Social Cognition, Vol. 16, No. 1, 1998, pp. 31–55

IN SEARCH OF SIMILARITY: STEREOTYPES AS NAIVE THEORIES IN SOCIAL CATEGORIZATION

BERND WITTENBRINK
University of Chicago

JAMES L. HILTON
University of Michigan

PAMELA L. GIST
Mount Saint Mary's College

Traditionally, models of categorization have been based on the premise that, in categorization, stimuli are grouped together because they appear similar to each other. In the social domain, where categorization processes are thought to play an important role in the phenomenon of stereotyping, such a similarity-based conception of category organization seems inadequate. Stereotypes, as categorical knowledge associated with social groups, generally reflect the perceiver's subjective construal of similarity relations in the social environment, instead of these similarity relations being based on an a priori structure of attribute covariations. Moreover, people's knowledge about social categories generally goes beyond assumptions regarding the presence or absence of category attributes that presumably define similarity relations. In many instances, social categorical knowledge includes important assumptions about how group attributes are related to one another; and the grouping of the social environment therefore reflects the perceiver's inferences and causal attributions based on this knowledge. Two experiments are reported to illustrate these arguments. The possibility of conceptualizing stereotypes as social knowledge organized by the perceiver's naive theories is explored as an alternative to a purely similarity-based categorization model.

> A young woman said to me: "I have had the most horrible experiences with furriers; they robbed me, they burned the fur I entrusted to them. Well, they were all Jews." But why did she choose to hate Jews rather than furriers? (Sartre, 1948, pp. 11–12)

Address correspondence to Bernd Wittenbrink, Center for Decision Research, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637; or, e-mail: bernd.wittenbrink@gsb.uchicago.edu.

> No one quite knows why related ideas in our minds tend to cohere and form categories.
> (Allport, 1954, p. 167)

Over the past two decades, social psychology has linked the phenomenon of stereotyping with the perceiver's tendency to group stimuli into meaningful entities and to base judgments and inferences on knowledge about the group of stimuli as a whole rather than on the individual stimulus (Allport, 1954; Tajfel, 1969). By grouping stimuli into categories, perceivers are able to accomplish two important goals (Bruner, 1957). Specifically, categorization allows perceivers to reduce the flow of incoming information (i.e., simplification) and simultaneously allows perceivers to go beyond the information given in any single instance (i.e., enrichment). Just as we could infer that a new species of bird would likely fly, given what we know about birds in general, so too could we infer that a particular woman would likely be passive and nurturing, assuming that we hold a traditional stereotype of women (Eagly, 1987).

Clearly, this categorization approach to stereotyping has greatly advanced our understanding of the phenomenon. In addition to directing the perceiver's attention (Bodenhausen, 1988; Cohen, 1981; Hilton, Klein, & von Hippel, 1991), stereotypes appear to affect recall (Hastie & Kumar, 1979; Rothbart, Evans, & Fulero, 1979; Stangor & Duan, 1991), and to influence the interpretation of perceived data (Banaji, Hardin, & Rothman, 1993; Biernat, Manis, & Nelson, 1991; Darley & Gross, 1983; Duncan, 1976; Kunda & Sherman-Williams, 1993). But what is it that actually makes a social category? How do we decide to partition our social environment into separate entities and why do we group the people around us the way we do? Why, in fact, does Sartre's anti-Semite decide to group other people according to their religious or cultural backgrounds, and thus blame the Jew, instead of the furrier? Given the fundamental nature of these questions, and given the current substantive body of research in the area, one would expect some ready answers. However, the issue turns out to be as complicated as it is fundamental. In the present article, we will discuss some of these complications, and we will suggest that reflection upon what holds social groupings together has important ramifications for the questions to be asked and answered by stereotype research.

SIMILARITY AS THE ORGANIZING PRINCIPLE UNDERLYING SOCIAL CATEGORIZATION

One possible and frequently offered answer to the question of what determines the partitions between groupings in the social environment

is that category coherence lies within the stimulus environment itself. According to this perspective, stimuli are grouped together simply because they appear similar to each other. Rosch (1975, 1978), for example, suggested that the categories people use to represent the world around them reflect existing similarity clusters in the environment. That is, certain stimulus attributes co-occur more frequently than others (e.g., tall and blond haired vs. tall and long nosed) and the categories that we use map this covariation of attributes—cutting the world at its naturally occurring "joints" (Rosch, 1978). Moreover, at what Rosch refers to as the "basic" level of abstraction (e.g., bird), categories maximize within-category similarity relative to between-category similarity.

Rosch's notion that discontinuities in the stimulus environment determine the partitions for people's categories has also been influential in theorizing about social categorization. Several authors have argued that individual attributes that are particularly salient and obvious (e.g., gender, race, age) serve as basic categories that are employed universally and more or less automatically (Brewer, 1988; Brewer & Lui, 1989; Fiske & Neuberg, 1990). Moreover, in the area of stereotyping, this realistic view that categories map existing covariation is reflected in the assumption that stereotypic conceptions of social groups presumably form around "a grain of truth" (Campbell, 1967). Eagly and her colleagues (Eagly, 1987; Eagly & Steffen, 1984), for example, have argued that stereotypic conceptions of males (e.g., agentic) and females (e.g., communal) result in part from the fact that traditional gender roles assign males and females to different occupational activities in society (e.g., breadwinner vs. homemaker).

