10

Common-sense psychology

Daniel M. Wegner and Robin R. Vallacher
Trinity University and Illinois Institute of Technology

Where did the study of social cognition begin? Some might trace the roots of the field to sociological ground, noting its precursors in symbolic interactionism and phenomenological sociology. Others might look for the starting points in the history of psychology, discerning the important Gestalt and Lewinian influences. Still others might argue that the area has only recently come into its own with the wholesale borrowing of concepts and paradigms from cognitive and information-processing psychology. Although each of these claims deserves some credence, we believe the present unity and intensity of the field can best be ascribed to a single idea, one that has repeatedly

surfaced as a catalyst for the integration of these many streams of thought. This is the idea that each person's understanding of the social world can be likened to a common-sense psychological theory.

One purpose of this chapter is to chart the history and recurrent vitality of the "everyman as psychologist" metaphor in social cognitive psychology. A second and broader purpose, however, is to use this background to introduce a distinction between two very different meanings of the idea in past writings. In what we term the *implicit* form of common-sense psychology, theorists have used "common sense" as a simple synonym for social cognition; common-sense psychology is seen as *how people think* about the social world. The *explicit* form of common-sense psychology, in turn, uses "common sense" to refer to the descriptions, explanations and accounts that people give for psychological phenomena; common-sense psychology is seen as *how people think they think* about the social world. As we hope to show in these pages, this distinction is more than a simple clarification of previous work. Indeed, it offers a useful way of conceptualizing much of what is currently known or surmized about the role of metacognitive activity in social cognition.

Psychology and common sense

There are many areas of interplay between common sense and psychology. Philosophers of science tell us that the beginnings of any science, psychology included, must lie in common sense (e.g. Kaplan, 1964); students of psychology often remark on the common-sense nature of what their psychology professors have to say. Writers of psychological theories find their largest and most contrary audiences when some aspect of common sense is challenged by their positions. Perhaps with such observations in mind, a number of commentators have noted specific parallels between the everyman's common-sense understanding of the social world and the professional psychologist's scientific theories.

A BRIEF HISTORY

The earliest writings that bear a clear connection to common-sense psychology occurred in the initial attempts to distinguish scientific and common-sense understanding in the social sciences. One line of this

work began with Max Weber's (1922) observation that while the everyman and social scientist might in fact have very similar subject matters to understand, their strategies and modes of analysis might differ because they have different purposes in mind (see also Schutz, 1935/1967). Though both the everyman and the social scientist achieve social understanding in everyday life in the service of personal survival, the social scientist may achieve an expanded understanding in scientific settings devoid of survival implications. A related thread of thinking began with William James' (1935) realization that the very "obviousness" of certain features of human life might lead social scientists to ignore their importance. In this sense, the mere similarity of the concerns of the everyman and the social scientist could lead the social scientist astray.

These two strains of thought were first combined in the context of "social perception" by Gustav Ichheiser (1949). His analyses presaged many later ideas, but have generally been less visible than three statements of the theme that appeared in the mid-1950s. Jerome Bruner and Renato Tagiuri (1954) proposed that a person's "implicit personality theory", like the psychologist's more formal counterpart, might be identified as the source of the person's inferences about the characteristics of others. George A. Kelly (1955) developed a theory of personality that centered on the "man as scientist" metaphor, and Fritz Heider (1958) offered a compelling assortment of ideas about "naïve psychology", the individual's common-sense understanding of social relations. With these contributions, common-sense psychology became the root metaphor for the field of social cognition. The almost magnetic effect of the metaphor resulted in the assimilation of work related to each of these contributions into the common-sense psychology tradition. So, for example, studies in the impression-formation paradigm of Asch (1946), research on clinical judgment, work on the "halo effect", and many other studies of person perception became identified with the tradition of implicit personality theory (see Rosenberg and Sedlak, 1972; Schneider, 1973). In the same way, most cognitive structural research in clinical psychology became linked with Kelly's "man as scientist" position, and much of the research on cognitive balance and consistency became associated with Heider's "naïve psychology".

The common-sense psychology theme has become increasingly commonplace in social cognitive writings over the years, often because

it provides a way to conceptualize the more rational, "scientific" side of human nature. Kelley (1967, 1971), for instance, helped to orchestrate the tremendous outpouring of research on lay attribution by conceptualizing the everyman's causal inference processes as data-driven, inductive and "scientific". Epstein (1973) suggested that the self-concept might be profitably viewed as a theory about one's own personal functioning, and that such "self theories" develop in ways that parallel the advance of scientific understanding. Wegner and Vallacher (1977) wrote a text on social cognition under the title of "Implicit Psychology", using the subdivisions of the field of professional psychology to categorize the everyman's areas of common-sense expertise.

To be sure, the common-sense psychology theme also allows for an emphasis on the shortcomings of lay judgment, and indeed, several writers have employed formal psychology as a standard against which the everyman's psychology is evaluated. Wrightsman (1974), for instance, explored common-sense psychology by collecting people's assumptions about human nature, assumptions that often appear naïve in comparison to formal theories. Ajzen (1977) studied "intuitive theories of events", and showed that causal attribution is heavily influenced by preconceptions that interfere with pure data-driven inference. Ross (1977) reviewed a number of biases in attribution and prediction with a view toward examining "the intuitive psychologist and his shortcomings". Nisbett and Wilson (1977) ventured the idea that people offer information about their own psychological processes by appealing to "a priori causal theories", and provided evidence that these theories are demonstrably inaccurate in several circumstances. Snyder and Swann (1978) detailed processes whereby the everyman tests hypotheses about the attributes of people, and demonstrated how these processes often operate as self-fulfilling prophesies. Nisbett and Ross (1980) reviewed the human judgement literature, comparing the relative acuity of everyman and scientist in a number of domains, and finding the everyman a poor second in most. In sum, the intriguing parallel between the psychologist and the everyman that becomes available upon embracing a common-sense psychology perspective has yielded important insights into both the everyman's social cognitive talents and his social cognitive faults.

