



## PROCESSES OF OPINION CHANGE\*

BY HERBERT C. KELMAN

Attitude and opinion data provide a basis for inferring the meaning of opinions held by individuals and groups and also for predictions about their future behavior. Such inferences and predictions, if they are to be made effectively, require a theoretical foundation which explains the processes by which people adopt and express particular opinions. Here is a theory of three processes by which persons respond to social influence.

Herbert C. Kelman is Lecturer on Social Psychology at Harvard University. Currently, he is spending the year at the Institute for Social Research in Oslo and devoting part of his time to a further study of the effects of a year's sojourn in the United States on the self-images of Scandinavian students. His theoretical and experimental work during the past eight years on the problems reported in this article will be published in a forthcoming book, an early draft of which was awarded the Socio-Psychological Prize of the American Association for the Advancement of Science in 1958.

A PERSISTENT concern in the analysis of public opinion data is the "meaning" that one can ascribe to the observed distributions and trends—and to the positions taken by particular individuals and segments of the population. Clearly, to understand what opinion data mean we have to know considerably more than the direction of an individual's responses or the distribution of responses in the population. We need information that will allow us to make some inferences about the characteristics of the observed opinions—their intensity, their salience, the level of commitment that they imply. We need information about the motivational bases of these opinions—about the functions that they fulfill for the individual and the motivational systems in which they are embedded.<sup>1</sup> We need information about the cognitive links of the opinions—the amount and the nature of information that supports them, the specific expectations and evaluations that surround them.

The need for more detailed information becomes even more apparent when we attempt to use opinion data for the prediction of subsequent behavior. What is the likelihood that the opinions observed in a particular survey will be translated into some form of concrete action?

\* This paper is based on a research program on social influence and behavior change, supported by grant M-2516 from the National Institute of Mental Health.

<sup>1</sup> For discussions of the different motivational bases of opinion see I. Sarinoff and D. Katz, "The Motivational Bases of Attitude Change," *Journal of Abnormal and Social Psychology*, Vol. 49, 1954, pp. 115-124; and M. B. Smith, J. S. Brainer, and R. W. White, *Opinions and Personality*, New York, Wiley, 1955.

What is the nature of the actions that people who hold a particular opinion are likely to take, and how are they likely to react to various events? How likely are these opinions to persist over time and to generalize to related issues? What are the conditions under which one might expect these opinions to be abandoned and changed? Such predictions can be made only to the extent to which we are informed about the crucial dimensions of the opinions in question, about the motivations that underlie them, and about the cognitive contexts in which they are held.

#### INFERRING THE MEANING OF OPINIONS

In a certain sense, the need for more detailed information about opinions can (and must) be met by improvements and refinements in the methodology of opinion assessment. A great deal of progress in this direction has already been made in recent years. Thus, many widely accepted features of interviewing technique are specifically designed to elicit information on which valid inferences about the meaning of opinions can be based: the creation of a relaxed, nonjudgmental atmosphere; the emphasis on open-ended questions; the progressive funneling from general to specific questions; the use of probes, of indirect questions, and of interlocking questions; and so on. These procedures facilitate inferences (1) by maximizing the likelihood that the respondent will give rich and full information and thus reveal the motivational and cognitive structure underlying the expressed opinions, and (2) by minimizing the likelihood that the respondent will consciously or unconsciously distort his "private" opinions when expressing them to the interviewer.

Similarly, when attitudes are assessed by means of questionnaires, it is possible to approximate these methodological goals. In part, this is accomplished by the instructions, which can motivate the subject to respond fully and honestly and assure him of confidentiality or anonymity. In part it is accomplished by the use of indirect and projective questions, and by the inclusion of a series of interrelated items in the questionnaire. And, in part, it is possible to make inferences about the meaning of opinions by the use of various scaling devices in the analysis of the data.

There is no question about the importance of these methodological advances, but in and of themselves they do not solve the problem of inference. They increase the investigator's ability to obtain rich and relatively undistorted information on which he can then base valid inferences. But, no matter how refined the techniques, they do not provide direct information about the meaning of the opinions and do

not permit automatic predictions to subsequent behavior: the investigator still has to make inferences from the data.

To make such inferences, the student of public opinion needs a theoretical framework which accounts for the adoption and expression of particular opinions on the part of individuals and groups. Such a framework can serve as a guide in the collection of data: it can provide a systematic basis for deciding what information is relevant and what questions should be asked in order to permit the drawing of inferences. Similarly, it can serve as a guide for interpreting the data and deriving implications from them.

The need for such a framework is particularly apparent when one attempts to make predictions about subsequent behavior on the basis of opinion data. For example, in a relaxed interview situation a particular respondent may express himself favorably toward socialized medicine. What are the chances that he will take the same position in a variety of other situations? To answer this, we would need a theoretical scheme for the analysis of interaction situations, in terms of which we could make some inferences about the structure and meaning of this particular interview situation as compared to various other situations in which the issue of socialized medicine might arise. How would we expect this same respondent to react to a concerted campaign by the Medical Association which links Federal insurance programs with creeping socialism? To answer this, we would need a theory of opinion formation and change, in terms of which we could make some inferences about the characteristics of opinions formed under different conditions.