Although intuitively appealing, there are several serious difficulties with the assumption that actual patterns of similarity provide partitions for categorizing the stimulus environment. One problem arises from the fact that the number of attributes that are potentially descriptive of any stimulus is simply endless (Quine, 1960; Zajonc, 1960). As Murphy and Medin (1985) have pointed out, even such dissimilar objects as plums and lawnmowers share a number of attributes (e.g., both have a smell, both can be dropped, etc.), and it is unclear how the perceiver would pick, without additional constraints, the "correct" covariation out of the infinite number of possible covariations among attributes in the environment.

A second problem arises from the fact that it is simply false that people's categories generally reflect existing patterns of covariation. Assertiveness, for example, is a common attribute in gender stereotypes, despite the fact that the effect size for actual differences between males and females on this dimension is relatively low (Eagly, 1987). Yet, other attributes that do show significant differences between males and females, such as body height, are conspicuously absent from the dominant

stereotypes of women and men (Hoffman & Hurst, 1990), nor do they seem to be relevant for people's assessment of how well various female exemplars fit their ideas of a prototypic woman (Armstrong, Gleitman, & Gleitman, 1983; see Molnar, 1992 for similar arguments regarding racial categories).

Hoffman and Hurst (1990) recently have raised similar criticisms, arguing that, instead of being a crude but truthful representation of reality, social stereotypes exist to justify behavior and attitudes toward the stereotype target (for corresponding arguments, see also Allport, 1954; Jost & Banaji 1994; Tajfel, 1981). Rather than reflecting a kernel of truth about female characteristics, for example, gender stereotypes are seen as serving to "rationalize, justify, or explain the sexual division of labor" (Hoffman & Hurst, 1990, p. 199). To support their position, Hoffman and Hurst presented participants with trait descriptions of individuals from a fictional society. Half of these targets were said to occupy the role of city workers, while the other half of the targets were described as child raisers. Consecutively, participants received descriptions of various members of these two groups, indicating their role and three personality traits for each target. Although these descriptions were designed such that any given trait would appear with equal probability for either of the two groups, participants showed systematic distortions as to what traits they associated with the two groups. While the city workers were believed to be more assertive, self-reliant, and so forth, the child raisers were thought to be more compassionate and sensitive. Hoffman and Hurst interpret these results as evidence that gender stereotypes develop in part as a consequence of people's attempts to explain the existing role distribution between the sexes. In other words, gender stereotypes are thought to be based on other relevant world knowledge, and the association of stereotypic attributes with males and females is determined by whether or not these attributes are consistent with this knowledge, rather than whether or not they are empirically correlated.

Unfortunately, Hoffman and Hurst's data are not as conclusive as one might wish. After all, it remains doubtful whether participants did indeed form new stereotypes for the fictional characters described in the stimulus materials, or whether they simply used their existing gender stereotypes in answering questions as to what traits would go along with child raising or bread winning, respectively. Despite these ambiguities, Hoffman and Hurst's point is well-taken, and the first experiment was designed to pursue this idea from a slightly different perspective. Specifically, this experiment addressed the question of whether perceptions of similarity are in fact dependent upon the perceiver's background knowledge.

STUDY 1

The specific focus of the first experiment is on Rosch's proposition that the actual covariation of stimulus attributes determines clusters of similar stimuli onto which people's categories are mapped, and our assertion that the partitioning of the stimulus environment is the result of an active construal of such covariation patterns. To examine these two perspectives, participants were asked to sort a set of stimuli, consisting of children's figure drawings, into groups that they considered to be similar. In two experimental conditions, participants were given differential background knowledge about these drawings. For both conditions, this background knowledge was presented in a fictional context (i.e., drawings made by aliens from outer space), in order to limit the relevance of participants' general world knowledge. If background knowledge is indeed relevant for the partitions used to categorize the stimulus environment, then this manipulation of available background knowledge should lead to different results in the sorting task. That is, different stimuli should be grouped together as similar. In contrast, the similarity-based model of categorization assumes that perceived similarity is determined sufficiently by the objective similarities contained in the stimulus material, and thus predicts identical groupings of the stimuli in both experimental conditions.

METHOD

PARTICIPANTS

One hundred and fifty-two Introductory Psychology students participated in the experiment in partial fulfillment of their course requirements.

MATERIALS

Children's figure drawings obtained from the Draw-a-Person test (DAP) were used as stimulus materials (Harris, 1963). The set that we used consisted of 20 black-and-white drawings collected from clinical literature (Harris, 1963; Koppitz, 1984), as well as from stimulus materials used in a study by Medin, Nakamura, and Wisniewski (reported in Medin, 1989). These drawings provided a relatively rich set of stimuli that included a large number of various details (e.g., hat; motion of the figure, etc.), thus containing numerous features that participants could have potentially used for their similarity judgments.