Looking through this literature in search of the general definition of "common-sense psychology", we have found a marked tendency for

proponents of the idea to adopt either one or the other of two quite distinct interpretations. As a way of dispelling the false integration of these meanings perpetuated by the use of "common-sense psychology" to refer to both, we suggest a partition in terms of implicit and explicit forms of common-sense psychology. In the next sections, we first define what is meant by implicit theory in common-sense psychology, and then consider what is signified by explicit theory in common-sense psychology.

IMPLICIT THEORY

When a psychologist makes some theoretical inference about a client or a research subject, he is typically not thinking about the theory itself. Rather, the psychologist's attention is focussed on the specific contents of the inference (e.g. "This person's failure to solve the problem indicates poor intelligence"). Although the psychologist may often be explicitly concerned with properties of the theory (e.g. "Does this problem tap intelligence?), and may also be able to identify or explain the theory at another point in time, the theory is *implicit* in the inference at the time the inference is being made. It is precisely this implicit quality of a theory-in-use that can lead psychologists to argue that they are "dust bowl empiricists", discoverers unguided by any sort of preconception (cf. Kaplan, 1964). And, it is this same property of the everyman's systems of social cognition that is emphasized in the "implicit theory" interpretation of common-sense psychology.

Common-sense psychology, in this view, is a system that is responsible for the individual's conscious awareness of the phenomena of the social world; in essence, it is the entire array of social cognitive structures and processes by which the contents of the individual's phenomenal field (objects, events, people, situations, behaviours, relationships, and even the self) are constructed from the flow of experience.

However, because the individual understands aspects of the social world *through* his system of common-sense theory, and not by looking *at* the theory itself, the theory can be characterized as implicit in its operation. Just as the telescope one looks through may be implicit in one's conscious perception of a planet, or a telephone system one listens through may be implicit in one's awareness of the voice at the other end, one's own sensory, perceptual and cognitive systems are implicit as

they construct and lend meaning to one's conscious social experience (cf. Polanyi, 1969; Wegner, in press).

This view of common-sense psychology as implicit theory has been the one used most frequently by common-sense psychology analysts. In the case of "implicit personality theory" for example, Bruner and Tagiuri (1954) began with the simple observation that perceivers tend to make inferences about personality; given information indicating that a person has one trait (e.g. "George is defensive"), perceivers often go on to infer that the person has other, seemingly related traits (e.g. "He might also be paranoid"). By the Bruner and Tagiuri definition of implicit personality theory, however, the perceiver need not necessarily know that he is doing this. All that is necessary for the postulation of an implicit personality theory in a particular perceiver is the psychologist's observation of inferential regularity in the perceiver's judgments.

This tendency to identify common-sense psychological theory as implicit, and hence as undiscoverable by means other than scientific investigation by professional psychologists, now pervades several literatures. Common-sense theory is assumed to be implicit in most subsequent work on implicit personality theory (e.g. Rosenberg and Sedlak, 1972; Schneider, 1973), in much of the thinking on personal constructs that has derived from Kelly's (1955) personality theory (e.g. Duck, 1973), and in a variety of other frameworks that have used common-sense psychology as a convenient synonym for social cognitive structures and processes (e.g. Ajzen, 1977; Epstein, 1973; Kelley, 1967, 1971, 1980; Ross, 1977; Wegner and Vallacher, 1977). Because cognitive and information-processing psychology also involves the postulation of mental systems that are implicit in their operation (cf. Mandler, 1975; Shallice, 1978), those accounts of social cognition that draw upon cognitive psychology for their models of thought (e.g. Hastie et al., 1980) may be said to have an implicit theory perspective. In all these approaches, the matter of whether the everyman can describe or report on some aspect of his common-sense theoretical system is largely irrelevant to scientific decisions on the existence or form of such a system. To be sure, all the trappings of scientific examination and validation may be called into play for the purpose of testing the claim that a particular implicit theory is necessary to account for a set of social behaviours or judgments. But the everyman's own observations are not counted.

EXPLICIT THEORY

When common-sense psychology is defined as explicit theory, the everyman's ideas about psychological processes are in centre stage. His accounts, descriptions, hypotheses, attributions and explanations themselves are seen as the corpus of common-sense psychological theory that is to be studied by professional psychologists. Although this definition of common-sense psychology has found occasional (and hence, misleading) representation in works that stress implicit theory, analyses of explicit theory unconfounded by this problem can be found as well. Perhaps the most popular such outlet is the social psychology textbook. A standard ploy in the introductory chapters of several such books is the presentation of a set of maxims, proverbs or bits of folk wisdom as "common-sense theories" of social psychology. Then, when certain pairs of maxims are shown to conflict (e.g. "Birds of a feather flock together" as against "Opposites attract"), and the utter senselessness of common-sense psychology has thereby been demonstrated, the writer is free to appraise students of the virtues of the scientific approach to these matters.

