The Role of Theories in Mental Representations and Their Use in Social Perception: A Theory-Based Approach to Significant-Other Representations and Transference

Serena Chen

University of Michigan, Ann Arbor

Social cognition researchers have made substantial progress in elucidating the nature of mental representations and their role in social perception. A diverse array of perspectives on representational content and structure exists (for reviews, see Carlton & E.R. Smith, 1996; E.R. Smith, 1998; Wyer & Carlton, 1994), along with a rich understanding of how and when representations are likely to be activated and brought to bear on the tasks of perceiving and judging the social world (e.g., Higgins, 1996; Wyer & Srull, 1986). Considerable progress has also been made regarding the role of mental representations in guiding social behavior (e.g., Bargh, 1990, 1997). Where do we go from here?

This chapter examines the content, structure, and use of mental representations. It focuses on the idea, long recognized on a theoretical level yet often absent from empirical work, that perceivers' stored mental representations of the social world may include "theories"—that is, explanatory or causal forms of knowledge. If we have not been examining such knowledge at all, then what have we been examining? Even a quick mental survey would reveal that in much social-cognitive research, representations are treated as if they were composed, essentially, of lists of features or attributes. Further, their activation and use are thought to depend, in part, on the fit or overlap between the features of a stimulus and features associated with perceivers' representations (Bruner, 1957a; Higgins, 1996; see also Bargh, 1997; Hardin & Rothman, 1997). For instance, the activation and use of a racial stereotype to interpret a stimulus person are often thought to occur by virtue of the fit between racial features of the person and ones stored as part of a perceivers' stereotype representation. Or, observing a stimulus person engage in a helpful behavior, such as helping an elderly person across a street, is thought to elicit the use of a perceivers' representation of the trait helpful due to the overlap between a feature of the stimulus and one stored as part of the perceivers' trait representation.

In recent years, however, challenges to such a feature-based approach have arisen in the social-cognitive literature (e.g., Wittenbrink, Gist, & Hilton, 1997), paralleling trends in cognitive and developmental work (e.g., Murphy & Medin, 1985). These challenges argue that theories are basic to mental representations and their use. Unlike features, which are often treated as unrelated units of knowledge, theories are seen as embodied in explanatory relations linking bits of knowledge about an entity. To illustrate the potential value of theory-based views, this chapter presents a theory-based approach to mental representations of significant others. It is an approach that brings together theory-based ideas in the cognitive, developmental, and social-cognitive domains with the growing body of evidence for the role of significant-other representations in interpersonal perception and relations.

The role that representations of significant others play in interpersonal perception and relations reflects transference, originally a clinical notion (Freud, 1912/1958; Sullivan, 1953), but one that also has been conceptualized in social-cognitive terms (Andersen & Glassman, 1996). In these terms, transference refers to the phenomenon whereby something about a new person activates a perceivers' representation of a significant other, leading him or her to interpret and respond to the person in ways derived from prior experiences with the significant other. Applying theory-based ideas to research on significant-other representations and transference, the main thrust of the theory-based approach proposed in this chapter is that perceivers are especially likely to have theories stored about their significant others, which implies that the content and structure of significant-other representations involve such theories, and that these theories play an important role in the transference
A BRIEF HISTORY

The content, structure, and use of mental representations are long-standing topics in social cognition (e.g., Lingle, Altom, & Medin, 1984; Wyer & Gordon, 1984) and ones that continue to be debated (e.g., T. R. Smith, 1998). Because many of the issues surrounding them relate to ones discussed in the cognitive literature (e.g., Keil, 1989; Rosch & Mervis, 1975; E. E. Smith & Medin, 1981; Tversky, 1977), a brief history of relevant cognitive work is presented first, followed by a closer look at relevant social-cognitive work.

Features as the Building Blocks: Cognitive Views

The general assumption that features are the primary building blocks of representations has a long history in cognitive views on categories and categorization (for reviews, see Komatsu, 1992; Medin, 1989; Rosch, 1978; E. E. Smith & Medin, 1981). The classical view is often discussed as the starting point of this history (e.g., E. E. Smith & Medin, 1981). This view maintains that categories are defined by a set of necessary and sufficient features. For example, the features orange and round might be a subset of the features that define one's "basketball" category. According to this view, category membership is all-or-none. If an object possesses all of the necessary and sufficient features of a category, it is a category member; otherwise, it is not.

The most immediate criticisms of the classical view took the form of prototype or family resemblance approaches (e.g., Rosch, 1975, 1978; Tversky, 1977; Wittgenstein, 1953), which contend that categories are organized around a prototypical member whose features reflect the central tendency of the category or summary abstracted from various experiences with category examples. In such approaches, category membership is seen as probabilistic or "fuzzy", rather than all-or-none. Objects are members if they possess some, but not necessarily all, of the features associated with the category. Greater featural similarity or overlap with the prototype implies being a more prototypical category member or, put another way, having greater family resemblance. Although classical and prototype perspectives differ fundamentally in their definitions of categories and their views on categorization, they nonetheless share the general view that features are the main building blocks of categories.

Prototype approaches have been challenged by exemplar ones, which argue, in their strongest form, that representations of single instances, called exemplars, are stored in memory rather than prototypes abstracted from experience with various instances (e.g., Medin & Schaffer, 1976; E. E. Smith & Medin, 1981). From an exemplar perspective, then, decisions about category membership depend on the similarity or overlap between the features of an object and those associated with stored exemplars, rather than with a prototype. Although prototype and exemplar views differ in several other basic ways, these differences are beyond the present scope (for recent discussions, see Fiske & Taylor, 1991; Komatsu, 1992; E. R. Smith, 1998). Of key relevance is the fact that, by and large, features have been treated as the basic building blocks of categories in exemplar approaches, just as in classical and prototype ones.

Although assumptions about representational structure have tended to be less explicit than content ones, some inferences might be drawn about the former from the latter. In particular, feature-based approaches have often treated the featural content of representations as unrelated units of knowledge, which is a view that arguably suggests little or rather simple structure, or at least in terms of structure that lies in linkages between features (see Medin, 1989). However, the notion of correlated features has long been discussed (e.g., Chapman & Chapman, 1969; Rosch, Mervis, Gray, Johnson, & Bojes-Braem, 1976) and is one that arguably suggests a greater degree of structure. Nonetheless, neither the principles that account for why features correlations may exist, nor the nature of the linkages between features that such correlations imply, have been spelled out with much precision (for a review, see Murphy & Medin, 1985).

Relative to assumptions about structure, feature-based views on the use of representations have tended to be more explicit. Specifically, feature-based ideas about content and structure have often come hand in hand with the view that categorization, or the likelihood that a representation is activated and used to interpret a stimulus, depends on the number of matching and mismatching features between a stimulus and a stored category (e.g., Tversky, 1977). The greater the number of matching features and the fewer the number of mismatching ones, the higher the likelihood that the stimulus is classified as a category member. Said differently, categorization has often been conceptualized as reflecting some function of feature-based similarity between a stimulus and a category (for reviews, see Goldstone, 1994; Medin, Goldstone, & Gentner, 1993; see also Bruner, 1957a).

Over the last decade or so, several lines of cognitive and developmental research have offered alternatives to feature-based approaches. In broad strokes, these alternatives contend that the building blocks of representations
may extend beyond features. Before considering this work, featural assumptions in social-cognitive views on mental representations and their use are examined.

**Features as the Building Blocks: Social-Cognitive Views**

Featural assumptions have been pervasive not only in social-cognitive work on mental representations and their use (e.g., Bruner, 1957b; Higgins, 1989, 1996; Lingle et al., 1984; Sedikides & Skowronski, 1991; Wyer & Snell, 1986), but also in social-cognitive views of personality (e.g., Cantor & Kihlstrom, 1987, 1989). In these social domains, features have been widely defined as trait attributes (see Zuroff, 1986). For example, representations of professors might be treated as if they include educated and absent-minded as features, whereas representations of Asians might include shy and hardworking. Although it is likely that most social cognition researchers would not, on a theoretical level, endorse simple feature-based definitions of representational content and structure, featural assumptions are often built right into the research methodologies used. For example, the stereotyping literature is composed largely of research in which stereotypic features were presented as stimuli, and participants' stereotypes were assessed in feature-based terms. Nonetheless, tracing the history of social-cognitive work on mental representations and their use reveals that, although features (i.e., traits) have been widely treated, especially on an empirical level, as the basic units of perceivers' conceptions of others, there have been hints of forms of knowledge extending beyond such features from the start.