Progress in the analysis of public opinion, then, requires theoretical development along with methodological improvements. For this development, it should be possible to draw on some of the current theoretical thinking and associated research in social psychology. There are two foci of social-psychological theorizing and research that would appear to be particularly germane to the analysis of public opinion. One is the study of processes of social interaction as such. Such diverse approaches to the analysis of social interaction as those of Getzels,<sup>2</sup> Goffman,<sup>3</sup> and Jones and Thibaut,<sup>4</sup> for example, can be useful for con-

<sup>2</sup> J. W. Getzels, "The Question-Answer Process: A Conceptualization and Some Derived Hypotheses for Empirical Examination," *Public Opinion Quarterly*, Vol. 18, 1954, pp. 80-91.

<sup>3</sup> See, for example, E. Goffman, "On Face-work: An Analysis of Ritual Elements in Social Interaction," *Psychiatry*, Vol. 18, 1955, pp. 213-231; and "Alienation from Interaction," *Human Relations*, Vol. 10, 1957, pp. 47-60.

<sup>4</sup> E. E. Jones and J. W. Thibaut, "Interaction Goals as Bases of Inference in Interpersonal Perception," in R. Tagiuri and L. Petrullo, editors, *Person Perception and Interpersonal Behavior*, Stanford, Calif., Stanford University Press, 1958, pp. 151-178.

ceptualizing the determinants of *opinion expression*. Thus, by using one or another of these schemes, the investigator can make some formulations about the expectations that the respondent brought to the interview situation and the goals that he was trying to achieve in this interaction. On the basis of such a formulation, he can make inferences about the meaning of the opinions expressed in this situation and about their implications for subsequent behavior—for example, about the likelihood that similar opinions will be expressed in a variety of other situations.

The second relevant focus of social-psychological theorizing and research is the study of processes of social influence and the induction of behavior change. Theoretical analyses in this area can be useful for conceptualizing the determinants of *opinion formation* and *opinion change*. They can help the investigator in making formulations about the sources of the opinions expressed by the respondent—the social conditions under which they were adopted, the motivations that underlie them, and the social and personal systems in which they are embedded. On the basis of such a formulation, again, he can make inferences about the meaning and implications of the opinions ascertained.

The model that I shall present here emerged out of the second research focus—the study of social influence and behavior change. It is, essentially, an attempt to conceptualize the processes of opinion formation and opinion change. It starts with the assumption that opinions adopted under different conditions of social influence, and based on different motivations, will differ in terms of their qualitative characteristics and their subsequent histories. Thus, if we know something about the determinants and motivational bases of particular opinions, we should be able to make predictions about the conditions under which they are likely to be expressed, the conditions under which they are likely to change, and other behavioral consequences to which they are likely to lead. Ideally, such a model can be useful in the analysis of public opinion by suggesting relevant variables in terms of which opinion data can be examined and predictions can be formulated.

#### THE STUDY OF SOCIAL INFLUENCE

Social influence has been a central area of concern for experimental social psychology almost since its beginnings. Three general research traditions in this area can be distinguished: (1) the study of social influences on judgments, stemming from the earlier work on prestige suggestion;<sup>2</sup> (2) the study of social influences arising from small-group

<sup>2</sup> See, for example, S. E. Asch, *Social Psychology*, New York, Prentice-Hall, 1952.

interaction;<sup>6</sup> and (3) the study of social influences arising from persuasive communications.<sup>7</sup> In recent years, there has been a considerable convergence between these three traditions, going hand in hand with an increased interest in developing general principles of social influence and socially induced behavior change.

One result of these developments has been that many investigators found it necessary to make qualitative distinctions between different types of influence. In some cases, these distinctions arose primarily out of the observation that social influence may have qualitatively different effects, that it may produce different kinds of change. For example, under some conditions it may result in mere public conformity—in superficial changes on a verbal or overt level without accompanying changes in belief; in other situations it may result in private acceptance—in a change that is more general, more durable, more integrated with the person's own values.<sup>8</sup> Other investigators found it necessary to make distinctions because they observed that influence may occur for different reasons, that it may arise out of different motivations and orientations. For example, under some conditions influence may be primarily informational—the subject may conform to the influencing person or group because he views him as a source of valid information; in other situations influence may be primarily normative—the subject may conform in order to meet the positive expectations of the influencing person or group.<sup>9</sup>

My own work can be viewed in the general context that I have outlined here. I started out with the distinction between public conformity and private acceptance, and tried to establish some of the distinct determinants of each. I became dissatisfied with this dichotomy as I began to look at important examples of social influence that could

<sup>6</sup> See, for example, D. Cartwright and A. Zander, editors, *Group Dynamics*, Evanston, Ill., Row, Peterson, 1953.

<sup>7</sup> See, for example, C. I. Hovland, I. L. Janis, and H. H. Kelley, *Communication and Persuasion*, New Haven, Yale University Press, 1953.

<sup>8</sup> See, for example, L. Festinger, "An Analysis of Compliant Behavior," in M. Sherif and M. O. Wilson, editors, *Group Relations at the Crossroads*, New York, Harper, 1953, pp. 232-256; H. C. Kelman, "Attitude Change as a Function of Response Restriction," *Human Relations*, Vol. 6, 1953, pp. 185-214; J. R. P. French, Jr., and B. Raven, "The Bases of Social Power," in D. Cartwright, editor, *Studies in Social Power*, Ann Arbor, Mich., Institute for Social Research, 1959, pp. 150-167; and Marie Jahoda, "Conformity and Independence," *Human Relations*, Vol. 12, 1959, pp. 99-120.