Fortunately, a number of more thoughtful explorations of the nature of explicit theory have appeared that save the everyman from a perpetual role as a straw man. These explorations can be thought of as members of three general categories. First, there are a variety of sociological and philosophical accounts of explicit common-sense psychology that fail to identify themselves as such. Theorists who draw upon everyday discourse to describe the common meanings of social behaviour, emotion and thought often reveal themselves as documentors of explicit theory (e.g. Davis, 1973; Goffman, 1959). Second, there are several contributions that portray explicit theory in some detail through empirical investigation. Wrightsman's (1974) questioning of individuals on their assumptions about human nature and Bromley's (1977) collection of ordinary language descriptions of personality are examples. The third category of explicit theory analysis is the study of attribution originated by Heider (1958). Although Heider's prime concern was the detailed exposition of the natural forms of explicit theory, the approach of attribution theorists and researchers who followed (e.g. Jones et al., 1971) has more often centred on examining the covariation between the person's explicit theory and his behaviour.

In all of these approaches, a person's explicit theory is completely known when the person has described and explained a psychological event to his own satisfaction. So, when the person says he bought a new house because ghosts infested the old one, his explicit theory of the house-buying behaviour is determined. Although a social cognitive psychologist studying this person might ask for some embellishment of the explicit theory, the psychologist's own assessment of the accuracy, coherence, descriptiveness or completeness of the person's statement has no bearing on whether it is an explicit theory. It is by definition. With this in mind, it is interesting to return to our earlier example of implicit personality theory. Recall that a person who has an implicit personality theory linking, say, physical attractiveness and friendliness, is likely to go about inferring that beautiful people are friendly; this has been determined by a professional psychologist's observation of covariation in his judgments of these characteristics. The person could then have an explicit theory that corresponds with this inference ("I think attractive people are friendly"); he could have an explicit theory that says nothing ("I think attractive people are good-looking"), or he could have an explicit theory that opposes his observed inference ("I think attractive people are snobs"). For that matter, each of these explicit theories could also exist in the absence of an implicit theoretical connection between attractiveness and friendliness. The point, in sum, is that these two ways of conceptualizing common-sense psychology bear no definitional relation to one another.

Our task for the remainder of this chapter is putting common-sense psychology back together again. Although implicit theory and explicit theory cannot be reconciled at the definitional level, the question of their actual interrelation in the individual's everyday functioning is a fascinating one. How does the person come to know explicitly the features of implicit theory? To what extent does this happen? And given explicit knowledge, can implicit theories be changed or managed as a result? Organizing current social cognitive theory and research to address such questions, we consider first the translation of implicit theory to explicit theory, and then turn to the manner in which explicit knowledge may influence implicit theory.

From implicit to explicit

Suppose we asked a woman on the street to tell us why she liked her husband, and this woman then proceeded to recount in full detail the genetic, environmental, motivational, cognitive, emotional and psychological determinants of her attraction. We should either be amazed at her insight or suspect she had sneaked a look at an introductory psychology textbook. But in either case, we would credit her with a certain amount of *self-knowledge* because her explicit theory corresponds in many respects to what we as psychologists believe her implicit theory to be. Our concern in this section is how and to what degree such self-knowledge may arise.

THE LOGIC OF SELF-KNOWLEDGE

Although the problem of self-knowledge has always been of interest to philosophers (e.g. Dennett, 1978; Hofstadter, 1979; Ryle, 1949), it has also recently become a pressing issue for cognitive psychologists (e.g. Ericcson and Simon, 1980; Mandler, 1975) and social cognitive psychologists (e.g. Bem, 1972; Nisbett and Wilson, 1977b; Pryor, 1980). As a way of drawing the contributions of these different groups into our discussion, it is useful to define what we mean by self-knowledge in the context of common-sense psychology. We can sketch this definition in broad strokes by resorting for the moment to a symbolic representation of the now-familiar implicit personality theory example.

Suppose we give information to a perceiver, Sam, indicating that a person he is about to meet, Lulu, is intelligent. Call this information A. Now, we ask Sam a question (Q_1): Is Lulu likely to be fat or thin? He answers with the information (B) that she will be thin. After observing Sam go through variations on this sequence, each time answering questions like Q_1 with explicit signs that the same inference has taken place, we are likely to propose that Sam has an implicit theory of the form "A implies B". Call this implicit theory I. Suppose we then ask Sam a new question (Q_2): Do you generally believe that intelligent people are thin? He may well be able to generate an explicit theory (E) that has precisely the same form as his implicit theory (I): "A implies B". It is at this point that many commentators on self-knowledge stop the analysis. The apparent parallel of E with I is so tempting that it

leads to the abrupt conclusion that E was produced by some sort of direct access to I in Sam's mind.

Resisting this temptation for the moment, let us continue our line of inquiry by asking Sam a new question (Q_3): Do you believe that "thinking intelligent people are thin" is good or bad? Sam answers "good", providing us with answer C. He does this frequently, and we now infer that he is the proud owner of another implicit theory (I'), this one of the form "$(A$ implies $B)$ implies C". This new implicit theory seems to be of a higher order than the original (I), because instead of only answering questions about B (such as Q_1), it answers questions about "A implies B" (specifically, Q_3). We can now see the crucial point about the logic of self-knowledge that this example makes available. If we ask what kind of implicit theory might be able to answer our question Q_2, and so expose E, it becomes evident that one of the form of I' would be matched to the task, while one of the form of I would not. Translating this conclusion into English, we find that the person must have a higher-level implicit theory to make implicit his understanding of a lower-level implicit theory. Unless we are willing to extend multiple, homunculus-like functions to every implicit theory, we are forced to conclude that explicit self-knowledge of such a theory must arise from yet another implicit theory.