**Elemental and Holistic Approaches.** The beginnings of social-cognitive work on mental representations might be traced back to work on impression formation. Early on in this work, an important distinction was drawn between elemental and holistic approaches (e.g., Fiske & Taylor, 1991; Leyens & Fiske, 1994). Elemental approaches emphasize the elements or pieces of information that presumably underlie perceivers' impressions of others (e.g., N.H. Anderson, 1974). These elements are treated as separate entities, ones that generally do not exert any influence on one another. In this regard, elemental approaches are fairly akin to feature-based ones in the cognitive literature, discussed earlier, that view the content and structure of categories in terms of lists of features and that tend to treat the features associated with a given category as independent entities.

In contrast, the essence of holistic approaches is that perceivers' impressions of others should be viewed in terms of wholes rather than separate elements. Said differently, impressions reflect coherent configurations of individual elements. Moreover, unlike elemental approaches, holistic ones emphasize the subjective nature of perceivers' construction of reality, thus arguing that perceivers subjectively construe coherent meanings among separate elements. Gestalt psychology, which reflects this holistic view, was a major influence in early social-cognitive work. From a Gestalt perspective, individual elements (e.g., traits) are thought to influence one another, implying that relations exist among these elements such that each element only has meaning in the context of others.

Asch's (1946) classic experiments clearly demonstrated the major propositions of the holistic approach, as did subsequent work in the same tradition (Asch & Zukier, 1984; see also Heider, 1958). For example, emerging from Asch's early work was the concept of central traits, which refer to traits that provide organizing themes around which the individual elements (i.e., traits) of one's knowledge about another person are interpreted and understood. These themes allow the presence of traits related to a central trait to be inferred, and the result is unified and coherent impressions reflecting these themes.

**Implicit Personality Theories.** Compared with elemental views, then, holistic approaches hinted early on that any view of features as the sole building blocks of representations was likely to fall short. The concept of implicit personality theory has a similar flavor in that it refers to the very idea that relations exist among features (for reviews, see Leyens & Fiske, 1994; Schneider, 1973). Indeed, implicit personality theories are thought to be embodied in the linkages between traits. The concept was introduced by Bruner and Taguiri (1954) in part to account for how perceivers are able to form unified impressions around central themes and to infer the presence of traits from the presence of others to which they are related.

It appears, then, that even some of the earliest findings from research reflecting a holistic approach to impression formation suggested, if considered in the language of mental representations, that features may not be the sole units of analysis for representational content, structure, and use. Indeed, as social-cognitive theorizing on representations developed, with the introduction of the schema concept, as well as the importation of the cognitive concepts of prototypes and exemplars, holistic themes persisted.

**Schemas, Prototypes, and Exemplars.** The schema concept emerged in part from holistic notions. A schema is a "cognitive structure that represents knowledge about a concept or type of stimulus, including its attributes and
relations among those attributes" (Fiske & Taylor, 1991, p. 98; see also Brewer & Nakamura, 1984; Carlston & E.R. Smith, 1996; Fiske & Linville, 1980; Markus & Zajonc, 1985; E.R. Smith, 1998). On the one hand, it is important to note in this definition that features/attributes are, once again, treated as the basic building blocks of representations. Yet, on the other hand, there is also an explicit recognition of relations among features, just as there is in the notion of implicit personality theory, implying the existence of internal organization and structure. Indeed, schemas have often been likened to theories that perceivers have about a given concept or stimulus, which harks back to the Gestalt notion of unifying themes or configurations (e.g., Markus & Zajonc, 1985).

The term prototype has often been used interchangeably with schema in the social-cognitive literature (see Fiske & Taylor, 1991; E.R. Smith, 1998), which suits the present purposes. As in the cognitive literature, prototypes in social-cognitive work have referred to categories organized around a summary, central member or prototype, abstracted from experience with various instances (e.g., Cantor & Mischel, 1977, 1979). Likewise, this prototype has been characterized by a "fuzzy" set of features, and category membership has thus been seen as probabilistic rather than all-or-none.

Also paralleling trends in cognitive work, exemplar approaches arose in the social domain as an alternative to prototype or schema views, suggesting that representations of people, groups, and social situations can be highly specific, designating individual instances rather than abstracted summaries of multiple instances (e.g., Lewicki, 1985; Linville & Fischer, 1993; E.R. Smith & Zarate, 1990, 1992). In the social realm, as in the cognitive arena of object categories, the assumptions of schema (or prototype) and exemplar views differ fundamentally in various ways (for a discussion of these differences, see E.R. Smith, 1998). Of central relevance here, however, are the generally converging assumptions these approaches make about how perceivers represent the social world, which are discussed next.

So Where Do Social Cognition Researchers Stand? Taken as a whole, theories in social cognition have tended to treat features, typically in the form of trait attributes, as the basic building blocks of perceivers' impressions of others and, presumably, of the representations that correspond to them. Indeed, in line with featural assumptions about content and structure, prevailing social-cognitive views on the use of representations have also tended to be feature-based. The relevant literature here is that concerned with principles of knowledge accessibility and use. In this body of work, it is widely agreed that, in addition to chronic and transient sources of knowledge accessibility, the activation and use of stored knowledge depend on applicability, which is defined as activation arising from the featural match or overlap perceived between the attended-to features of a stimulus and featural knowledge stored about a knowledge construct (Higgins, 1989, 1996; see also Bargh, 1997; Bruner, 1957b; Hardin & Rothman, 1997; Sedikides & Skowronska, 1991; Wyer & Srull, 1986). By serving as a basis for the contribution of applicability to knowledge activation and use, features have been accorded a central role in governing the use of representations in social perception (see Chen, Andersen, & Hinkley, 1999). In fact, paralleling feature-based views of categorization in the cognitive literature, described earlier, it has been explicitly argued that "the greater the number of shared features [between a stimulus and construct], the more the stimulus would contribute to a construct's activation level" (Higgins, 1996, p. 156).

Yet as noted, there has long been the recognition that perceivers construe, and presumably store in memory, more meaning than implied by a list of independent trait attributes. This can be seen in Asch's (1946) early research on impression formation, in later work in the same Gestalt spirit (e.g., Heider, 1958; Lewin, 1951), such as work on implicit personality theories, and still later in the vast literature on schemas and prototypes. Each of these lines of thinking seems to suggest, in some manner, that organizing themes or configural meanings indeed, theories emerge out of individual attributes. Such emergent meanings imply that attributes are in some way related to one another and, in this regard, that the content, structure, and use of mental representations are likely to involve more than simple lists of unrelated features. Exemplar approaches in the social-cognitive literature, in their defining emphasis on separate instances, also suggest that relations between attributes may exist in being able to account for feature clusters or correlations.

Before drawing the conclusion that social-cognitive work departs from prevailing feature-based approaches in the cognitive literature, it is critical to note several aspects of this work that suggest this conclusion is not yet warranted, at least on empirical grounds. First, as noted, the bulk of social-cognitive research has focused almost exclusively on trait attributes, despite considerable evidence that perceivers' conceptions of others also include knowledge about others' attitudes, thoughts, goals, feelings, and so on (e.g., Andersen, Glassman, & Gold, 1998; Fiske & Cox, 1979; Karmiol, 1986; Park, 1986; Prentice, 1990; Read, Jones, & Miller, 1990; Trzebinski, 1989). Thus, if representational content,
structure, and use involve themes or meanings that extend beyond trait attributes, the social-cognitive literature cannot speak very extensively about these yet.

Second, in most research in the tradition of implicit personality theories, the focus has been limited to trait attributes, and theories have been conceptualized simply in terms of trait covariations or a small number of dimensions (for reviews, see C.A. Anderson & Sedikides, 1991; Schneider, 1973). That pairs or bundles of traits may covary does little to elaborate on, in a substantive way, the view that features (i.e., trait attributes) are the main building blocks of representations. The same holds true even when viewing implicit personality theories in terms of dimensions because such a view is still based on the notion of trait covariations, with traits that fall within or along the same dimension expected to covary. Overall, then, although long-standing social-cognitive theorizing in the spirit of holistic approaches—such as Gestalt views, implicit personality theories, as well as schemas, prototypes, and exemplars—clearly hinted at the notion that features are not the only building blocks of representations, in most existing social-cognitive research, alternatives to a feature-based approach to representational content, structure, and use have not been fully realized.

BEYOND FEATURES: THE ROLE OF THEORY-BASED FORMS OF KNOWLEDGE

Having provided some historical context, recent work on mental representations and their use is examined. The focus is on theory and research which propose, in one form or another, that beyond featural units of knowledge the content, structure, and use of representations involve theoretical, explanatory, or causal forms of knowledge. As a group, these bodies of work are referred to as theory-based approaches despite variations among them. After briefly touching on relevant cognitive and developmental work, emphasis is given to social-cognitive work on theories and theory-based processes so as to lay the groundwork for describing a theory-based approach to significant-other representations and transference.

