Medin & Schaeffer, 1978), memory (e.g., Murdock, 1993), learning (e.g., Spence, 1936), and induction across categories (Osherson, Smith, Shafir, Gualtierotti, & Biolsi, 1995; Sloman, 1993). A fundamental assumption made by all these models is that concepts can be meaningfully reduced to sets of features; i.e., to constituent aspects, parts, and attributes. These features are treated as independent in the sense that they make separable—usually additive—contributions to the model's output. However, these models need not consider features independent in another sense: The importance of a feature may be a function of its relations to other features. Features may differ in their centrality with respect to a concept and the only viable accounts of centrality appeal to features' roles in networks of knowledge.

A compelling illustration of this hypothesis was provided by Keil (1989) in the domain of categorization (see also Rips, 1989). Keil reported a series of experiments in which children and adults were asked to categorize stimuli having the perceptual features of one category but the internal features of another. Although all participants tended to categorize artifacts according to their perceptual features, older children and adults were more likely to categorize natural kinds according to their internal features. Keil argued that the internal features of natural kinds became weightier in category judgments to the extent that the participants had developed explanatory biological theories. Features were differentially weighted in categorization decisions, and the weighting was somehow a product of people's knowledge about the interrelations between features.

To explain these data and others, theorists like Keil (1989, 1994), Carey (1985), and Murphy (1993; Murphy & Medin, 1985) argue that features are weighted in categorization decisions according to the centrality of the feature in an "intuitive theory" of the category. Unfortunately, a precise model of centrality with respect to an intuitive theory has yet to be articulated. The language of intuitive theories has not been sufficiently well-formulated to support a precise notion of centrality. We will try to offer such a notion. To do so, we will ignore, or at least abstract from, much of the structured knowledge that the intuitive theory view seems to presuppose. We will rely on the simple mechanics of constraint satisfaction. Our model of centrality can thus be understood as a bridge between the intuitive theory view of conceptual structure and the more impoverished, but more clearly articulated, feature-based approach.

Mutability and Dependency Structure

The centrality of a feature represents the degree to which the feature is integral to the mental representation of an object, the degree to which it lends conceptual coherence. We will therefore measure the degree of coherence associated with a feature by asking people how easily they can transform their mental representation of an object by eliminating the feature, or by replacing the feature with a different value, without changing other aspects of the object's representation. We call such judgments measures of "mutability" (cf. Kahneman & Miller, 1986), because they reflect how much a feature of a concept of an object can be mentally transformed.

This paper has two objectives. First, we will argue that the mutability of conceptual features can be represented as a single, multiple-valued dimension. We will show that the features of a concept can be reliably ordered with respect to the degree to which people are