This conclusion is consistent with the assumption of most implicit theory analysts that a person only knows *anything* about the social world (including the self's own processes) by virtue of an implicit theory that brings meaning to experience. From this perspective, one's own implicit theories are only known consciously and explicitly when other of one's implicit theories are focussed on them. This conclusion is also consistent with the several analyses of self-knowledge noted at the outset of this section. Ryle's (1949) philosophy of mind, for example, uses a variety of arguments quite divorced from ours in developing the idea that a simple, "direct access" form of introspective self-knowledge is an impossibility. Hofstadter's (1979) assemblage of insights on self-knowledge gleaned from the fields of artificial intelligence and mathematics suggests in different but compatible terminology the necessity of special "access" systems for the occurrence of any explicit understanding of a thinking machine.

The question of whether explicit common-sense psychology is "accurate", in this light, is the question of whether the understanding reached through any implicit theory is accurate. Just as we might

question the accuracy of a person's implicit personality theory, noting several biases that could manifest themselves in its operation (cf. Wegner and Vallacher, 1977), we may question the accuracy of the higher-level implicit theories that come into play when the person is asked for an explicit theory of his implicit personality theory. Indeed, just such questioning has been the basis of an empirical demonstration of the fallibility of explicit theory. Nisbett and Wilson (1977a) manipulated the warmth *vs* coldness of a stimulus person's style in a between-subjects experimental design, and found that this manipulation had a sizable impact on observers' judgments of the person's mannerisms, appearance and speech accent: all qualities that were not changed as part of the warm-cold manipulation. Yet when these observers were asked to report the process by which they came to judge the person as warm or cold, many indicated that their evaluations of the person's mannerisms, appearance or accent contributed to this judgment. Their explicit theories held that the direction of inference was opposite to that of their implicit theories.

PERSONAL FACTORS IN SELF-KNOWLEDGE

Although Nisbett and Wilson (1977b) have taken this and other findings to indicate that explicit theory is correct at best by chance, their position appears to be too extreme. Counterarguments and new data have arisen from several quarters (e.g. Bowers, in press; Ericcson and Simon, 1980; Smith and Miller, 1978) to suggest conditions under which the individual's mental systems may in fact provide correct explicit accounts of themselves. In this section, we wish to expose the common-sense psychology approach to this controversy by outlining several personal factors, aspects of the person's social cognitive functioning, that impinge on the likelihood that a person could submit a correct explicit theory. Each of these factors represents a way in which an implicit theory may be more or less open to analysis and observation by a higher-level implicit theory that is called on to produce an explicit account of it.

The first personal factor that the accuracy of explicit theory may depend on is the *recency of observation* of an implicit theory in operation (cf. Ericcson and Simon, 1980; White, 1980). Just as a scientist asked for an explanation of an event in history may offer an incorrect *a posteriori* interpretation, the everyman asked for an explicit theory of his

own past activities may err because of the time lag alone. Aspects of the operation of the implicit theory that could be made available by the focussing of a higher-level implicit theory upon it during its operation (if we allow "parallel processing" by the two implicit theories) or shortly thereafter (if we insist on "serial processing") are likely to fall away fairly quickly. Not only will short-term memories fade, but longer-term memories may be modified or even replaced with observations made at a later time (Loftus, 1979). So, if we are to find the everyman with even a rudimentary grasp of his implicit theory, we must extend him the courtesy of prompt questioning.

A second personal factor in the correctness of explicit theory is the degree to which *planned observation* is allowed. Just as we as scientists would like to know ahead of time when we are conducting an experiment, the everyman could be in a much better position to offer an accurate explicit theory of an implicit theory he was planning to observe. In fact, the typical experimental demonstrations of explicit theory gone awry depend critically on unplanned, impromptu self-observation for their effects. Studies in Bem's (1972) self-perception tradition (cf. Enzle, 1980) and in the Schachter and Singer (1962) cognitive labelling of emotion framework (cf. Pennebaker, 1980) regularly show that individuals develop incorrect causal explanations for their activities and feelings under certain conditions. Led to write a counterattitudinal essay in response to subtle social pressure, for example, a person may engage in Bemian self-perception afterward and develop an explicit theory suggesting that he actually agreed with the essay topic from the start. Since this kind of erroneous explicit theory-making does not occur when the person is given information about his prior attitude (Bem and McConnell, 1970) (information that would also become available with planned observation) such demonstrations may be quite misleading regarding the everyman's actual capacities. Although individuals may only infrequently plan their self-observations in everyday life, it seems that the encouragement of such planning in the laboratory could yield more valid statements of explicit theory.

A third personal factor in the accuracy of explicit theory is the extent to which an implicit theory is *disconfirmable*. Somewhat ironically, individuals are most likely to be able to describe the nature of implicit theory when that theory has been proven incorrect. Like the convert who realizes the error of his prior worldview, or the scientist who recognizes unfounded assumptions in a rejected theory, the individual

is in a good position to make implicit theory explicit when the implicit theory is wrong. As a rule, implicit theories that are repeatedly confirmed, as a result of their correctness, as a result of improper hypothesis-testing procedures (cf. Snyder and Swann, 1978) or as a result of their tendency to produce the very effects that they predict (as in a self-fulfilling prophecy), are likely to produce only little observable evidence of their operation. Implicit theories that are not disconfirmable, then, tend to hide themselves such that individuals may be ignorant of their character and unable to offer explicit accounts of their nature.

