to assert himself"; "appeared responsible in interview." The analytic therapists who saw a job applicant called him "attractive and conventional looking"; "candid and innovative"; "ordinary, straightforward"; "upstanding, middle-class-citizen type, but more like a hard hat"; "probably of lower or blue-collar class origins"; "middle-class protestant ethic orientation; fairly open—somewhat ingenious." The analytic therapists that saw a patient described him as a "tight, defensive person . . . conflict over homosexuality"; "dependent, passive-aggressive"; "frightened of his own aggressive impulses"; "fairly bright, but tries to seem brighter than he is . . . impulsivity shows through his rigidity"; "passive, dependent type"; "considerable hostility, repressed or channeled."

The fact that the different labels set the analytic therapists to look for very different behaviors may be further exemplified by reviewing typical responses to the question "What do you think might explain Mr. Smith's outlook on life? Do you think he is realistic?"

Analytic therapists viewing a patient said: "Doesn't seem to be realistic because he

seems to use denial (and rationalization and intellectualization) to center his problems in situations and other people," "seems afraid of his own drives, motives . . . outlook not based on realities of 'objective world,'" "anxiety about his ability and adequacy," "basically fear of his aggressive drives and in particular as they are related to his fear of women."

Those analytic therapists viewing a job applicant said: "His attitudes are consistent with a large subculture in the U.S. . . . the silent majority"; "he seems fairly realistic," "fairly reality oriented; recognizes injustices of large systems but doesn't seem to think he can individually do anything to change them"; "realistic to some degree, he knows how to conform but finds it difficult"; "he seems to be perceptive and realistic about politicians"; "values capitalist system."

Behavior therapists given either label respond very much like the latter group of analytic therapists: "His previous experience working in bureaucratic organizations might account for his distrust of authority. . . . He is probably realistic"; "his desire to be a successful businessman may have been partly a function of the business orientation of his friends and family"; "his negative attitudes probably result from the frustrations of working in backward correctional or educational institutions"; "he seems fairly realistic and apparently wants to do something to help the kids he's working with"; "his pessimism is realistic"; "don't know what his outlook on life is, except that he thinks people should be more involved in their work, and that is realistic."

The present research does not tell us why the behavior therapists were apparently immune to the biasing effects of the mere label patient versus job applicant. Perhaps they tended to focus so heavily on the manifest behaviors in the interview that they barely even attended to background information such as labels; more likely, they actively noted the label patient but consciously discounted its relevance because their training explicitly encourages such discounting. Likewise, we do not know how the analytic therapists succumbed to the labeling effect, whether by using the patient label to filter

differentially the variety of cues on the tape or simply by superimposing their general concept of a sick person on the particular concept of this person gleaned from the tape. Whichever the case, it is interesting that the New York University group, trained (or self-selected) to subscribe more completely to the classical doctrine of mental illness than the Yale group, is more extreme in its susceptibility to the labeling effect. Our results are apparently sensitive to the particulars of different university training programs, although our sample of three programs is obviously quite limited.

In practical terms, the labeling bias may have unfortunate consequences whatever the specific details of its operation. Once an individual enters a therapist's office for consultation, he has labeled himself "patient." From the very start of the session, the orientation of the conversation may be quite negative. The patient discusses all the negative things he said, did, thought, and felt. The therapist then discusses or thinks about what is wrong with the patient's behavior, cognitions and feelings. The therapist's negative expectations in turn may affect the patient's view of his own difficulties, thereby possibly locking the interaction into a self-fulfilling gloomy prophecy. As in the study presented here, if the therapist were not given the label patient, he would see a very different range of behaviors or attribute the given behaviors to factors other than the patient's "illness." He might, for example, attribute the loss of a prior job by the interviewee to economic conditions on a national or state level, rather than to the interviewee's emotional problems. This factor was indeed often taken into account when "Mr. Smith" was described as a job applicant rather than as a patient.

Another way of viewing the present results, of course, is that the person on the tape did indeed have deep underlying problems to which the behavior therapists were not sensitive and which the traditional therapists only looked for when there was good reason, that is, when the individual was presumably a patient. Since the person being

discussed was able to cope with his environment and was in fact not a patient, this alternative seems to the present authors not too satisfactory.

Despite the questionable light in which the analytic therapist group was cast in the present study, one strongly suspects that conditions might be arranged wherein the behavior therapists would fall into some kind of error, as much as the traditionalists. No single type of orientation toward clinical training is likely to avoid all types of biases or blind spots. In any case, all we can claim to have shown is that the behavior therapists avoid the particular kind of bias in which the superficial cue "patient" produces drastically negative interpretations, even when an extended visual and verbal segment filled with personal cues is available.

REFERENCES

- ASCH, S. E. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 1946, 41, 258-290.
- BRAGINSKY, B., BRAGINSKY, D., & RING, K. L. *Methods of madness: The mental hospital as a last resort*. New York: Holt, Rinehart & Winston, 1969.
- KELLEY, H. H. The warm-cold variable in first impressions of persons. *Journal of Personality*, 1950, 18, 431-439.
- HOWARD, K. I., ORLINSKY, D. E., & HILL, J. A. Context of dialogue in psychotherapy. *Journal of Counseling Psychology*, 1969, 16, 396-404.
- HUGUENARD, T., SAGER, E. B., & FERGUSON, L. W. Interview time, interview set, and interview outcome. *Perceptual and Motor Skills*, 1970, 31, 831-836.
- RAPP, D. W. Detection of observer bias in the written record. Unpublished manuscript, University of Georgia, 1965. Cited in R. Rosenthal, *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966, p. 21.
- ROSENHAN, D. L. On being sane in insane places. *Science*, 1973, 179, 250-257.
- ROTHAUS, P., HANSON, P. G., CLEVELAND, S. E., & JOHNSON, D. L. Describing hospitalization: A dilemma. *American Psychologist*, 1963, 18, 85-89.
- TEMERLIN, M. K. Suggestion effects in psychiatric diagnosis. *Journal of Nervous and Mental Disease*, 1968, 147, 349-359.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.
- YATES, A. J. *Behavior therapy*. New York: Wiley, 1970.

(Received May 21, 1973)