Cognitive and Developmental Work

Despite their pervasive influence, alternatives to feature-based approaches to categories and categorization have arisen in the cognitive and developmental domains, challenging various aspects of feature-based views (e.g., Armstrong, Gleitman, & Gleitman, 1983; Murphy & Medin, 1985; Rips & Collins, 1993). Although an examination of the full range of these challenges is beyond the scope of this chapter (for discussions, see Goldstone, 1994; Medin et al., 1993), a few of the most widely questioned assumptions of feature-based approaches are considered next.

Challenging Key Assumptions of Feature-Based Approaches. Perhaps first and foremost, feature-based approaches have typically assumed that category members cohere due to their featural similarity to one another, and that classification of a new instance as a category member is thought to depend on the degree to which the instance is featurally similar to a category prototype or to category exemplars. In short, feature-based similarity serves as an organizing principle in many feature-based views on categories and categorization. Closely related to this assumption is the view that categories essentially reflect lists of features, thus implying that features are independent entities, as noted previously.

The assumption that feature-based similarity grounds categories and categorization has been called into question on numerous grounds, most notably by research showing its flexibility (for reviews, see Goldstone, 1994; Medin et al., 1993). For instance, research has shown that variations in the directionality of similarity comparisons may alter perceptions of featural similarity (e.g., Holyoak & Gordon, 1983; Snuf & Gao, 1983; see also Medin et al., 1993). Research on Tversky's (1977) contrast model, which views similarity between two objects as a weighted function of their common and distinctive features, has also shown the fluidity of judged similarity as a function of factors such as the nature of the task and stimuli being judged, as well as the scope of the stimulus context (e.g., Gati & Tversky, 1984). Several cognitive and social-cognitive models of categorization similarly recognize that featural similarity is not fixed (e.g., Nosofsky, 1987; E.R. Smith & Zarate, 1992; see also Higgins, 1996). These models posit that similarity depends on perceivers' allocation of attention to particular stimulus dimensions, which in turn depends on, for example, the context.

The malleability of feature-based similarity suggests that features may not be the sole building blocks of categories, nor the sole basis of categorization (e.g., Medin, 1989; Murphy & Medin, 1985; see also Goldstone, 1994). How so? If the degree to which members of a category are featurally similar and the degree to which a new instance is featurally similar to members of a category are not fixed, this implies that, under any given set of conditions, there must be knowledge extending beyond features that help provide coherence to categories and that serve as a basis for categorization, respectively. Such knowledge helps to define or constrain the basis for
determining what makes members of a particular category similar to one another, and what instances should be classified as members of a category (Murphy & Medin, 1985).

Indeed, a critical problem with featural similarity as an organizing principle is that it is unclear how the particular set of features that makes members similar, or that determines categorization, is decided (Medin, 1989; Murphy & Medin, 1985; Medin et al., 1993). There is an infinite number of features that make two stimuli similar or dissimilar to one another, or a new instance similar or dissimilar to the members of a category. As such, the assumption that categories reflect feature lists, and that categorization is based on some function of the number of shared and unshared features between an instance and a category, falls short because knowledge beyond features is needed to constrain what features are relevant to category membership (e.g., Murphy & Medin, 1985). Perhaps the clearest examples of this shortcoming are categories that group together instances that share little or even no obvious featural similarity, such as the category for "things to take out of one's house in case of a fire," which might include pets, money, jewelry, and photo albums as members (example cited in Murphy & Medin, 1985; see also Barsalou, 1983; Medin, 1989).

Feature-based approaches, and the assumption often inherent to them that feature-based similarity grounds categories and categorization, have also been challenged by research documenting the nonredundance of ratings of similarity and categorization judgments (Rips & Collins, 1993; for a review, see Goldstone, 1994; cf. Smith & Stolman, 1994). This work has shown that judgments of how similar an instance is to a category versus judgments of the likelihood that the instance is a member of the category are differentially influenced by variations in a given stimulus dimension. If similarity ratings and categorization judgments diverge, this implies that similarity, defined in featural terms, is not likely to be the only basis for categorization. If it were, judgments of how similar an instance is to a category prototype or exemplar should go hand in hand with judgments of the likelihood that the instance is a member of the category.

In essence, challenges to feature-based similarity as the sole organizing principle of categories and categorization counter the idea that feature lists adequately capture the content and structure of categories. This idea has, in fact, been explicitly called into question by various strands of theorizing and research in both the cognitive and developmental literatures (e.g., Armstrong et al., 1983; Barsalou, 1983; Carey, 1985; Keil, 1989; Medin, 1989; Murphy & Medin, 1985; Resch, 1978). For example, Armstrong et al. (1983) argued that knowledge about a given entity (e.g., bird) is likely to reflect not only a list of features associated with the entity (e.g., has wings, able to fly), but also relationships between these features (e.g., birds can fly because they have wings).

Taken as a whole, challenges to feature-based approaches tend to suggest that the glue that holds category members together may involve more than shared features among them, and that the basis on which categorization judgments are made may involve more than the matching and mismatching of such features. In fact, the mounting evidence against features as the sole basis of categories and categorization coincides with the growing recognition of a potential role for theory-based forms of knowledge in cognitive and developmental work (for recent reviews, see Komats, 1992; Medin, 1989).

Theory-Based Approaches as an Alternative. Although there is variation in the precision with which theories are defined, and in how directly theory-based processes in categorization are addressed, theory-based approaches in the cognitive and developmental domains maintain that people's lay theories about the world are fundamental to how categories are represented and used. Unlike feature-based views, which tend to treat features as unrelated units, such approaches argue that explanatory relations among category features exist and are what make a category coherent (e.g., Medin, 1989; Murphy & Medin, 1985; Rips, 1991). Although some have argued for merely treating theoretical relations as kinds of features (e.g., Goldstone, 1994), such a resolution continues to imply that categories are grounded simply in features, rather than also, potentially, in the theoretical mortar binding the featural bricks of categories together (Armstrong et al., 1983; Medin, Wattenmaker, & Hampson, 1987; see also Gentner, 1989; Medin et al., 1993; Markman & Gentner, 1993). Put differently, it is argued that recognizing that theoretical relations may exist is 'to concede that attributes may have a complex internal structure...this internal structure means that one is working with more than a list of simple attributes and that constraints and explanatory power will derive from this rich structure' (Murphy & Medin, 1985, p. 297).

Support for theory-based approaches cuts across a wide range of domains and has been discussed in detail elsewhere (see Murphy & Medin, 1985). As alluded to before, particularly notable is evidence for the fluidity of featural similarity, which various researchers suggest implies that perceivers must bring knowledge extending beyond features to bear on judgments of similarity and categorization, and thus that categories and categorization are not simply feature-based (e.g., Goldstone, 1994; Medin et al., 1993). It is worth noting that much of the
8. THEORIES ABOUT SIGNIFICANT OTHERS

Evidence for the malleability of featural similarity has involved judgments of the similarity between a pair of stimuli, rather than judgments of the featural similarity between a stimulus and stored category, with the latter speaking more directly to the possible role of theory-based forms of knowledge in categorization.

However, some direct evidence for the influence of theories on categorization judgments also exists. For instance, developmental research has shown that children's theories about category membership may at times override the perceptual (i.e., featural) similarity of a stimulus to stored, category information as the primary basis of their categorization judgments (Gelman & Markman, 1986; for reviews, see Carey, 1985; Keil, 1989; cf. Jones & Smith, 1993). Research on category learning also points to a role for theories in categorization (Winston et al., 1994). In this work, participants were shown two sets of children's drawings, which were either given neutral labels ("Group 1" vs. "Group 2") or theoretical labels ("drawn by creative children" vs. "drawn by noncreative children"). They were then asked to categorize a set of new drawings, providing a rule for each of their categorization decisions. Category learning varied depending on whether a theory had been activated. For instance, whereas the categorization rules of the neutral-label group referred mainly to concrete features, those of the theoretical-label group referred to relations between these features and more abstract ones that reflected previously assessed lay theories about what drawings done by creative versus noncreative children should look like.

As a final example, cognitive work on analogical reasoning fits the notion that categorization may at times be theory-based (for recent reviews, see Gentner & Markman, 1997; Holyoak & Thagard, 1995). Specifically, some researchers have argued that relational (vs. attribute) similarity between stored knowledge and a stimulus may trigger the retrieval and use of the knowledge to interpret the stimulus (Gentner & Markman, 1997; see also Gick & Holyoak, 1980, 1983; Holyoak & Thagard, 1995; Markman & Gentner, 1993). In other words, structural "parallels" rather than simply attribute-based matches may at times fuel categorization judgments (Gentner & Markman, 1997).