These personal factors represent only a few of the most crucial determinants of the accuracy of explicit theory that can be drawn from the analogy of person as scientist. Studies of the accuracy of causal attribution (e.g. Fischhoff, 1976; Monson and Snyder, 1977) and examinations of verbal reports on mental processes (e.g. Ericcson and Simon, 1980; Nisbett and Wilson, 1977b) suggest many more. In the next section, we turn from this emphasis on the internal, personal workings of the everyman's investigations to a concern with the more social and interactional factors that can influence translations from implicit to explicit forms of common-sense psychology. Just as the development of science hinges on processes of communication and consensual validation, the development of the person's explicit theories is tied closely to interaction with others.

SOCIAL FACTORS IN SELF-KNOWLEDGE

Why should the individual ever make his implicit theories explicit? Along with Mead (1934) and others of the symbolic interactionist viewpoint, we would argue that explicit self-knowledge is available to the individual only because of the social functions such knowledge may serve. In essence, an explicit understanding of oneself may be necessary for understanding, predicting and controlling both self and others in the course of social interaction. So, just as we might examine the social milieu of the scientist to determine why he publishes what he does (cf. Kuhn, 1970; Harré, Chapter 9), we can consider the social environs of the everyman as a way of understanding the determinants of explicit theory.

A first social factor in the accuracy of explicit theory is the *availability of descriptive terms* in the social lexicon. Although a scientist may invent

descriptive terms for some newly observed phenomenon, he may often adopt jargon from previously existing theories and descriptions. In the same way, the everyman's explicit statements may be formed, supplemented and even distorted by the ways in which his colleagues commonly describe their own implicit theories. To be sure, the operation of many implicit theoretical systems at the neuronal level of brain function may never be susceptible to explicit description because neither their data nor their hypotheses are represented in language (cf. Hofstadter, 1979; Mandler, 1975). But in the description of higher-level implicit theories that yield verbalizable conclusions or hypotheses (e.g. implicit personality theory), available terms play a central role. As the terminology of scientific psychology becomes assimilated into popular culture, it is possible that social cognitive psychologists will become increasingly amazed at the "insights" of the everyman that are produced by this new common lexicon.

The second social factor in explicit theoretical accuracy is the degree to which social forces promote *attention* to implicit theory. In a process somewhat akin to the case in which a scientist's attention is moved toward some phenomena and away from others by the urgings of his colleagues, the everyman's attention to implicit theory may be strengthened or weakened by the social situation in which his observations are made. The mere presence of an audience, for example, or of other items that remind the person of his status as an object of attention (e.g. a mirror or a camera) seems sufficient to induce a degree of self-awareness (Duval and Wicklund, 1972). This awareness, in turn, has been shown in several investigations to increase the validity of a person's self-reports of attitudes and behavioural dispositions (Pryor, 1980). Although such explicit reports of implicit theories seem to be offered easily no matter where the person's attention is directed, their accuracy increases when attention is focussed on the self. An analogous increase in the correctness of explicit theory may also arise when the person focusses attention on others in attempts to ascertain their implicit theories; increased precision in explicit judgments (Vallacher, 1978) and enhanced memory for the other's characteristics (Wegner and Giuliano, in press) have both been found in investigations of such attention. In short, the different perspectives that are engaged by a person in the enterprise of social interaction may impinge on the correspondence of explicit and implicit theories.

A third social determinant of explicit accuracy is the *value of accuracy*

in social interaction. Although the ideals of science suggest that accuracy should be valued without question, accuracy may have a much less central standing in the everyman's priorities. In the quest for prediction and control in social relations, the person may often find it useful to develop erroneously flattering explicit accounts of his own implicit theories. He may deny implicit theories that appear irrational, avoid explicit communication of implicit theories that might be seen as immoral or improper, and even try to obscure explicit understanding when this would place him in a poor position to maintain his worldview (cf. Snyder and Wicklund, in press). It is the small value placed on accuracy in social interaction that is probably most responsible for explicit theory's reputation as a body of *post hoc* interpretation systems that are disengaged in a fundamental way from experience (cf. Heider, 1958). Because explicit theory may be communicated for a variety of purposes, only a few of which might promote accurate representations, the fact that it is offered up regularly in social interaction is no guarantee that it will be true.

To summarize this discussion of self-knowledge, we would like to make the theme underlying our thinking explicit. Quite simply, we refuse to join the stream of commentators on this topic who have taken its appearance as a cue for polarization. Instead of arguing that self-knowledge is near perfect (e.g. Ashworth, 1979) or near impossible (e.g. Nisbett and Wilson, 1977b), we argue that it is near neither extreme. Self-knowledge, as indexed by the capacity of explicit theory to recover key features of implicit theory, is a variable quantity whose personal and social determinants are of special interest. This is true because, as we detail in the next sections, explicit theory may have a profound impact on the nature and function of implicit theory.

From explicit to implicit

Suppose we stopped a woman on the street and gave her a six-week intensive course in social cognitive psychology, detailing current professional knowledge about the nature of her every conceivable implicit theory. If this explicit knowledge, the best we have available, were to have an impact on the functioning of her implicit theories themselves, how would it do so? We can envisage two general ways in which such translation may take place. First, this woman might learn something of her *capacities* as a theorist, and so develop ways of coping

with those aspects of her implicit theories that are inherent and unchangeable. Second, the woman could find out something of her *preferences* as a theorist, and so undertake to revise those preferences that, once explicit, seem like ones she would prefer not to have. In this section, we consider each of these translation processes in turn, and then explore some complicating factors that limit the degree to which explicit theory can influence implicit theory in these ways.

