# Causal Status, Coherence, and Essentialized Categories

**ShinWoo Kim (shinwoo.kim@nyu.edu)**
**Bob Rehder (bob.rehder@nyu.edu)**
Department of Psychology, New York University
6 Washington Place, New York, NY 10003 USA

People possess theoretical knowledge about categories beyond what they observe. For example, they not only know that birds fly, have wings, and build nests in trees, but also that they build nests in trees because they can fly and can fly because they have wings. When features are arranged in a causal chain (e.g., X→Y→Z), two effects on classification have been established. First, a *causal status effect* obtains in which the root of the causal chain (X) has greater influence on classification than Y (and, in some studies, Y has greater influence than Z) (Ahn et al., 2000; Rehder & Kim, 2006). Second, a *coherence effect* obtains in which objects are better category members when their combination of features is consistent with the causal laws (e.g., X and Y both present or both absent) rather than inconsistent (e.g., X present and Y absent or vice versa) (e.g., Rehder, 2003).

The robustness of these effects has been questioned, however. First, Rehder (2003) suggested that the causal status effect arises only when the causal chain of features is thought to be caused by underlying feature(s) considered to be *essential* to category membership. On this account, X is more influential because it is more likely to be generated than Y or Z. Rehder & Kim (2006) also suggested that the causal status effect may depend on the fact that in most studies features are listed in order (XYZ) on the computer screen. Second, Marsh & Ahn (2006) argued that because Rehder (2003) instructed subjects on categories like *Lake Victoria Shrimp*, most of which were described as having a "high body weight" whereas some had a "normal body weight," the repetition of "normal" may have magnified coherence effects for atypical exemplars (e.g., 0000).

To test these claims, subjects were taught categories exhibiting one of the three structures in Fig. 1. In the Essential-Chain condition, the causal chain of features is described as being caused by an "essential" feature E. For example, subjects who learned Lake Victoria Shrimp learned that all category members have "high amounts of the ACh neurotransmitter" and that no other kind of shrimp did. High ACh thus served as an essential feature (E). In contrast, the causal chain was not caused by E in the Chain condition and there were no causal links in the Control condition. Causal links were described as having a strength of 80%. Subjects rated the degree of category membership of 8 exemplars that can be formed by binary combinations of X, Y, and Z. We predicted a causal status effect would obtain in the Essential-Chain condition only and that coherence effects would obtain in both the Essential-Chain and Chain conditions. Note that the presentation order of features was randomized over subjects and "normal" features were not used (e.g., subjects were told that most Lake Victoria Shrimp had a high body weight and that some had a *low* body weight).
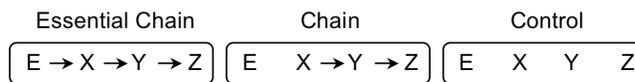


Figure 1

The ratings were subjected to regression with predictors that coded the presence or absence of each feature (which provided a measure of feature importance) and predictors that coded the 2-way interactions between features (which provided a measure of coherence). The results are presented in Fig. 2. First, a stronger causal status effect (feature weight X > Y > Z) obtained in the Essential-Chain condition than in the Chain condition. This result confirms Rehder's (2003) claim that the causal status effect depends on the features being caused by underlying essential properties. Second, feature weights both in the Chain and Control condition did not differ from one another. This result confirms Rehder & Kim's (2006) claim that past demonstrations of a causal status effect arose because features were presented in causal order. Third, a coherence effect obtained in both causal conditions, indicated by the large positive interactions between features directly causally linked (XY and YZ) and weaker interactions between features indirectly linked (XZ). Thus, contra the suggestion of Marsh & Ahn (2006), a strong and robust coherence effect does not depend on the "normal" feature wording. These results are consistent with a generative theory of classification in which classifiers estimate how likely an object's features were to have been generated by a category's causal model.
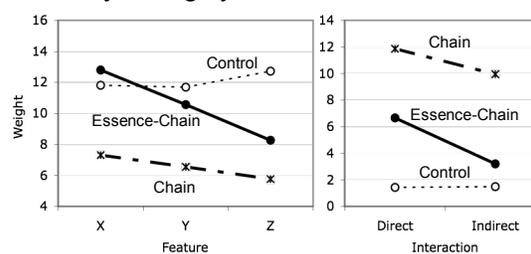


Figure 2

## References

Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41*, 361-416.

Marsh, J. K., & Ahn, W. (2006). The role of causal status versus interfeature links in feature weighting. *The 26th Annual Conference of the Cognitive Science Conference*. (pp. 561-566).

Rehder, B. (2003). A causal-model theory of conceptual representation and categorization. *JEP:LMC, 29*, 1141-1159

Rehder, B. & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *JEP:LMC, 32,* 659-683.