Social-Cognitive Work

Social-cognitive support for theory-based approaches also takes various forms. First, as described, some of the earliest contributions to social cognition suggested that perceivers' impressions of others reflect unified configurations or Gestalts rather than sets of separate elements or features (Asch, 1946). As noted, a major idea emerging from this work is that meaningful relations exist among the features (i.e., trait attributes) associated with an entity, a notion also embodied in the concept of implicit personality theory (e.g., Bruner & Tagui, 1954; Schneider, 1973). Beyond this, early work on attribution theory also reflected Gestalt influences, as seen quite clearly in Heiderian views on the formation of causal units composed of persons and acts (Heider, 1958).

More recent work in the Gestalt tradition has explicitly conceptualized relations between traits as theoretical in nature. Specifically, research has shown that one way people make sense of incongruent information about others is to impose a theory—that is, call on knowledge to serve expressly as an explanation—for why the incongruence exists (Asch & Zuckier, 1984; see also Hastie, Schroeder, & Weber, 1990; Kunda, Miller, & Claire, 1990; Park, DeKay, & Kraus, 1994). Along the same lines, it has been argued that perceivers tend to integrate inconsistencies in their experiences into internally coherent narratives or stories, which similarly implies that perceivers may think in terms of explanatory relations (e.g., Baumeister & Newman, 1994; Murray & Holmes, 1993, 1994, 1996; Pennington & Hastie, 1992, 1993; Read & Miller, 1993; Thagard, 1989; Zuckier, 1986).

Recent work on person models also has the flavor of a theory-based approach (Park et al., 1994; see also Wittenbrink, Park, & Judd, 1998). These researchers argue that perceivers' impressions of others resemble stories or narratives consisting of a central concept that serves as an organizing structure, in which pieces of information about others are linked in ways consistent with the central concept. More specifically, they demonstrate that perceivers may extract different central concepts from an identical pool of target information, which leads to the formation of distinct person models that, while having certain target attributes or features in common, nonetheless imply, overall, very different portraits of the same target (Park et al., 1994). Thus, this work fits the idea that conceptions of others may include linkages among the attributes associated with them, some of which may be explanatory in nature, that transform the meaning of attributes considered in isolation.

In the same spirit is recent research conceptualizing implicit personality theories as person types (Sedikides & C.A. Anderson, 1994; see also Andersen & Klatsky, 1987; C.A. Anderson & Sedikides, 1991; Cantor & Mischel, 1977, 1979). The main argument of this research is that perceivers often conceptualize others in terms of person types, which are composed of causally linked traits. While acknowledging the possibility of
various other types of glue bonding traits together (e.g., evaluative), this research has explicitly shown that causal connections are one type of link between traits. In this regard, this work speaks quite directly to the idea that representational content and structure involves causal or explanatory forms of knowledge.

Social-cognitive work on analogy also exists. This research offers rather direct support for theory-based approaches by demonstrating that explanations are stored in memory, and are capable of being retrieved and used (Read & Cesa, 1991; see also Schank, 1982, 1986). Specifically, it found that retrieval of a previously generated explanation was facilitated when participants were faced with the task of explaining a new, unexpected event similar in causal structure to the one for which the explanation was originally constructed (Read & Cesa, 1991). Thus, stimulus cues reflecting causal relations helped trigger and render applicable the stored explanation for the prior event.

In a similar vein, recent research has argued that stereotype representations include knowledge about cause-effect relations, and that this knowledge may influence the processing of stereotype-relevant information (Wittenbrink et al., 1997; Wittenbrink, Hilton, & Gist, 1998; see also Kunda & Thagard, 1996; Shoda, Mischel, & Wright, 1989). This work focused on perceivers' theories about the cause of African-Americans' socioeconomic disadvantage (Wittenbrink et al., 1997). Specifically, scoring high on the Modern Racism Scale (McConahay, Hardée, & Batts, 1981) were thought to hold the cause-effect belief that African-Americans are responsible for their disadvantage, whereas those scoring low were thought to hold the belief that African-Americans are the victims of structural disadvantages. In two studies, the two participant groups varied in their perceptions of responsibility (i.e., causality) in a dispute involving an African-American target in a manner reflecting the nature of the cause-effect belief each held (Wittenbrink et al., 1997, Experiments 1 & 2). In a third study, the impact of participants' cause-effect beliefs on perceptions of causality was found only when the pool of to-be-interpreted information contained cause-effect cues that matched these beliefs – that is, when the information was presented in a way that facilitated versus inhibited construing causal structure among them (Wittenbrink et al., 1997, Experiment 3). Overall, this research suggests that matches between cause-effect relations (i.e., theories) stored in memory about a stereotype with cues in the stimulus environment that embody these relations contributed to the activation and use of the stereotype (Wittenbrink et al., 1997).

In a somewhat different vein, theory-based notions can also be seen in connectionist models of social perception (Read, Vanman, & Miller, 1997; see also Miller & Read, 1991; Read & Miller, 1993; cf. Kunda & Thagard, 1996; E.R. Smith, 1996, 1998). Reflecting basic connectionist assumptions, these approaches generally argue that mental models of persons and social groups contain interconnected pieces of knowledge about others' traits, goals, beliefs, and so on, of some of which are explanatory. Reeminent of Gestalt principles, these approaches view social perception as conditional – that is, people perceive the characteristics, emotions, and behaviors of others in the context of one another, as a configuration, with the aim of achieving a coherent understanding of the social world. That is, each element or piece of knowledge about a person or group only has meaning in the context of all the other elements to which it is connected (Read et al., 1997).

Conceiving of individuals in such conditionalized terms lies at the heart of the long-standing cognitive-social model of personality (Mischel, 1968, 1973, 1990; Mischel & Shoda, 1995). Countering widely held assumptions about the stable and traitlike nature of personality, this model takes an interactionist approach, positing IF-THEN relations as the basic units of personality. IFs refer to situations or conditions, and THENs refer to the responses that are exhibited in them. The notion that personality is composed of IF-THEN relations raises the possibility that perceivers may conceive of other individuals and social groups in terms of such IF-THEN relations, recognizing that what others think, feel, and do may be contingent on the situation (Shoda & Mischel, 1993).

The idea that perceivers may possess knowledge about the IF-THEN relations of others served as a primary basis for the theory-based approach to significant-other representations and transference presented later. To provide a conceptual framework for this approach, the social-cognitive model of transference is first described (Andersen & Glassman, 1996; Zhen & Andersen, 1999). The main thrust of this model is that the phenomenon of transference can be understood in terms of the activation and use of significant-other representations in interpersonal perception and relations.

THE SOCIAL-COGNITIVE MODEL OF TRANSFERENCE

The concept of transference refers to the idea that past experiences with and assumptions about significant others may resurface in relations with new others (Freud, 1912/1958). Said differently, transference entails the reemergence of aspects of past relationships in present
ones. This general notion can be seen in various psychological domains, among them the clinical, personality, developmental, and social literatures (e.g., Baldwin, 1992; Bartholomew & Horowitz, 1991; Bowlby, 1969; Bugental, 1992; Feeney & Noller, 1990; Greenberg & Mitchell, 1983; Guidano & Liotti, 1983; Hazan & Shaver, 1987; Horney, 1939; Horowitz, 1991; Luborsky & Crits-Christoph, 1990; Rogers, 1951; Safran & Segal, 1990, Shaver & Rubenstein, 1980; Sullivan, 1953), and forms the basis of the social-cognitive model of transference (Andersen & Glassman, 1996).

Basic Assumptions

According to this model, mental representations of significant others are stored in memory, housing considerable knowledge, including characteristics of significant others, as well as feelings, motivations, and roles experienced in relation to them (for a review, see Chen & Andersen, 1999). Relative to representations of stereotypes, traits, and nonsignificant others, significant-other representations have been shown to be associatively richer, in that perceivers spontaneously list a greater number of idiosyncratic features to characterize them (Andersen & Cole, 1990, Studies 1 & 2). Further, significant-other features have been shown to be distinctive compared with features of related stereotypes and traits, meaning they are less likely to be associated with stereotypes and traits compared with the reverse (Andersen & Cole, 1990, Study 2). Featural richness and distinctiveness suggest the greater structural complexity of significant-other representations relative to the other representations (Tversky, 1977; see also Andersen & Klatzky, 1987; Sull & Gaelick, 1983).

Significant-other representations designate single individuals, rather than categories or types of individuals, and are thus exemplars (e.g., E.R. Smith & Zarate, 1990, 1992). As alluded to in prior sections, research has shown that exemplars are activated and used in social perception, just as are other constructs designating, for example, traits or stereotypes (e.g., Andersen, Klatzky, & Murray, 1990; Bargh & Thein, 1985; Higgins, Robles, & Jones, 1977). As such, transference can be understood in terms of the activation and use of a significant-other representation to interpret and respond to a new person in representation-derived ways.