THE MANAGEMENT OF CAPACITIES

Explicit theories often seem to serve as "operating instructions" for the human being. Like the wrinkled slip of paper in the bottom of the box that tells us how to run our new appliance, explicit theory can be conceptualized as a list of features of ourselves and others that is useful in the enterprise of operating each. Because many of the implicit in the enterprise of operating each. Because many of the implicit theories that explicit theory documents are the result of our biological inheritance or lifetime training, or are otherwise unsusceptible to change, they represent capacities of implicit theory to which we may adapt. One might come to an explicit understanding in the course of everyday life, for example, that one is "stupid". Although this global characterization of one's implicit theories might not be accepted with delight, it still might be accepted for its usefulness. Given this information, one might select explicitly "smart" friends to share the load of intellectual tasks, avoid settings in which "stupidity" is tied in explicit theory to "failure", and otherwise arrange one's environment to allow for this shortcoming. To the extent that the explicit theory is true, this pattern of activities represents an important practical application of it.

The findings of formal social cognitive psychology might also be applied. In the same way that management processes can overcome the limitations of implicit theory that are made explicit to us in daily life, certain allowances might be made in accord with limitations we understand explicitly as a result of exposure to the literature of social cognitive psychology (cf. DeSoto, 1979). After all, the parade of biases, shortcomings and irrational tendencies of implicit theory that have been identified by social cognitive psychologists in recent years (see Fischoff, 1976; Nisbett and Ross, 1980; Ross, 1977; Wegner and Vallacher, 1977) seem like specific "stupidities" whose explicit representation might arouse attempts at management. Just as children

develop better memory management capacities with greater explicit knowledge of the nature of storage and retrieval mechanisms (Flavell and Wellman, 1977), individuals might gainsay their own social cognitive incapacities through explicit contact with them.

Of the several attempts social cognitive psychologists have made to intervene with implicit theoretical biases, making them explicit has seemed most effective. Although offering large monetary incentives and allowing unlimited exposure to stimulus information has proven ineffectual in removing one bias, the illusory correlation (Chapman and Chapman, 1969), the provision of explicit knowledge of the bias, the primary effect (Luchins, 1957) and the impression perseverance effect (Ross et al., 1975) have been useful in reducing two others. In future research, varying biases and intervention strategies independently, it may well be determined that explicit knowledge has a moderating effect on implicit theoretical errors. If the scientific study of social cognition is to have any beneficial impact on the social cognitive activities of the everyman, this is one way in which it might do so.

THE REVISION OF PREFERENCES

Explicit theories often reveal our desires and preferences to us. We may be only dimly aware of some motive or preference that an implicit theory guides us to have, as when we prefer to avoid certain people or situations that are implicitly hypothesized to be dangerous or unpleasant. When such preferences are made explicit through self-observation or social feedback, their operation becomes a known entity which is now also subject to our preference (cf. Vallacher, 1980). Knowing that we are "prejudiced" against a certain group, for example, may provide the basis for a "meta-preference" to avoid further thoughts or acts of prejudice. In this view, common-sense explicit theorizing about oneself is the foundation of self-control.

We suspect this process plays a major role in the development and revision of implicit theory. In daily life, a person may think and behave in a certain way on the basis of an implicit theory of the form "A implies B". If A were the presence of a person of a different race, B might be the inference that the person is "inferior". The operation of this implicit theory might very well go unchallenged if it were never acknowledged explicitly. However, if the person obtains the explicit knowledge that "I think people of different races are inferior", there is a possibility of

change. A higher-level implicit theory concerned with this explicit quality of self could lead the person to monitor expressions of the undesirable inferences and behaviours emanating from the lower-level implicit theory, frequently stopping their expression or replacing them with more meta-preferable alternatives. Over time, this monitoring process might well result in a smoothly-running "automatic" implicit theory of the form "A implies C". C might be an inference that the person is "disadvantaged", it might be an inference that "I am cruel", or it might be any of a number of replacements for B. The crucial point is that the initial implicit theory is changed through its explicit representation.

To some degree, this is the kind of change that is sought through "insight" in psychotherapy. A person with some odd or maladaptive set of preferences is given the opportunity to make them explicit to himself, no matter how explicit they may already seem to the therapist. Armed with this information, the person then has the possibility of change in view. Although such explicit common-sense theory does not require change, it does make it available. This, then, is another process whereby the explicit knowledge gained from social cognitive research might find its way into the activities of the everyman. In those aspects of implicit theory that can be represented as preferences rather than capacities, implicit theory itself may be changed through explicit understanding.

LIMITS OF APPLICATION

Given the benefits of self-control and self-management, it might seem that people would be eager and willing to scrutinize their implicit theories and to change them when ineffective, immoral or undesirable features are discovered. The widespread occurrence of self-defeating behaviour, faulty judgment and downright nastiness, however, reminds us that people are often very reticent to question their values and assumptions, even when these features of implicit theory are demonstrably wrong. It is important, then, to consider the functions that may be served by *failing* to use explicit theory as a tool.

Though unexamined aspects of our mental systems contribute to a lack of self-knowledge, such ignorance is essential in promoting one's sense of free will (cf. Hofstadter, 1979). Because our thought-producing structures and processes are necessarily implicit as they are engaged,

we can not know at all times what specific images, judgments, inferences or desires may result from them. And even if we begin to discern these things through higher-level implicit theories that make them explicit, we are still at a loss to understand explicitly the operation of these higher-level implicit theories. It is natural to equate this lack of complete self-knowledge with a lack of determinism. Just as an incomplete understanding of the mechanistic substrate of a computer may allow us to imbue it with goals, purposes and other harbingers of free agency, an incomplete understanding of ourselves allows us the luxury of believing in our own freedom of action. While free will of this sort is clearly an illusion, it is no doubt an adaptive illusion (de Charms, 1968). To a degree, we may even avoid self-knowledge in the pursuit of a feeling of freedom.