<sup>9</sup> See, for example, M. Deutsch and H. B. Gerard, "A Study of Normative and Informational Social Influence upon Individual Judgment," *Journal of Abnormal and Social Psychology*, Vol. 51, 1955, pp. 629-636; J. W. Thibaut and L. Strickland, "Psychological Set and Social Conformity," *Journal of Personality*, Vol. 25, 1956, pp. 115-129; and J. M. Jackson and H. D. Saltzstein, "The Effect of Person-Group Relationships on Conformity Processes," *Journal of Abnormal and Social Psychology*, Vol. 57, 1958, pp. 17-24.

not be encompassed by it. I was especially impressed with the accounts of ideological conversion of the "true believer" variety, and with the recent accounts of "brainwashing," particularly the Chinese Communist methods of "thought reform."<sup>10</sup> It is apparent that these experiences do not simply involve public conformity, but that indeed they produce a change in underlying beliefs. But it is equally apparent that they do not produce what we would usually consider private acceptance—changes that are in some sense integrated with the person's own value system and that have become independent of the external source. Rather, they seem to produce new beliefs that are isolated from the rest of the person's values and that are highly dependent on external support.

These considerations eventually led me to distinguish three processes of social influence, each characterized by a distinct set of antecedent and a distinct set of consequent conditions. I have called these processes *compliance*, *identification*, and *internalization*.<sup>11</sup>

### THREE PROCESSES OF SOCIAL INFLUENCE

*Compliance* can be said to occur when an individual accepts influence from another person or from a group because he hopes to achieve a favorable reaction from the other. He may be interested in attaining certain specific rewards or in avoiding certain specific punishments that the influencing agent controls. For example, an individual may make a special effort to express only "correct" opinions in order to gain admission into a particular group or social set, or in order to avoid being fired from his government job. Or, the individual may be concerned with gaining approval or avoiding disapproval from the influencing agent in a more general way. For example, some individuals may compulsively try to say the expected thing in all situations and please everyone with whom they come in contact, out of a disproportionate need for favorable responses from others of a direct and immediate kind. In any event, when the individual complies, he does what the agent wants him to do—or what he thinks the agent wants him to do—because he sees this as a way of achieving a desired response from him. He does not adopt the induced behavior—for example, a particular opinion response—because he believes in its content, but because it is instrumental in the production of a satisfying social effect. What the individual learns, essentially, is to say or do the expected thing in special situations, regardless of what his private beliefs may

<sup>10</sup> For instance, R. J. Lifton, "'Thought Reform' of Western Civilians in Chinese Communist Prisons," *Psychiatry*, Vol. 19, 1956, pp. 173-195.

<sup>11</sup> A detailed description of these processes and the experimental work based on them will be contained in a forthcoming book, *Social Influence and Personal Belief: A Theoretical and Experimental Approach to the Study of Behavior Change*, to be published by John Wiley & Sons.

be. Opinions adopted through compliance should be expressed only when the person's behavior is observable by the influencing agent.

*Identification* can be said to occur when an individual adopts behavior derived from another person or a group because this behavior is associated with a satisfying self-defining relationship to this person or group. By a self-defining relationship I mean a role relationship that forms a part of the person's self-image. Accepting influence through identification, then, is a way of establishing or maintaining the desired relationship to the other, and the self-definition that is anchored in this relationship.

The relationship that an individual tries to establish or maintain through identification may take different forms. It may take the form of classical identification, that is, of a relationship in which the individual takes over all or part of the role of the influencing agent. To the extent to which such a relationship exists, the individual defines his own role in terms of the role of the other. He attempts to be like or actually to *be* the other person. By saying what the other says, doing what he does, believing what he believes, the individual maintains this relationship and the satisfying self-definition that it provides him. An influencing agent who is likely to be an attractive object for such a relationship is one who occupies a role desired by the individual—who ~~possesses~~ those characteristics that the individual himself lacks—such as control in a situation in which the individual is helpless, direction in a situation in which he is disoriented, or belongingness in a situation in which he is isolated.

The behavior of the brainwashed prisoner in Communist China provides one example of this type of identification. By adopting the attitudes and beliefs of the prison authorities—including *their* evaluation of *him*—he attempts to regain his identity, which has been subjected to severe threats. But this kind of identification does not occur only in such severe crisis situations. It can also be observed, for example, in the context of socialization of children, where the taking over of parental attitudes and actions is a normal, and probably essential, part of personality development. The more or less conscious efforts involved when an individual learns to play a desired occupational role and imitates an appropriate role model would also exemplify this process. Here, of course, the individual is much more selective in the attitudes and actions he takes over from the other person. What is at stake is not his basic sense of identity or the stability of his self-concept, but rather his more limited "professional identity."

The self-defining relationship that an individual tries to establish or maintain through identification may also take the form of a reciprocal role relationship—that is, of a relationship in which the roles of the two

parties are defined with reference to one another. An individual may be involved in a reciprocal relationship with another specific individual, as in a friendship relationship between two people. Or he may enact a social role which is defined with reference to another (reciprocal) role, as in the relationship between patient and doctor. A reciprocal-role relationship can be maintained only if the participants have mutually shared expectations of one another's behavior. Thus, if an individual finds a particular relationship satisfying, he will tend to behave in such a way as to meet the expectations of the other. In other words, he will tend to behave in line with the requirements of this particular relationship. This should be true regardless of whether the other is watching or not: quite apart from the reactions of the other, it is important to the individual's own self-concept to meet the expectations of his friendship role, for example, or those of his occupational role.