Chronic and Priming Sources of Accessibility

In viewing transference in terms of the activation and use of significant-other representations, it is important to consider principles of knowledge accessibility and use. In line with prevailing social-cognitive models (e.g., Bargh, 1997; Higgins, 1989, 1996), the triggering of significant-other representations depends on chronic and transient sources of accessibility. Chronic sources arise from the frequent past activation of a representation. They are often referred to as the chronic accessibility of a representation, or its chronic readiness to be brought to bear on social perception. Research has documented the chronic accessibility of significant-other representations by showing that participants made stronger inferences about a target person based on a significant-other representation, compared with various control representations, even in the absence of transient sources of accessibility (Andersen, Glassman, Chen, & Cole, 1995).

One possible transient source of accessibility lies in cues in the environment that temporarily increase accessibility levels. This source is widely referred to as priming (e.g., Higgins et al., 1977; Sull & Wyer, 1979, 1980). Research has shown that priming contributes to the activation and use of significant-other representations (Andersen et al., 1995, Study 1). Specifically, although participants made stronger inferences about a target person based on a significant-other representation relative to various control representations, even in the absence of priming, inferences derived from participants' significant-other representations were still greater when these representations were primed. This finding is similar to ones from research on trait constructs showing that chronic and transient sources of accessibility may combine to trigger the use of trait constructs, as seen in the enhanced strength of trait-consistent inferences about new others in the presence of both sources versus either source alone (e.g., Bargh, Bond, Lombardi, & Tota, 1986).

Applicability Sources of Accessibility

The contribution of chronic accessibility and priming to the activation and use of a stored representation to interpret and respond to a stimulus have typically been examined using stimuli to which the representation is somehow applicable or relevant (for reviews, see Hardin & Rothman, 1997; Higgins, 1996). For example, research has used ambiguous targets of perception who have features that are equally similar to an activated representation, as they are to some alternative one (e.g., Higgins et al., 1977). Other research has used vague targets who have enough features to render a representation's use appropriate but not inevitable (e.g., Sull & Wyer, 1979). Beyond this, the potential for a to-be-interpreted stimulus to contribute to the activation and use of a representation has not been widely discussed (for exceptions, see
In terms of transference, cues in a new person that match knowledge stored about a significant other should enhance the activation and use of the representation of this other to interpret and respond to the person. Research has found direct evidence for such applicability-based triggering of transference, with applicability arising from a target person’s featural resemblance to a perceiver’s significant other (Andersen et al., 1995, Study 2). Specifically, participants’ inferences about the target based on a significant-other representation were stronger in the presence of applicability relative to its absence.

The Basic Transference Paradigm

To empirically examine the transference phenomenon, a combination of idiographic and nomothetic methods have been used. In the typical research paradigm, participants idiosyncratically generate features to describe a significant-other representation, as well as various control representations, in a pretest session. For example, they may generate descriptive features such as "Ramona is smart" or "The nerd studies hard." In an ostensibly unrelated session, which takes place at least 2 weeks later, participants’ idiosyncratic information is used as stimuli in a nomothetic experimental design.

In the learning phase of the experimental session, participants “meet” one or several fictional target persons by being presented with a series of descriptive features about them. For each target, some of these features are ones that participants had generated in the pretest session to describe a given representation. One target is always characterized in part by features that participants had generated earlier to describe a significant other. Thus, applicability-based triggering of significant-other representations is operationalized in terms of feature-based cues in a target person that overlap with features associated with a significant-other representation. These cues serve as feature-based triggers for the activation and use of the representation to interpret the target, and are viewed as an analog for the cues that perceivers confront in the context of actual, face-to-face encounters with others (Chen et al., 1999).

After the learning phase, participants are asked to complete a recognition-memory test about each target; they are presented with a series of descriptive sentences and asked to rate their confidence about having learned each one earlier (see Cantor & Mischel, 1977). Some of these are ones that were actually presented earlier to characterize a given target, and some are ones that were not actually presented earlier but that were derived from the representation to which the target corresponds. For the target corresponding to the participant’s own significant other, recognition-memory confidence ratings for representation-derived descriptors that were not presented earlier serve as an index of transference—that is, of inferences about the target that are derived from the participant’s representation of a significant other (e.g., Andersen & Baum, 1994; Andersen & Cole, 1990; Andersen et al., 1995, Andersen, Reznik, & Mansella, 1996; Baum & Andersen, 1999; Chen et al., 1999; Hinkley & Andersen, 1996).

Overall, a growing body of evidence supports the social-cognitive model of transference. However, as in the broader social-cognitive literature, featural assumptions have been embedded in research on transference. Yet such assumptions may not pertain equally to all mental representations; instead, theories and theory-based processes may better characterize some representations and their use. In particular, this may be the case for representations of significant others, which is the central argument of the theory-based approach to significant-other representations and transference presented next.

A THEORY-BASED APPROACH TO SIGNIFICANT-OTHER REPRESENTATIONS AND TRANSFERENCE

The present approach draws in part from the cognitive-social model of personality and social perception (Mischel & Shoda, 1995; Shoda & Mischel, 1993). As noted, this model takes an interactionism approach, arguing that IF-THEN relations are basic to both personality and social perception. It’s refer to situations and THENs refer to the responses exhibited in them. In terms of personality, the model contends that the nature of IFs depends on the acquired meaning of a given situation for an individual (Mischel & Shoda, 1995)—that is, it depends on the individual’s construal of the situation. Thus, for example, the IF designating a given situation may be construed as “at a party” for one person, whereas for another it is seen as “when meeting new people.” Regardless of how a situation is construed, the model maintains that situations activate particular sets of cognitive-affective units (e.g., feelings, goals, expectancies) that represent the psychological mediators of the IF-THEN relations that constitute personality.
8. THEORIES ABOUT SIGNIFICANT OTHERS

Applied to social perception, the model suggests that perceivers may perceive others in terms of IF-THEN relations, viewing others' feelings, thoughts, and behaviors as contingent on the situation (Shoda & Mischel, 1993). The present approach extends this basic idea by proposing that knowledge about the IF-THENs of others may at times represent theories—that is, IFs may serve expressly as explanations for THENs. What kind of IF-THEN knowledge might be considered a theory? It is argued that knowledge about the observable situations (e.g., "in school") in which another person exhibits particular responses is not especially explanatory because it simply reflects prior empirical observations of the person. At times, however, perceivers may take an explanatory step beyond simply making observations about another person to making inferences about the psychological situations that produce his or her responses—that is, the psychological contingencies of these responses. Psychological situations refer to internal, private states such as feelings, wishes, or motives. In the present approach, IF-THENs in which the IFs refer to the psychological situations that elicit a person's THENs are seen as theories because, essentially, they reflect knowledge about the most proximal causes of the person's responses. That is, they reflect conceiving of the IFs of the person not simply in terms of the observable situations that, according to the cognitive-social model, activate psychological mediating variables, but rather directly in terms of the psychological mediators linking observable IFs and THENs.

The central argument of the current approach is that perceivers are especially likely to conceptualize the IFs of their significant others at the level at which they perceive these situations are construed by these important others. In terms of the previous example, rather than merely viewing the IF as an observable situation such as "at a party" or "when meeting new people," for a significant other, perceivers are particularly likely to view the situation in terms of the psychological situation of the person (e.g., "when feeling inadequate"). Thus, rather than simply conceiving of IFs in terms of the observable situations that presumably activate particular psychological situations, which in turn elicit particular THENs, perceivers are especially likely to understand the IFs of significant others directly in terms of these psychological situations.

In summary, the present approach defines theories in terms of IF-THENs, in which the IFs represent inferences about the psychological situations of another person believed to produce his or her responses or THENs (e.g., "IF Ramona feels insecure, THEN she doesn't talk in class"). Put another way, it proposes that perceivers' conceptions of others may include not only unrelated units of factual knowledge, but also psychological-state theories that take the form of short, explanatory IF-THENs, each composed of a psychological IF and a THEN to which it is causally linked. The overriding argument is that perceivers are especially likely to make sense of significant others in terms of such psychological-state theories. This implies that the content and structure of significant-other representations, as well as the use of these representations in transference, are likely to involve such theories. These assertions are bolstered by various strands of existing evidence, which are considered next.