In a broader sense, it can be argued that implicit theories constitute our means of making contact with reality. While people may be willing to examine and perhaps discard specific implicit theories in incidental areas of their lives, this is far less likely to be the case for deeply-ingrained theories that provide entire worldviews. As Kuhn (1970) has noted in the context of scientific theory, there is a strong tendency for the most basic theories or paradigms to persist even after their ability to represent reality has been undermined by disconfirming evidence. Scientists tend to look through their theories, not at them, so in the face of such negative evidence, they may focus on shortcomings in the data: that is, in the nature of the reality the theory is designed to represent. Because the theory is the only available means of imposing order and meaning on reality, even a discredited theory may be preferred over no theory at all.

Carried one step further, the analogy between implicit and scientific theory provides a way of understanding when a person will overcome tendencies to avoid explicit self-knowledge and so undergo significant changes in his implicit theories. Kuhn (1970) has suggested that a discredited theory is abandoned only when a better theory becomes available. In a similar fashion, personality theorists and clinicians argue that meaningful changes in cognitive systems cannot be achieved by simply discrediting a person's existing system (e.g. Kelly, 1955). If the management of capacities or the revision of preferences is to be attained, explicit theory of the individual's present implicit theories must be accompanied by explicit theory detailing more useful and desirable implicit theories. In essence, our common sense as

psychologists must be better than that of the everyman if he is to accept our attempts to help.

Conclusion

This analysis of common-sense psychology has drawn in detail on the history of the idea in the field of social cognition. We found that while this history indicated a pivotal role for common-sense psychology in almost every advance in the field, the scientific understanding of common-sense psychology itself was still incomplete. In essence, past writers were basing their frameworks on two common-sense psychologies, one comprised of implicit theory and the other comprised of explicit theory. On making this distinction, we found implications of common-sense psychology for two general areas of social cognitive functioning to which it had not been previously applied: *self-knowledge* (movement from implicit to explicit theory) and *self-control* (movement from explicit theory to implicit theory). It is our hope that this advance may maintain the position of common-sense psychology as an integral part of the continued study of human social cognition.

References

Ajzen, I. Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 1977, **35**, 303–314.

Asch, S. E. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 1946, **41**, 258–290.

Ashworth, P. D. Social Interaction and Consciousness. Chichester, England: Wiley, 1979.

Bem, D. Self-perception theory. In *Advances in Experimental Social Psychology*, Berkowitz, L. (Ed.), Vol. 6, New York and London: Academic Press, 1972.

Bem, D. and McConnell, H. K. Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology*, 14, 23–31, 1970.

Bowers, K. S. Knowing more than we can say leads to saying more than we can know: On being implicitly informed. In Toward a Psychology of Situations. Magnusson, D. (Ed.), Hillsdale, N.J.: Erlbaum (in press).

Bromley, D. B. *Personality Description in Ordinary Language*. London: Wiley, 1977.

Bruner, J. S. and Tagiuri, R. The perception of people. In Handbook of Social Psychology. Lindzey, G. (Ed.), Cambridge, Mass.: Addison-Wesley, 1954.

Chapman, L. J. and Chapman, J. P. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 1969, **74**, 271–280.

Davis, M. S. That's interesting! *Philosophy of Social Science*, 1971, **1**, 309–344.

Davis, M. S. Intimate Relations. New York: Free Press, 1973.

de Charms, R. Personal Causation. New York and London: Academic Press, 1968.

Dennett, D. C. Brainstorms. Montgomery, Vermont: Bradford, 1978.

DeSoto, C. B. Everyman as psychologist. *Contemporary Psychology*, 1979, **24**, 455–456.

Duck, S. W. Personal Relationships and Personal Constructs. London: Wiley, 1973.

Duval, S. and Wicklund, R. A. A Theory of Objective Self Awareness. New York and London: Academic Press, 1972.

Enzle, M. E. Self-perception of motivation. In The Self in Social Psychology. Wegner, D. M. and Vallacher, R. R. (Eds), New York: Oxford, 1980.

Epstein, S. The self-concept revisited: Or a theory of a theory. *American Psychologist*, 1973, **28**, 404–416.

Ericcson, K. A. and Simon, H. A. Verbal reports as data. *Psychological Review*, 1980, **87**, 215–251.

Fischoff, B. Attribution theory and judgment under uncertainty. In *New Directions in Attribution Research*. Harvey, J. H., Ickes, W. J. and Kidd, R. F. (Eds), Vol. 1. Hillsdale, N.J.: Erlbaum, 1976.

Flavell, J. H. and Wellman, H. M. Metamemory. In *Perspectives on the Development of Memory and Cognition*. Kail, R. V. and Hagen, J. W. (Eds), Hillsdale, N.J.: Erlbaum, 1977.

Goffman, E. *The Presentation of Self in Everyday Life*. New York: Doubleday, 1959.

Hastie, R., Ostrom, T. M., Ebbesen, E. B., Wyer, R. S. Jr., Hamilton, D. L. and Carlston, D. E. (Eds), *Person Memory*. Hillsdale, N.J.: Erlbaum, 1959.

Heider, F. *The Psychology of Interpersonal Relations*. New York: Wiley, 1958.