Thus, the acceptance of influence through identification should take place when the person sees the induced behavior as relevant to and required by a reciprocal-role relationship in which he is a participant. Acceptance of influence based on a reciprocal-role relationship is similar to that involved in classical identification in that it is a way of establishing or maintaining a satisfying self-defining relationship to another. The nature of the relationship differs, of course. In one case it is a relationship of identity; in the other, one of reciprocity. In the case of reciprocal-role relationships, the individual is not identifying with the other in the sense of taking over *his* identity, but in the sense of empathically reacting in terms of the other person's expectations, feelings, or needs.

Identification may also serve to maintain an individual's relationship to a group in which his self-definition is anchored. Such a relationship may have elements of classical identification as well as of reciprocal roles: to maintain his self-definition as a group member an individual, typically, has to model his behavior along particular lines and has to meet the expectations of his fellow members. An example of identification with a group would be the member of the Communist Party who derives strength and a sense of identity from his self-definition as part of the vanguard of the proletarian revolution and as an agent of historical destiny. A similar process, but at a low degree of intensity, is probably involved in many of the conventions that people acquire as part of their socialization into a particular group.

Identification is similar to compliance in that the individual does not adopt the induced behavior because its content per se is intrinsically satisfying. Identification differs from compliance, however, in that the individual actually believes in the opinions and actions that he adopts. The behavior is accepted both publicly and privately, and its mani-



festation does not depend on observability by the influencing agent. It does depend, however, on the role that an individual takes at any given moment in time. Only when the appropriate role is activated—only when the individual is acting within the relationship upon which the identification is based—will the induced opinions be expressed. The individual is not primarily concerned with pleasing the other, with giving him what he wants (as in compliance), but he is concerned with meeting the other's expectations for his own role performance. Thus, opinions adopted through identification do remain tied to the external source and dependent on social support. They are not integrated with the individual's value system, but rather tend to be isolated from the rest of his values—to remain encapsulated.

Finally, *internalization* can be said to occur when an individual accepts influence because the induced behavior is congruent with his value system. It is the content of the induced behavior that is intrinsically rewarding here. The individual adopts it because he finds it useful for the solution of a problem, or because it is congenial to his own orientation, or because it is demanded by his own values—in short, because he perceives it as inherently conducive to the maximization of his values. The characteristics of the influencing agent do play an important role in internalization, but the crucial dimension here—as we shall see below—is the agent's credibility, that is, his relation to the content.

The most obvious examples of internalization are those that involve the evaluation and acceptance of induced behavior on rational grounds. A person may adopt the recommendations of an expert, for example, because he finds them relevant to his own problems and congruent with his own values. Typically, when internalization is involved, he will not accept these recommendations *in toto* but modify them to some degree so that they will fit his own unique situation. Or a visitor to a foreign country may be challenged by the different patterns of behavior to which he is exposed, and he may decide to adopt them (again, selectively and in modified form) because he finds them more in keeping with his own values than the patterns in his home country. I am not implying, of course, that internalization is always involved in the situations mentioned. One would speak of internalization only if acceptance of influence took the particular form that I described.

Internalization, however, does not necessarily involve the adoption of induced behavior on rational grounds. I would not want to equate internalization with rationality, even though the description of the process has decidedly rationalist overtones. For example, I would characterize as internalization the adoption of beliefs because of their congruence with a value system that is basically irrational. Thus, an authori-

tarian individual may adopt certain racist attitudes because they fit into his paranoid, irrational view of the world. Presumably, what is involved here is internalization, since it is the content of the induced behavior and its relation to the person's value system that is satisfying. Similarly, it should be noted that congruence with a person's value system does not necessarily imply logical consistency. Behavior would be congruent if, in some way or other, it fit into the person's value system, if it seemed to belong there and be demanded by it.

It follows from this conception that behavior adopted through internalization is in some way—rational or otherwise—integrated with the individual's existing values. It becomes part of a personal system, as distinguished from a system of social-role expectations. Such behavior gradually becomes independent of the external source. Its manifestation depends neither on observability by the influencing agent nor on the activation of the relevant role, but on the extent to which the underlying values have been made relevant by the issues under consideration. This does not mean that the individual will invariably express internalized opinions, regardless of the social situation. In any specific situation, he has to choose among competing values in the face of a variety of situational requirements. It does mean, however, that these opinions will at least enter into competition with other alternatives whenever they are relevant in content.

It should be stressed that the three processes are not mutually exclusive. While they have been defined in terms of pure cases, they do not generally occur in pure form in real-life situations. The examples that have been given are, at best, situations in which a particular process predominates and determines the central features of the interaction.

#### ANTECEDENTS AND CONSEQUENTS OF THE THREE PROCESSES

For each of the three processes, a distinct set of antecedents and a distinct set of consequents have been proposed. These are summarized in the table below. First, with respect to the antecedents of the three processes, it should be noted that no systematic quantitative differences between them are hypothesized. The probability of each process is presented as a function of the same three determinants: the importance of the induction for the individual's goal achievement, the power of the influencing agent, and the prepotency of the induced response. For each process, the magnitude of these determinants may vary over the entire range: each may be based on an induction with varying degrees of importance, on an influencing agent with varying degrees of power, and so on. The processes differ only in terms of the *qualitative* form that these determinants take. They differ, as can be seen in the table,

SUMMARY OF THE DISTINCTIONS BETWEEN  
THE THREE PROCESSES

	<i>Compliance</i>	<i>Identification</i>	<i>Internalization</i>
<b>Antecedents:</b>			
1. Basis for the importance of the induction	Concern with social effect of behavior	Concern with social anchorage of behavior	Concern with value congruence of behavior
2. Source of power of the influencing agent	Means control	Attractiveness	Credibility
3. Manner of achieving prepotency of the induced response	Limitation of choice behavior	Delineation of role requirements	Reorganization of means-ends framework
<b>Consequents:</b>			
1. Conditions of performance of induced response	Surveillance by influencing agent	Salience of relationship to agent	Relevance of values to issue
2. Conditions of change and extinction of induced response	Changed perception of conditions for social rewards	Changed perception of conditions for satisfying self-defining relationships	Changed perception of conditions for value maximization
3. Type of behavior system in which induced response is embedded	External demands of a specific setting	Expectations defining a specific role	Person's value system

in terms of the *basis* for the importance of the induction, the *source* of the influencing agent's power, and the *manner* of achieving prepotency of the induced response.