The Case for Theories about Significant Others

As reviewed earlier, a growing body of social-cognitive work suggests that perceivers' conceptions of others involve explanatory or causal knowledge (e.g., Sedikides & C.A. Anderson, 1994; Wittenbrink et al., 1997). Such knowledge implies prior causal thinking on the part of perceivers. But why would perceivers be especially likely to engage in such thinking about their significant others? Causal Thinking About Significant Others. A wide range of theory and evidence suggest that, given the relevance of significant others to one's emotional and motivational outcomes, perceivers are likely to be especially willing to exert effort to make sense of these important individuals (see Andersen et al., 1997, 1998). Indeed, it is widely agreed that having one's outcomes dependent on others, such as is the case with significant others, engenders more effortful processing of information about these others than if there were no such outcome dependency (e.g., Erber & Fiske, 1984; Fiske & Neuberg, 1990). This desire to know significant others is likely to involve coming up with explanations for the behaviors of significant others (e.g., Fletcher & Fincham, 1991). Such explanations should enable perceivers to better regulate their own thoughts, feelings, and behaviors (Andersen et al., 1997, 1998). Said differently, figuring out why significant others think, feel, and act as they do is likely to serve a self-regulatory function, informing perceivers about how to feel, what goals to pursue, what to expect and hope for, and so on in relationships with these important others. Indeed, considerable evidence supports the basic idea that significant others are relevant to self-regulation, essentially demonstrating that the self is experienced in part in relation to significant others (e.g., Aron, Aron, Tudor, & Nelson, 1991; Baldwin, 1992; Downey & Feldman, 1996; Hinkley & Andersen, 1996; Moretti & Higgins, 1999).

Direct evidence that perceivers engage in causal
thinking about significant others can be found in close relationships work (for a review, see Chen & Andersen, 1999). For example, research examining attributional tendencies and relationship satisfaction indicates that relationship partners engage in attributional thinking about each other and their relationship (e.g., Fletcher, Fincham, Cramer, & Heron, 1987; Miller & Read, 1991; Read & Miller, 1989), and that the types of attributions that partners make about each other may be associated with satisfaction in the relationship (for a review, see Bradbury & Fincham, 1990). The association found between attribution type and satisfaction has been found both when attributions are assessed using direct solicitation procedures, as well as when they are assessed spontaneously, suggesting that attributional thinking in relationships may reflect controlled, conscious processing, as well as relatively automatic processing (see Fletcher & Fincham, 1991).

Building on the attributions and satisfaction work is research examining the association between attributional tendencies and attachment style (Collins, 1996). This work suggests that individuals with different attachment styles are inclined to interpret and explain relationship events in ways consistent with their particular working models of self and other. Stated another way, it proposes that working models of the self and others entail explanatory tendencies, developed in part on the basis of past experiences with significant others (see also Downey & Feldman, 1996). For instance, this research has found that individuals with a preoccupied attachment style tend to formulate explanations for relationship events reflecting a more negative interpretation, as well as more negative views of relationship partners, than do securely attached individuals (Collins, 1996). In turn, such variations in explanatory tendencies have been shown to influence emotional responding, as well as self-reported behavioral intentions.

Still more evidence that perceivers engage in causal thinking about significant others lies in research showing that perceivers construct theories about significant others to construct and maintain positive illusions about these individuals, thus bolstering feelings of confidence in relationships with them (e.g., Murray & Holmes, 1993, 1994, 1996; Murray, Holmes, & Griffin, 1996; see also Holmes & Rempel, 1989). In more concrete terms, this research has found that, when faced with a romantic partner’s fault, people formulate theories to “encapsulate or diffuse the negativity” of the fault (Murray & Holmes, 1993, 1994). For instance, they reconstrue the fault into a virtue, in this way forging protective, explanatory linkages between the faults and virtues of their significant others.

Analogous data have emerged in research on transference (Andersen et al., 1996). Specifically, one study found that, on encountering a target person characterized by some negative descriptors associated with a positively evaluated significant other, participants responded to these descriptors with especially positive facial affect while encoding them. The inconsistency of the negative features in relation to the overall positive affective tone of the significant-other representation apparently led perceivers to transform the negative into a positive response — arguably a reflection of the influence of their positive theories about their significant others. The construction of theories that cast significant others in a favorable light may reflect a basic motivation to remain attached to these important others (Andersen et al., 1997). That is, they may reflect motivated attempts to bolster a perceiver’s sense that a significant other loves him or her (see also Murray & Holmes, 1993, 1994, 1996).

Overall, considerable research indicates that perceivers engage in causal thinking about significant others. Such thinking suggests that significant-other representations contain more than unrelated featural units of knowledge; they may also include knowledge reflecting causal relations — essentially, theories. But what evidence is there that such knowledge is especially likely to reflect psychological-state theories?

**Psychological-State Theories About Significant Others.** On perhaps the most fundamental level, several lines of research have shown that representations of significant others contain more knowledge about internal, private states (e.g., thoughts, feelings) than do representations of various others, such as nonsignificant others (e.g., Andersen et al., 1998; Prentice, 1990). This finding bolsters the argument that perceivers are particularly likely to have psychological-state theories stored about their significant others insofar as it indicates that they have an especially large pool of psychological-state knowledge to draw from to construct such theories for these important individuals.

Also relevant to this argument is recent research in which close relationships have been defined in terms of the inclusion of the other in the self (Aron & Aron, 1996; Aron et al., 1991). This research suggests that, to the extent a significant other is incorporated as part of the self, there should be greater similarity in the nature of cognition about the self and this other (Aron & Aron, 1996). For example, actor/observer differences in perspective are likely to be diminished (e.g., Andersen et al., 1998; Aron et al., 1991), resulting in greater similarity in the attributions made for one’s own behaviors and those made for the other’s behaviors (Nisbett, Caputo,
8. Theories About Significant Others

Legant, & Marecek, 1973; see also Prentice, 1990; Sande, Goethals, & Radloff, 1988; Zuroff, 1982; cf. Taylor & Koivumaki, 1976). Similarity in cognition about the self and significant others implies some degree of similarity in representations of self and significant others. Although recent work has shown that self-representations contain more knowledge about private aspects (e.g., thoughts, motives) than do significant-other representations, it has also shown that significant-other representations resemble self-representations more so than nonsignificant-other representations in several ways (Andersen et al., 1998). Specifically, significant-other representations are more similar to the self than to nonsignificant-other representations in terms of the distinctiveness of knowledge about private aspects relative to public aspects (e.g., behaviors), as well as in terms of the accessibility of knowledge about private aspects.

The notion that there is some degree of similarity in knowledge about the private aspects of the self and significant others is pertinent because, coupled with research indicating that perceivers generally view cognitive/affective information to be more informative than behavioral information and that this is especially true for the self versus others (Andersen & Ross, 1984), it suggests that information about significant others’ internal, private experiences is likely to be seen as especially informative, just as for such information about the self. Overall, this coheres with the research noted earlier suggesting that perceivers are likely to be especially motivated to engage in effortful thinking about their significant others (e.g., Erber & Fiske, 1984), and extends it by suggesting that inferences about others’ inner, psychological states are particularly likely to satisfy perceivers’ desire to know their significant others, just as one’s own inner workings and experiences are thought to best reflect one’s “true” self (Andersen & Ross, 1984; see also Andersen et al., 1998; Johnson & Boyd, 1995).

Overall, the bulk of existing evidence suggests that the knowledge perceivers store about the psychological states of significant others is substantial. Although it is less than what is stored about the self, it is nonetheless comparable in various ways to such knowledge about the self and is more available than such knowledge stored about nonsignificant others. Together, the clear availability of psychological-state knowledge about significant others, and the informational value perceivers place on such knowledge, along with evidence suggesting that perceivers are especially motivated to know their significant others, suggest that perceivers are particularly likely to draw on psychological-state knowledge to explain the behaviors of their significant others. In short, although it may often be sufficient to know that an acquaintance or a “nerd” is smart and skinny, it is unlikely that such simple, feature-based conceptions would suffice in making sense of a significant other. Instead, the present approach proposes that perceivers are especially likely to know their significant others in terms of psychological-state theories.

Evidence for a Theory-Based Approach to Significant-Other Representation and Transference

Research examining this approach has focused on three main hypotheses. First, the argument that perceivers are especially likely to conceptualize their significant others in terms of psychological-state theories suggests that such theories should be more available in memory about significant others relative to various control targets. Second, it suggests that these theories should be especially accessible from memory. Finally, if significant-other representations are especially likely to involve psychological-state theories, then cues in a new person that reflect such theories should be especially likely to trigger transference. Thus, it was hypothesized that applicability-based triggering involving psychological-state theories would serve as a particularly powerful source for the activation and use of significant-other representations to interpret and respond to new others—that is, transference.

Three studies have been conducted to examine these hypotheses. In each, “theories” referred to participants’ inferences or beliefs about the psychological situations (i.e., IFs) that explain the responses (i.e., THENs) of others. Also, in each study, participants were asked to idiosyncratically identify a significant other (e.g., romantic partner) and, as control conditions, a stereotype (e.g., jock) and nonsignificant other (e.g., friend of a friend).