Hofstadter, D. R. *Godel, Escher, Bach*. New York: Basic Books, 1979.

Ichheiser, G. Misunderstandings in human relations. *American Journal of Sociology*, 1949, **55**, 2, Part 2.

James, W. On a certain blindness in human beings. In *Selected Papers on Philosophy*. New York: Dutton, 1935.

Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S. and Weiner, B. (Eds), *Attribution: Perceiving the Causes of Behavior*. Morristown, N.J.: General Learning Press, 1971.

Kaplan, A. *The Conduct of Inquiry*. San Francisco: Chandler, 1964.

Kelley, H. H. Attribution in social psychology. *Nebraska Symposium on Motivation*, 1967, **15**, 192–238.

Kelley, H. H. *Attribution in Social Interaction*. Morristown, N.J.: General Learning Press, 1971.

Kelley, H. H. The causes of behavior: Their perception and regulation. In *Retrospections on Social Psychology*. Festinger, L. (Ed.), New York: Oxford, 1980.

Kelly, G. A. *The Psychology of Personal Constructs*. New York: Norton, 1955.

Kuhn, T. S. *The Structure of Scientific Revolutions* (2nd ed.). Chicago: University of Chicago Press, 1970.

Loftus, E. F. *Eyewitness Testimony*. Cambridge, Mass.: Harvard University Press, 1979.

Luchins, A. S. Primacy-recency in impression formation. In *Order of Presentation in Persuasion*. Hovland, C. (Ed.), New Haven, Conn.: Yale University Press, 1957.

Mandler, G. *Mind and Emotion*. New York: Wiley, 1975.

Mead, G. H. *Mind, Self, and Society*. Chicago: University of Chicago Press, 1934.

Monson, T. C. and Snyder, M. Actors, observers, and the attribution process: Toward a reconceptualization. *Journal of Experimental Social Psychology*, 1977, **13**, 89–111.

Nisbett, R. and Ross, L. *Human inference*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

Nisbett, R. and Wilson, T. D. The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 1977a, **35**, 250–256.

Nisbett, R. and Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 1977b, **84**, 231–259.

Pennebaker, J. W. Self-perception of emotion and internal sensation. In The Self in Social Psychology. Wegner, D. M. and Vallacher, R. R. (Eds), New York: Oxford, 1980.

Polanyi, M. Knowing and Being. Chicago: University of Chicago Press, 1969.

Pryor, J. B. Self-reports and behavior. In *The Self in Social Psychology*. Wegner, D. M. and Vallacher, R. R. (Eds), New York: Oxford, 1980.

Rosenberg, S. and Sedlak, A. Structural representations of implicit personality theory. In *Advances in Experimental Social Psychology*, Vol. 6. Berkowitz, L. (Ed.), New York and London: Academic Press, 1972.

Ross, L. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in Experimental Social Psychology*, Vol. 10. Berkowitz, L. (Ed.), New York and London: Academic Press, 1977.

Ross, L., Lepper, M. R. and Hubbard, M. Perseverence in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 1975, **32**, 880–892.

Ryle, G. *The Concept of Mind.* London: Hutchinson, 1949.

Schachter, S. and Singer, J. Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 1962, **69**, 379–399.

Schneider, D. J. Implicit personality theory: A review. *Psychological Bulletin*, 1973, **79**, 294–309.

Schutz, A. *The Phenomenology of the Social World.* Evanston, Ill.: Northwestern University Press, 1937/1967.

Shallice, T. The dominant action system: An information processing approach to consciousness. In *The Stream of Consciousness*. Pope, K. S. and Singer, J. L. (Eds), New York: Plenum, 1978.

Smith, E. R. and Miller, F. S. Limits on perception of cognitive processes: A reply to Nisbett and Wilson. *Psychological Review*, 1978, **85**, 355–367.

Snyder, M. and Swann, W. B. Jr. Hypothesis testing processes in social interaction. *Journal of Personality and Social Psychology*, 1978, **36**, 1202–1212.

Snyder, M. L. and Wicklund, R. A. Attributional ambiguity. In *New Directions in Attribution Research*, Vol. 3. Harvey, J. H., Ickes, W. J. and Kidd, R. F. (Eds), Hillsdale, N.J.: Erlbaum (in press).

Vallacher, R. R. Objective self awareness and the perception of others. *Personality and Social Psychology Bulletin*, 1978, **4**, 63–67.

Vallacher, R. R. An introduction to self theory. In *The Self in Social Psychology*. Wegner, D. M. and Vallacher, R. R. (Eds), New York: Oxford, 1980.

Weber, M. *Gesammelte Aufsätze zur Wissenschaftslehre.* Mohr, Tübingen.

Wegner, D. M. Justice and the awareness of social entities. In *Equity and Justice in Social Behavior*. Greenberg, J. and Cohen, R. L. (Eds), New York and London: Acdemic Press (in press).

Wegner, D. M. and Giuliano, T. The forms of social awareness. In *Personality, Roles, and Social Behavior*. Ickes, W. J. and Knowles, E. S. (Eds), New York: Springer-Verlag (in press).

Wegner, D. M. and Vallacher, R. R. Implicit Psychology: An Introduction to Social Cognition. New York: Oxford, 1977

White, P. Limitations on verbal reports of internal events: A refutation of Nisbett and Wilson and of Bem. *Psychological Review*, 1980, **87**, 105–112.

Wrightsman, L. S. *Assumptions about Human Nature: A Social-Psychological Approach.* Monterey, Calif.: Brooks-Cole, 1974.