1. The processes can be distinguished in terms of the basis for the importance of the induction, that is, in terms of the nature of the motivational system that is activated in the influence situation. What is it about the influence situation that makes it important, that makes it relevant to the individual's goals? What are the primary concerns that the individual brings to the situation or that are aroused by it? The differences between the three processes in this respect are implicit in the descriptions of the processes given above: (a) To the extent that the individual is concerned—for whatever reason—with the *social effect* of his behavior, influence will tend to take the form of compliance. (b) To the extent that he is concerned with the *social anchorage* of his behavior, influence will tend to take the form of identification. (c) To the extent that he is concerned with the *value congruence* of his be-

havior (rational or otherwise), influence will tend to take the form of internalization.

2. A difference between the three processes in terms of the source of the influencing agent's power is hypothesized. (a) To the extent that the agent's power is based on his *means control*, influence will tend to take the form of compliance. An agent possesses means control if he is in a position to supply or withhold means needed by the individual for the achievement of his goals. The perception of means control may depend on the agent's *actual* control over specific rewards and punishments, or on his *potential* control, which would be related to his position in the social structure (his status, authority, or general prestige). (b) To the extent that the agent's power is based on his *attractiveness*, influence will tend to take the form of identification. An agent is attractive if he occupies a role which the individual himself desires<sup>12</sup> or if he occupies a role reciprocal to one the individual wants to establish or maintain. The term "attractiveness," as used here, does not refer to the possession of qualities that make a person likable, but rather to the possession of qualities on the part of the agent that make a continued relationship to him particularly desirable. In other words, an agent is attractive when the individual is able to derive satisfaction from a self-definition with reference to him. (c) To the extent that the agent's power is based on his *credibility*, influence will tend to take the form of internalization. An agent possesses credibility if his statements are considered truthful and valid, and hence worthy of serious consideration. Hovland, Janis, and Kelley<sup>13</sup> distinguish two bases for credibility: expertness and trustworthiness. In other words, an agent may be perceived as possessing credibility because he is likely to *know* the truth, or because he is likely to *tell* the truth. Trustworthiness, in turn, may be related to over-all respect, likemindedness, and lack of vested interest.

3. It is proposed that the three processes differ in terms of the way in which prepotency is achieved. (a) To the extent that the induced response becomes prepotent—that is, becomes a "distinguished path" relative to alternative response possibilities—because the individual's choice behavior is limited, influence will tend to take the form of compliance. This may happen if the individual is pressured into the induced response, or if alternative responses are blocked. The induced response thus becomes prepotent because it is, essentially, the only response permitted: the individual sees himself as having no

<sup>12</sup> This is similar to John Whiting's conception of "Status Envy" as a basis for identification. See J. W. M. Whiting, "Sorcery, Sin, and the Superego," in M. R. Jones, editor, *Nebraska Symposium on Motivation*, Lincoln, University of Nebraska Press, 1959, pp. 174-195.

<sup>13</sup> *Op.cit.*, p. 21.

choice and as being *restricted* to this particular alternative. (b) To the extent that the induced response becomes prepotent because the requirements of a particular role are delineated, influence will tend to take the form of identification. This may happen if the situation is defined in terms of a particular role relationship and the demands of that role are more or less clearly specified; for instance, if this role is made especially salient and the expectations deriving from it dominate the field. Or it may happen if alternative roles are made ineffective because the situation is ambiguous and consensual validation is lacking. The induced response thus becomes prepotent because it is one of the few alternatives available to the individual: his choice behavior may be unrestricted, but his opportunity for selecting alternative responses is limited by the fact that he is operating exclusively from the point of view of a particular role system. (c) Finally, to the extent that the induced response becomes prepotent because there has been a reorganization in the individual's conception of means-ends relationships, influence will tend to take the form of internalization. This may happen if the implications of the induced response for certain important values—implications of which the individual had been unaware heretofore—are brought out, or if the *advantages* of the induced response as a path to the individual's goals, compared to the various alternatives that are available, are made *apparent*. The induced response thus becomes prepotent because it has taken on a new meaning: as the relationships between various means and ends become restructured, it emerges as the preferred course of action in terms of the person's own values.

Depending, then, on the nature of these three antecedents, the influence process will take the form of compliance, identification, or internalization. Each of these corresponds to a characteristic pattern of internal responses—thoughts and feelings—in which the individual engages as he accepts influence. The resulting changes will, in turn, be different for the three *processes*, as indicated in the second half of the table. Here, again, it is assumed that there are no systematic quantitative differences between the processes, but rather qualitative variations in the subsequent histories of behavior adopted through *each* process.