Theory-Based Content in Significant-Other Representations. A first study tested the hypothesis that psychological-state theories are especially likely to be stored as part of significant-other representations (Chen, 1997, Study 1). To do so, this study examined two kinds of IF-THEN knowledge. In the first, the IFs referred to observable situations (e.g., “at a party”). To assess this knowledge, participants were given a list of sentence prompts for each representation. Each prompt designated an observable situation (i.e., IF), and participants were asked to complete the sentence if they felt they knew the most likely response (i.e., THEN) of the significant other, stereotype, or nonsignificant other in the observable situation. Thus, these IF-THENs assessed the number of empirical observations participants have stored about
each representation (e.g., "IF at a party, THEN Amy talks to everyone"). In the second kind of IF-THEN knowledge, the IFs referred to psychological situations. To elicit such IF-THENs, participants were asked to consider each IF-THEN observation they had listed and provide an explanation for the observation if they felt they had one. More specifically, they were asked to provide a BECAUSE phrase (i.e., explanation) following each observation they felt they could explain (e.g., "IF at a party, THEN Amy talks to everyone BECAUSE..."). Participants' explanations were then coded as ones reflecting psychological states, such as thoughts, feelings, or goals (e.g., "...because he wants to get ahead" or "...because she thinks others don't like her") or ones that did not reflect such states (e.g., "...because he's in school" or "...because it's her job"). BECAUSE phrases referring to psychological states served as an indirect measure of psychological-state theories, in that they reflect participants' inferences about the psychological situations that elicit the responses exhibited in particular, observable situations. Put another way, those BECAUSE phrases reflect what participants perceive as the psychological mediators between observable situations (i.e., observable IFs) and the responses (i.e., THENs) that occur in them. In this regard, these psychological-state explanations represent participants' perceptions of the most proximal IFs of THENs.

Participants were expected to provide more psychological-state theories to explain the IF-THEN observations listed for their significant other relative to their stereotype and nonsignificant other. This would support the hypothesis that perceivers are especially likely to go beyond simply making IF-THEN observations about their significant others to formulating theories that reflect beliefs about the psychological situations that mediate these observable, IF-THEN relations. Although the difference between the significant-other representation versus the control representations was expected to be largest in comparisons involving such theories, participants were also expected to list more BECAUSE phrases overall for their significant other relative to control representations, regardless of whether these phrases were based on information about psychological states. This prediction reflects the general idea that perceivers are especially likely to engage in causal thinking about their significant others.

Several findings emerged from this study. First, participants listed more IF-THEN observations for their significant other relative to the control targets, a finding that fits prior evidence indicating that perceivers are more likely to view familiar versus less familiar others in situation-specific terms (e.g., Prentice, 1990; Zuroff, 1982). The situations examined in the earlier work were also observable ones, such as "in class" or "with family" (Prentice, 1990, p. 383). Extending this earlier work, a greater proportion of participants' IF-THEN observations about their significant others relative to those about the control targets was backed by an explanation (i.e., BECAUSE phrase). Moreover, the explanations given for IF-THEN observations about significant others were especially likely to refer to psychological states. In short, participants were especially likely to provide explanations about their significant others, and these explanations were particularly likely to reflect psychological-state theories.

Despite strong support for this study's availability hypothesis, one limitation of this study was the nomothetic nature of the sentence prompts that were used. These prompts were used to measure the number of IF-THEN observations participants have stored, controlling for the observable situations considered. This also allowed an assessment of differences across representations in the degree to which participants' IF-THEN observations were backed by an explanation—based on the same set of situations. To address this limitation, a second study conceptually replicated the first using a fully idiographic procedure in which participants were asked to freely generate psychological-state theories for each representation (Chen, 1997, Study 2). The findings of this study converged with those of the first. Namely, participants spontaneously generated more psychological-state theories (i.e., IF-THENs with psychological IFs) to describe their significant other versus their stereotype and nonsignificant other. Taken together, these data clearly suggest that the content of significant-other representations is particularly likely to include psychological-state theories defined in IF-THEN terms.

Theory-Based Structure in Significant-Other Representations. If perceivers tend to think about their significant others in terms of psychological-state theories, it was hypothesized that the linkages connecting stored knowledge about the psychological IFs and THENs of these individuals should be especially strong—that is, activation should spread relatively quickly across them. Thus, the component parts of psychological-state theories should be retrieved from memory in tandem quite readily. Put another way, psychological-state theories about significant others should be particularly accessible, suggesting that perceivers should find it especially easy to retrieve the psychological IFs that explain the THENs of these important others.

To test this hypothesis, Chen (1997, Study 2) had participants generate psychological-state theories to describe each of the three representations on a computer,
which recorded their generation latencies. More specifically, participants generated each theory by first typing in a response (i.e., THEN) true of the individual (significant other or nonsignificant other) or stereotype being described, followed by the psychological situation (i.e., psychological IF) that they believe produces the response.

As predicted, participants showed significantly greater ease in coming up with psychological IFs to explain THENs (i.e., shorter average latency within IF-THEN units) about their significant other relative to the control representations. This was true even when controlling for the average latency between IF-THEN units — verifying the distinctiveness of the average latency within IF-THEN units, and bolstering the interpretation that the latter reflects the existence of IF-THEN units in memory that, for significant-other representations, possess particularly strong linkages. These results support the accessibility hypothesis and suggest that psychological-state theories about significant others take the form of especially tightly linked IF-THEN units, each composed of a psychological state and the piece of data it explains. Said differently, they indicate that the theory-based content of significant-other representations has structural implications.

A Theory-Based Approach to Applicability-Based Triggering in Transference. Applicability’s role in the activation and use of a representation can be understood in learning terms (Bruner, 1957b; Higgins, 1989, 1996; see also Bargh, 1997; Hardin & Rothman, 1997; see also Lewicki, 1986; E.R. Smith, 1990). Specifically, a representation’s applicability to a stimulus depends on the stimulus patterns that have most frequently led to the activation of the representation in the past. Thus, applicability is based in part on a representation’s activation history. For instance, some research has found that the effect of priming gender-stereotypic traits on ratings of male and female targets on these trait dimensions was greater when there was stereotype-consistent fit between the primed trait and the gender of the target (Banaji, Hardin, & Rothman, 1993). Presumably, target gender enhanced the applicability of the corresponding gender stereotype because of the gender-consistent nature of the activation history of gender stereotypes (see also Higgins et al., 1977; Leyens, Zebryt, & Corneille, 1996; Oakes, Turner, & Haslam, 1991).

By and large, the stimulus patterns or cues thought to contribute to the activation and use of a representation have been conceptualized in featural terms. That is, applicability has been widely defined in terms of the match between features in a stimulus and feature-based knowledge stored as part of a representation, with greater match corresponding to greater applicability (Higgins, 1989, 1996; see also Chen et al., 1999), as noted in prior sections. But if perceivers conceive of their significant others in terms of psychological-state theories, the stimulus cues that should render a significant-other representation most applicable are ones that embody these theories. In other words, psychological-state theories about significant others should serve as a powerful basis for applicability’s role in triggering transference. For example, observing the IF-THEN relation "when criticized, a target person gets angry" should activate a perceivers’ representation of a significant other because the IF-THEN embodies the psychological-state theory that her significant other lashes out at others if feeling inadequate. In short, the main basis of applicability-based triggering should reflect the level at which perceivers tend to think about the entity designated by the representation (Hardin & Rothman, 1997).

A third study examined the hypothesis that psychological-state theories about significant others play a distinctly powerful role in triggering transference (Chen, 1997, Study 3). This study used a variant of the basic transference paradigm and compared the contribution of theory-based applicability to that of feature-based applicability to the activation and use of participants’ significant-other representations. More specifically, participants in this study learned about several target persons, one of whom was described by descriptors that participants had generated in a pretest session to describe a significant other. The other targets resembled various control representations. Thus, as in all research on transference, applicability-based triggering of participants’ representation of a significant other took the form of a target’s resemblance to this important other. The critical difference in this study was that applicability was based on either features or theories associated with the significant other. In the feature-based applicability condition, participants saw feature-based stimuli in the learning phase (e.g., "Terry is an intelligent person"). In contrast, participants in the theory-based condition were presented with stimuli reflecting psychological-state theories (e.g., "Terry is nice to others when he wants to get ahead").

In both conditions, participants were given a recognition-memory test about each target in which they were presented with a series of descriptors and were asked to rate their confidence about having learned each one earlier (see Cantor & Mischel, 1977). Some of these descriptors had been presented earlier to describe the target, and some had not actually been presented but were derived from the representation that the target resembled. Confidence ratings for representation-derived
descriptors that were not presented earlier were averaged as a measure of the degree to which representation-derived inferences about the target were made based on the corresponding representation. Inferences based on a significant-other representation served as an index of transference, as in prior research.

Regardless of applicability condition, participants made stronger representation-derived inferences about the target resembling their own significant other relative to the control targets—an effect that constitutes evidence for transference. In the feature-based condition, this effect replicates prior research in which applicability-based triggering of transference has always taken the form of feature-based cues in a target that matches features associated with participants' significant other. However, if perceivers tend to represent their significant others in terms of psychological-state theories, then theory-based cues in a new person should produce a stronger transference effect. Although some deviations from this strict prediction emerged, the transference effect—defined in terms of greater representation-derived inferences about the significant-other target relative to the control targets—was in fact significantly stronger in the theory-based applicability condition relative to the feature-based one. This finding supports this study's key hypothesis that theory-based applicability plays an especially powerful role in triggering transference.