1. It is proposed that the processes differ in terms of the subsequent conditions under which the induced response will be *performed* or expressed. (a) When an individual adopts an induced response through compliance, he tends to perform it only under conditions of *surveillance* by the influencing agent. These conditions are met if the agent is physically present, or if he is likely to find out about the individual's actions. (b) When an individual adopts an induced response through identification, he tends to perform it only under conditions of *salience* of his relationship to the agent. That is, the occurrence of the behavior

will depend on the extent to which the person's relationship to the agent has been engaged in the situation. Somehow this relationship has to be brought into focus and the individual has to be acting within the particular role that is involved in the identification. This does not necessarily mean, however, that he is consciously aware of the relationship; the role can be activated without such awareness. (c) When an individual adopts an induced response through internalization, he tends to perform it under conditions of *relevance of the values* that were initially involved in the influence situation. The behavior will tend to occur whenever these values are activated by the issues under consideration in a given situation, quite regardless of surveillance or salience of the influencing agent. This does not mean, of course, that the behavior will occur every time it becomes relevant. It may be out-competed by other responses in certain situations. The probability of occurrence with a given degree of issue relevance will depend on the strength of the internalized behavior.

2. It is hypothesized that responses adopted through the three processes will differ in terms of the conditions under which they will subsequently be abandoned or changed. (a) A response adopted through compliance will be abandoned if it is no longer perceived as the best path toward the attainment of social rewards. (b) A response adopted through identification will be abandoned if it is no longer perceived as the best path toward the maintenance or establishment of satisfying self-defining relationships. (c) A response adopted through internalization will be abandoned if it is no longer perceived as the best path toward the maximization of the individual's values.

3. Finally, it is hypothesized that responses adopted through the three processes will differ from each other along certain qualitative dimensions. These can best be summarized, perhaps, by referring to the type of behavior system in which the induced response is embedded. (a) Behavior adopted through compliance is part of a system of external demands that characterize a specific setting. In other words, it is part of the rules of conduct that an individual *learns* in order to get along in a particular situation or series of situations. The behavior tends to be related to the person's values only in an instrumental rather than an intrinsic way. As long as opinions, for example, remain at that level, the individual will tend to regard them as not really representative of his true beliefs. (b) Behavior adopted through identification is part of a system of expectations defining a particular role—whether this is the role of the other which he is taking over, or a role reciprocal to the other's. This behavior will be regarded by the person as representing himself, and may in fact form an important aspect of himself. It will tend to be isolated, however, from the rest of the person's values—

to have little interplay with them. In extreme cases, the system in which the induced response is embedded may be encapsulated and function almost like a foreign body within the person. The induced responses here will be relatively inflexible and stereotyped. (c) Behavior adopted through internalization is part of an internal system. It is fitted into the person's basic framework of values and is congruent with it. This does not imply complete consistency: the degree of consistency can vary for different individuals and different areas of behavior. It does mean, however, that there is some interplay between the new beliefs and the rest of the person's values. The new behavior can serve to modify existing beliefs and can in turn be modified by them. As a result of this interaction, behavior adopted through internalization will tend to be relatively idiosyncratic, flexible, complex, and differentiated.

#### RESEARCH BASED ON THE MODEL

The model itself and its possible implications may be seen more clearly if I present a brief summary of the research in which it was used. This research has moved in three general directions: experimental tests of the relationships proposed by the model, application of the model to the study of personality factors in social influence, and application of the model to the analysis of a natural influence situation.

*Experimental tests of the proposed distinctions between the three processes.* The relationships proposed by the model can be tested by experiments in which the antecedents postulated for a given process are related to the consequents postulated for that process. The first experiment on this problem<sup>14</sup> varied one of the antecedents—the source of the influencing agent's power—and observed the effects of this variation on one of the consequents—the conditions of performance of the induced response. Subjects (Negro college freshmen) were exposed to a tape-recorded interview dealing with an aspect of the Supreme Court decision on school segregation. Four versions of this communication were developed and played to different groups of subjects. The four communications contained the same message, but they differed in the way in which the communicator was introduced and presented himself at the beginning of the interview. These differences were designed to vary the source and degree of the communicator's power: in one communication the speaker was presented as possessing high means control, in the second as possessing high attractiveness, in the third as possessing high credibility, and in the fourth (for purposes of comparison) as being low in all three of these sources of power.

<sup>14</sup> H. C. Kelman, "Compliance, Identification and Internalization: Three Processes of Attitude Change," *Journal of Conflict Resolution*, Vol. 2, 1958, pp. 51-60.

The subjects filled out attitude questionnaires designed to measure the extent of their agreement with the communication. To vary the conditions of performance, we asked each subject to complete three separate questionnaires, one under conditions of salience and surveillance, one under conditions of salience of the communicator—but without surveillance, and a third under conditions of nonsurveillance and nonsalience. It was predicted that attitudes induced by the communicator high in means control would tend to be expressed only under conditions of surveillance by the communicator (the mediating process here being compliance), attitudes induced by the communicator high in attractiveness would tend to be expressed only when the subject's relationship to the communicator was salient (the mediating process here being identification), and attitudes induced by the communicator high in credibility would tend to be expressed when they were relevant in content, regardless of surveillance or salience (the mediating process here being internalization). These predictions were confirmed to a most encouraging degree.

One implication of this study for the analysis of public opinion is that we can make certain predictions about the future course of a given opinion if we know something about the interpersonal circumstances under which it was formed. An interview might reveal the predominant dimensions in terms of which the respondent perceives those individuals and groups to whom he traces the opinion in question. For example, does he see them primarily as potential sources of approval and disapproval? Or as potential reference points for his self-definition? Or as potential sources of information relevant to his own concern with reality testing and value maximization? From the answers to these questions we should be able to predict the future conditions under which this opinion is likely to come into play.

The study also suggests possible "diagnostic" devices that would make it possible to infer the process by which a particular opinion was adopted and hence the level at which it is held. If, for example, an opinion is expressed only in the presence of certain crucial individuals, one can assume that it is probably based on compliance and one can make certain further inferences on that basis. In other words, by observing the "conditions of performance of the induced response" (one of the consequents in our model), we can deduce the process on which this response is based.

It would, of course, be considerably easier and safer to make such inferences if several diagnostic criteria were available. It would be useful, therefore, to derive—from the list of consequents postulated by the model—further indicators in terms of which compliance-based, identification-based, and internalized opinions can be distinguished



from one another, and to test the validity of these indicators. This is particularly true for identification and internalization. Since both of these processes, presumably, produce changes in "private belief," it is difficult to pin down the distinction between opinions based on them. There is a need, therefore, to develop a number of indicators that can capture the qualitative differences in the nature of opinions produced by these two processes, subtle though these differences may be. A second experiment addressed itself to this problem.<sup>15</sup>

The experimental situation, again, involved the use of tape-recorded communications. Three versions of the communication were used, each presented to a different group of college students. In each of the communications a novel program of science education was described and the rationale behind it was outlined. The basic message was identical in all cases, but the communications differed in terms of certain additional information that was included in order to produce different orientations. In one communication (*role-orientation* condition) the additional information was designed to spell out the implications of the induced opinions for the subject's relationship to certain important reference groups. Positive reference groups were associated with acceptance of the message, and—in a rather dramatic way—negative reference groups were associated with opposition to it. The intention here was to create two of the postulated antecedents for *identification*: a concern with the social anchorage of one's opinions, and a delineation of the requirements for maintaining the desired relationship to one's reference groups (see the table). In the second communication (*value-orientation* condition) the additional information was designed to spell out the implications of the induced opinions for an important value—personal responsibility for the consequences of one's actions. The communication argued that acceptance of the message would tend to maximize this value. The intention here was to create two of the postulated antecedents of *internalization*: a concern with the value congruence of one's opinions, and a reorganization of one's conception of means-ends relationships. The third communication was introduced for purposes of comparison and contained only the basic message.

On the basis of the theoretical model it was predicted that the nature of the attitude changes produced by the two experimental communications would differ. Role orientation would presumably produce the consequences hypothesized for identification, while value orientation would produce the consequences hypothesized for internalization. A number of measurement situations were devised to test these predic-

<sup>15</sup> H. C. Kelman, "Effects of Role-orientation and Value-orientation on the Nature of Attitude Change," paper read at the meetings of the Eastern Psychological Association, New York City, 1960.

tions: (1) In each group, half the subjects completed attitude questionnaires immediately after the communication, under conditions of salience, and half completed them a few weeks later, under conditions of nonsalience. As predicted, there was a significant difference between these two conditions of measurement for the role-orientation group but not for the value-orientation group. (2) The generalization of the induced attitudes to other issues involving the same values, and to other situations involving similar action alternatives, was measured. The prediction that the value-orientation group would show more generalization than the role-orientation group on the value dimension tended to be confirmed. The prediction that the reverse would be true for generalization along the action dimension was not upheld. (3) Flexibility of the induced attitudes was assessed by asking subjects to describe their doubts and qualifications. As predicted, the value-orientation group scored significantly higher on this index. (4) Complexity of the induced attitudes was assessed some weeks after the communication by asking subjects to list the things they would want to take into account in developing a new science education program. The total number of items listed was greater for the role-orientation group, but the number of items showing an awareness of relevant issues (as rated by a naïve judge) was clearly greater in the value-orientation group. (5) Half the subjects in each group were exposed to a countercommunication presenting a new consensus, the other half to a countercommunication presenting new arguments. It was predicted that the role-orientation group would be relatively more affected by the first type of countercommunication, and the value-orientation group by the second. The predicted pattern emerged, though it fell short of statistical significance.

The results of this study are not entirely unambiguous. They are sufficiently strong, however, to suggest that it should be possible to develop a number of criteria by which identification-based and internalized attitudes can be distinguished from one another. On the basis of such distinctions, one can then make certain inferences about the meaning of these attitudes and further predictions about their future course.

*The relation between personality factors and social influence.* This research starts with the assumption that the specific personality variables that are related to the acceptance of influence will depend on the particular process of influence involved. There is a further assumption that relationships depend on the type of influence situation to which the person is exposed. In other words, **the concern is with exploring the specific personality variables that predispose individuals to engage in each of the three processes, given certain situational forces.**

In the first study of this problem<sup>16</sup> we were interested in the relationship between one type of personality variable—cognitive needs and styles—and the process of internalization. We wanted to study this relationship in a situation in which people are exposed to new information that challenges their existing beliefs and assumptions. This is a situation in which at least some people are likely to re-examine their beliefs and—if they find them to be incongruent with their values in the light of the new information—they are likely to change them. A change under these particular motivational conditions would presumably take the form of internalization.

It was proposed that people who are high in what might be called the *need for cognitive clarity* would react more strongly to a situation of this type. They would be made uncomfortable by the incongruity produced by such a situation and the challenge it presented to their cognitive structures. The *nature* of their reaction, however, may differ. Some people may react to the challenge by changing their beliefs, while others may react by resisting change. Which of these directions an individual would be likely to follow would depend on his characteristic *cognitive style*. A person who typically reacts to ambiguity by seeking clarification and trying to gain understanding (a "clarifier") would be likely to open himself to the challenging information and perhaps to reorganize his beliefs as a consequence. A person who typically reacts to ambiguity defensively, by simplifying his environment and keeping out disturbing elements (a "simplifier"), would be likely to avoid the challenging information.