Taken together, the three studies described earlier provide support for a theory-based approach to significant-other representations and transference. The content, structure, and use of representations of significant others appear to be especially likely to involve theories, defined specifically as perceivers' IF-THEN knowledge in which the IFs refer to the psychological states that elicit the THENs.

ISSUES AND FUTURE DIRECTIONS

On one level, the theory-based approach presented earlier clearly resonates with the recurrent idea in the social-cognitive literature that perceivers strive for coherence in social perception, thus viewing others in terms of configurations, organized around central themes, rather than in terms of separate elements or features. Asch's notion of central traits, and later concepts such as implicit personality theories and schemas, embody this basic idea. It is an idea that leaves room for the possibility that theoretical relations exist among features and are basic to representational content, structure, and use.

On several other levels, however, the present approach extends existing social-cognitive concepts. Among them, it goes beyond viewing features simply in terms of trait attributes, which are the focus of most social-cognitive research. More precisely, it posits not only that representational content and structure may include explanatory relations, but also that these relations may involve inferences about others' feelings, motives, thoughts, and so forth. On a more basic level, the current approach offers a concrete view of the form that theories might take in the social realm, which is a level of specificity that is absent in most social-cognitive work consistent with, or even explicitly espousing, theory-based notions.

On a different note, the present approach is also distinct in arguing that representational content, structure, and use are likely to vary as a function of the significance of the "other" designated by representations. More specifically, it suggests that to understand the content and structure of perceivers' representations of others, as well as the processes that govern the role of these representations in interpersonal perception and relations, it is critical to consider the nature of perceivers' relations with these others (see also Hardin & Conley, chap. 1, this volume). Put differently, how the self is related to others is likely to influence how perceivers go about making sense of these others, and, as a result, how knowledge about these others is mentally represented and how and when it is likely to be used. By proposing that the nature of perceivers' relations with others influences mental representations and their use, the current approach is part of an expanding literature reflecting a cross-fertilization of social-cognitive theory and close relationships work.

Finally, in a related vein, the present approach extends existing work on the principles that govern knowledge activation and use by suggesting that applicability may at times be primarily feature-based, and at other times theory-based, depending on how perceivers conceive of the entity designated by a given representation (see Chen et al., 1999). Of course, many issues have yet to be addressed. Some of these are considered next, followed by a discussion of some future directions they suggest or compel.

What Is Meant By "Theories?"

Existing theory-based approaches vary considerably in terms of how and the precision with which theories have been defined. One dimension onto which many variations fall is the globality of theories. Whereas some theory-based approaches appear to define theories as single, global, or all-encompassing explanatory principles, others suggest a lower level of analysis, at which multiple
thus far only examined theories at a lower level of
analysis, namely, in the form of IF-THEN units. Defining
theories in such IF-THEN terms fits squarely with the
increasing recognition in the social-cognitive literature
that perceivers may at times view the social world in
interactionist, IF-THEN terms (e.g., Bargh, 1997;

What Is Meant By “Theory-Based Reasoning”?
Judgmental tasks. For example, perceivers may draw on
a global theory about what makes a given object what it
is in deciding whether an instance is a member of a given
category, or in making theory-derived inferences about
the instance (see Murphy & Medin, 1985). The latter
type of theory-based process is often what is implied in
the vast social-cognitive literature on schema-driven
processing. Schemas have often been likened to theories
that guide top-down processing, whereby perceivers
make schema-derived inferences about a stimulus person.

In contrast, the present approach, which defines
theories in terms of explanatory relations among featural
bits of knowledge, is specifically concerned with the role
of theory-based processes in knowledge activation and
perceivers themselves are the IFS of their significant
others’ THENs. Just as perceivers are likely to be moti-
vated to figure out the psychological situations that elicit
their significant others’ responses, they may also be
motivated to understand the relational dynamics that
elicit these responses. Examples of such relational IF-
THENs might be “When I am feeling frustrated, my
partner soothes me” or “When I am weak, my mother
comforts me” (for a review, see Baldwin, 1992). In the
transference realm has shown that activation of a
significant-other representation leads to changes in
perceivers’ self-evaluations (Baldwin, Carrell, & Lopez,
1990). It is argued that such self-evaluative changes
occur by virtue of self/significant-other linkages stored in
memory as part of perceivers’ “relational schemas” (for
a review, see Baldwin, 1992). In the transference domain,
research has shown that when a significant-other repre-
sentation is activated and used in an encounter with a
target person, stored expectations regarding the signifi-
cant other’s acceptance or rejection are likely to color
relations with the target such that the perceiver comes to
expect the target to accept or reject him or her just as his
or her significant other does (Andersen et al., 1996).
Similarly, stored motivations to approach or distance oneself from a significant other have been shown to come into play in relations with new others when transference occurs. Expectations and motives associated with significant others reflect knowledge about relational patterns between the self and significant others, just as relational IF-THENs do. Beyond this, research has shown shifts in self-definition and self-evaluation in the context of transference (Hinkley & Andersen, 1996). That is, activation of a significant-other representation accordingly activates aspects of a perceivers self that are associated with the relevant significant other, leading the perceivers to become the self he or she is when with the other — that is, the "self-with-significant-other" (see also Ogilvie & Ashmore, 1991).

Thus far, research on transference has conceptualized self/significant-other linkages in terms of what they imply for the nature of interpersonal perception and relations. In line with the current theory-based approach, however, future research might consider the role of relational IF-THENs in triggering transference. For example, having one's concerns about a particular issue belittled or dismissed by a new acquaintance may be a relational cue that reminds one of a pattern of interaction often experienced with a significant other, thereby triggering the activation and use of the relevant significant-other representation.

The idea that relational IF-THENs may serve as triggers of transference raises the questions of when and for whom different kinds of IF-THEN cues might elicit transference. Cross-cultural differences in social perception are interesting to consider in this regard. For example, research examining the distinct ways in which Americans and Hindus explain others' behaviors suggests that Hindus may be less likely to come up with explanations that refer to internal or psychological aspects of others (Miller, 1984). In general terms, this research found that Hindus were more likely than Americans to generate contextual explanations for an actor's behavior, such as explanations referring to social norms, social roles, or interpersonal relationships (e.g., "Because she is his aunt"), rather than explanations referring to something internal about the actor (e.g., "He was interested only in himself"). Such differences in the kinds of attributions made by Americans compared to Hindus were interpreted as reflecting the emphasis placed in Indian culture on "locating a person, object, or event in relation to someone or something else" (Miller, 1984, p. 968).

This work suggests that evidence for the pronounced availability and accessibility of theories about significant others, defined specifically in terms of IF-THENs with psychological Ifs, and the especially powerful role of such psychological-state theories in triggering transference, may hold more or less strongly among perceivers across different cultures.

**CONCLUDING REMARKS**

Because much of everyday social cognition involves significant others, it is critical to consider how perceivers go about making sense of these individuals. The theory-based approach presented in this chapter, and the findings it has yielded thus far, suggest that perceivers are especially likely to make sense of significant others in terms of psychological-state theories about them. This is reflected in both the content and structure of their representations of these important others, as well as in the triggering of these representations in encounters with new others. On a broader level, this work suggests the need to explicitly incorporate a role for theories in social-cognitive work on mental representations and the processes that govern their use in making sense of the social world. Social cognition may often be theory-based, and potentially in qualitatively different yet systematic ways, about the different persons and groups that make up the perceivers social world.

**ENDNOTES**

1. Although applicability has been theoretically defined in terms of overlap between a stored representation and cues in a to-be-interpreted stimulus (e.g., Hardin & Rothman, 1997; Higgins, 1996; see also Srull & Wyer, 1979, 1980), applicability has not been examined strictly in these terms in the small body of research examining its contribution to the activation and use of stored representations. In this work, much of which has focused on the applicability of a stored, trait construct to a target stimulus, the degree of applicability or match of the trait construct to the stimulus has most typically been operationalized on the basis of pilot participants' ratings of how much the stimulus implies the trait (e.g., Higgins & Brandl, 1995). Nonetheless, research on transference has taken a rather strict, feature-based approach in defining applicability, guided by theorizing on feature-based approaches from the social-cognitive domain (e.g., Higgins, 1996), as well as from the cognitive and developmental literatures (for recent reviews, see Komatsu, 1992; Medin et al., 1993). As noted, these approaches generally maintain that categorization reflects in part the degree of feature-based matches and mismatches between stored knowledge about a category and a to-be-classified stimulus.