Measures of cognitive need and cognitive style were obtained on a group of college students who were then exposed to a persuasive communication that presented some challenging information about American education. Change in attitudes with respect to the message of the communication was measured on two occasions for each subject: immediately after the communication, under conditions of salience, and six weeks later under conditions of nonsalience.

We predicted that, among people high in need for cognitive clarity, those whose characteristic style is clarification would be the most likely to manifest attitude change in the induced direction, while those whose characteristic style is simplification would be the most likely to manifest resistance to change and possibly even negative change. This difference should be especially marked under conditions of nonsalience, which are the conditions **necessary for a reasonable test of internalization**. Among the people who are low in need for cognitive clarity, it was

<sup>16</sup> H. C. Kelman and J. Cohler, "Reactions to Persuasive Communication as a Function of Cognitive Needs and Styles," paper read at the meetings of the Eastern Psychological Association, Atlantic City, 1959.

predicted that cognitive style would be unlikely to produce consistent differences since they are less motivated to deal with the ambiguity that the challenging information has created.

The results clearly supported these predictions. High-need clarifiers showed more change than high-need simplifiers (who, in fact, changed in the negative direction). This difference was small under conditions of salience, but became significant under conditions of nonsalience—suggesting that the difference between clarifiers and simplifiers is due to a difference in their tendency to internalize. Among low-need subjects, no consistent differences between the two style groups emerged.

This study suggests that one can gain a greater understanding of the structure of an individual's opinions on a particular issue by exploring relevant personality dimensions. In the present case we have seen that, for some of the subjects (those concerned with cognitive clarity), the opinions that emerge represent at least in part their particular solution to the dilemma created by incongruous information. In studies that are now under way we are exploring other personality dimensions that are theoretically related to tendencies to comply and identify. If our hypotheses are confirmed in these studies, they will point to other ways in which emerging opinions may fit into an individual's personality system. Opinions may, for example, represent partial solutions to the dilemmas created by unfavorable evaluations from others or by finding oneself deviating from the group. Since these relationships between opinions and personality variables are tied to the three processes of influence in the present model, certain predictions about the future course of the opinions for different individuals can be readily derived.

*The application of the model to the analysis of a natural influence situation.* We are currently engaged in an extensive study of Scandinavian students who have spent a year of study or work in the United States.<sup>17</sup> We are interested in the effects of their stay here on their self-images in three areas: nationality, profession, and interpersonal relations. Our emphasis is on learning about the processes by which changes in the self-image come about or, conversely, the processes by which the person's existing image maintains itself in the face of new experiences. Our subjects were questioned at the beginning of their stay in the United States and at the end of their stay, and once again a year after their return home.

This study was not designed as a direct test of certain specific hypotheses about the three processes of influence. In this kind of rich field situation it seemed more sensible to allow the data to point the

<sup>17</sup> Lotte Bailyn and H. C. Kelman, "The Effects of a Year's Experience in America on the Self-Image of Scandinavians: Report of Research in Progress," paper read at the meetings of the American Psychological Association, Cincinnati, 1959.

way and to be open to different kinds of conceptualizations demanded by the nature of the material. The model of influence did, however, enter into the formulation of the problem and the development of the schedules and is now entering into the analysis of the data.

In a preliminary analysis of some of our intensive case material, for example, we found it useful to differentiate four patterns of reaction to the American experience which may affect various aspects of the self-image: (1) An individual may change his self-image by a *reorganization* of its internal structure; here we would speak of a *change by means* of the process of *internalization*. (2) His self-image may *be changed* by a reshaping of the social relationships in which this image is anchored; *here* we would speak of a change by means of *identification*. (3) The individual may focus on the internal structure of the self-image but maintain it essentially in its original form; here we would speak of the process of *confirmation*. Finally, (4) he may maintain his self-image through a focus on its original social anchorage; here maintenance by the process of *resistance* would be involved. We have related these four patterns to a number of variables in a very tentative way, but the analysis will have to progress considerably farther before we can assess the usefulness of this scheme. It is my hope that this kind of analysis will give us a better understanding of the attitudes and images that a visitor takes away from his visit to a foreign country and will allow us to make some predictions about the subsequent history of these attitudes and images. Some of these predictions we will be able to check out on the basis of our post-return data.

#### CONCLUSION

There is enough evidence to suggest that the distinction between *compliance*, *identification*, and *internalization* is valid, even though it has certainly not been established in all its details. The specification of distinct antecedents and consequents for each of the processes has generated a number of hypotheses which have met the experimental test. It seems reasonable to conclude, therefore, that this model may be useful in the analysis of various influence situations and the resulting opinion changes. It should be particularly germane whenever one is concerned with the quality and durability of changes and with the motivational conditions that produced them.

I have also attempted to show the implications of this model for the analysis of public opinion. By tying together certain antecedents of influence with certain of its consequents, it enables us to infer the motivations underlying a particular opinion from a knowledge of its manifestations, and to predict the future course of an opinion from a knowledge of the conditions under which it was formed. Needless to say, the

**usefulness** of the model in this respect is limited, not only because it is still in an early stage of development but also because of the inherent complexity of the inferences involved. Yet it does suggest an approach to the problem of meaning in the analysis of public opinion data.