Qualitative and Quantitative Individual Differences in Semantic Categorization

Steven Verheyen (steven.verheyen@ppw.kuleuven.be)
Laboratory of Experimental Psychology, University of Leuven
Tiensestraat 102 - box 3711, B-3000 Leuven, Belgium

Gert Storms (gert.storms@ppw.kuleuven.be)
Laboratory of Experimental Psychology, University of Leuven
Tiensestraat 102 - box 3711, B-3000 Leuven, Belgium

Keywords: mixture models; IRT models; threshold models.

Introduction

When asked to indicate which items from a set of candidates belong to a particular category, inter-individual differences appear: Individuals disagree on the items that should be considered category members (e.g., Black, 1937; Hampton, Dubois, & Yeh, 2006; McCloskey & Glucksberg, 1978). Individuals might disagree about whether *hiking* and/or *darts* are *sports*, for instance. We will argue that inter-individual differences in semantic categorization come in two kinds. (i) Qualitative differences reflect a different organization of the candidate items with respect to the target category. (ii) Quantitative differences reflect a different propensity to endorse items as category members.

Qualitative differences represent different views on what are considered representative category members. Individuals who consider *hiking* a better example of *sports* than *darts* presumably find that the former meets the requirements of category membership better than the latter does. They, for instance, recognize that *hiking* is physically more demanding than *darts* is. Individuals who consider *darts* to be the better example must then employ different requirements for category membership. When judging category membership they place more emphasis on elements such as rules or competition, for instance. These requirements are better met by *darts* than by *hiking*.

Among individuals who agree on the organization of items with respect to the target category, categorization differences of a quantitative nature can arise. These are differences that pertain to the propensity to endorse items as category members. The item organization reflects the varying extents to which the items fulfill the requirements for category membership (e.g., *hiking* is physically more demanding than *darts* is). Certain individuals might want to see more evidence of these requirements than others. They might only deem *hiking* physical enough to be considered a *sport*, while others find both *darts* and *hiking* demanding enough.

Model

Our goal is to elucidate these two kinds of inter-individual categorization differences by means of the mixture item response theory model (Mislevy & Verhelst, 1990; Rost, 1990) that is expanded in Equation (1). Models like these are traditionally employed to assess individuals' aptitudes and dispositions in response to a number of test items. However, they have also been shown to be flexible tools to analyze semantic categorization data (e.g., Verguts, De Boeck, & Storms, 1998; Verheyen, Hampton, & Storms, 2010).

Here the model will be used to partition a large sample of categorizers in a number of groups that are maximally different in terms of their organization of the items with respect to the target category. These item organizations take the shape of scales along which all candidate items are positioned according to their likelihood of being endorsed. Different organizations capture qualitative categorization differences between participants from different groups: They reveal how a particular item might be a likely category member in one group, but an unlikely category member in another group. Participants that end up together in a group are understood to adopt the same item organization. These categorizers do not differ qualitatively, but can display varying degrees of propensity to endorse items as category members. These are inter-individual categorization differences of a quantitative nature. In the formal framework they take the shape of criteria that are imposed on the scales that organize the candidate items: They reveal how some individuals in a group might use very liberal criteria, while others employ very stringent criteria.

Binary categorization decisions Y_{ci} constitute the input for the model. Here the categorization data are comprised of member/non-member decisions Y by 250 categorizers c towards 24 items i in each of 8 natural language categories (*fish*, *fruits*, *furniture*, *insects*, *sciences*, *sports*, *tools*, *vegetables*).

Every one of these categorization decisions is considered the outcome of a Bernoulli trial with the probability of a *member* response:

$$\Pr(Y_{ci} = 1) = \frac{e^{\alpha_g(\beta_{g,i} - \theta_{gc})}}{1 + e^{\alpha_g(\beta_{g,i} - \theta_{gc})}} \tag{1}$$

In Equation (1) the betas capture the organization of the items with respect to the target category. g groups of categorizers are extracted, with separate item organizations that are maximally different. For each group the organization takes the shape of a scale along which all candidate items are positioned. $\beta_{g,i}$ indicates the position of item i along the scale for group g. Higher values for $\beta_{g,i}$ indicate likelier category members.

The thetas in Equation (1) capture the degree of liberalness/conservatism categorizers display. A separate indication of the propensity to endorse items as category members is extracted for each categorizer c. It takes the shape of a criterion that is positioned along the same scale that organizes the items for the group the participant belongs to. θ_{gc} indicates the position of the criterion for categorizer c along the scale for group g. Higher values for θ_{gc} indicate more conservative categorizers.

Unlike the betas and thetas, the alphas in Equation (1) can only take on positive values. A separate α_g for each group determines the shape of the response function that relates the unbounded difference $\beta_{g,i} - \theta_{gc}$ to the probability of a *member* response (bounded between 0 and 1). Indeed, the relative position of item and criterion along a scale determines the probability of a *member* response. If $\beta_{g,i}$ equals θ_{gc} the numerator of Equation (1) takes on the value of 1, while the denominator takes on the value of 2. The resulting probability is .50, indicating that the categorization decision can go either way. The odds change when item and criterion have a different position along the scale. If the item surpasses the criterion, the odds are that the categorizer will endorse it. The greater the distance between item and criterion, the greater the odds of a *member* decision. If the item does not surpass the criterion, the odds are that c will not endorse i. Under this circumstance, the odds of a *non-member* decision increase with the distance between item and criterion.

The one-group variant of the model in Equation (1) has been applied to semantic categorization by Verheyen et al. (2010). That particular model only allows for quantitative inter-individual categorization differences. Participants can differ in terms of the categorization criterion they employ, but not in terms of the scale along which the criteria are placed. They all adopt the same category organization. The model in Equation (1) is more general. It allows for qualitative differences in addition to quantitative ones. It relaxes the assumption that all participants adhere to the same category organization. Instead, it assumes that the participants divide in groups with a different item organization each. (One set of beta estimates is extracted for each group.) Within each group, individuals are still thought to differ in terms of the employed categorization criterion. (A theta estimate is extracted for every categorizer.) The model in Equation (1), then, is a mixture of differently parameterized quantitative differences-only models of the kind employed by Verheyen et al. (2010).

Findings

The analysis of the categorization data with the mixture model in Equation (1) yields evidence for both qualitative and quantitative inter-individual differences. For the categories of *fish*, *insects*, *sciences*, *sports*, and *tools* the sample of categorizers divides in distinct groups, who regard different items likely category members (i.e., qualitative differences). Within each of these groups categorizers differ in their propensity to provide membership responses (i.e., quantitative differences). The existence of multiple item organizations for a single category suggests that it might be improper to assume a default category representation that is the same for all language users. Rather, it would appear that there exist a number of these default representations, which emphasize different sets of category features. Indeed, a clear pattern emerged when we (i) determined to what extent features that participants consider important for category membership are true of the different candidate items, (ii) obtained a small number of principal components that convey the information that is contained in these feature applicability judgments, and (iii) regressed the item organizations of different groups upon these principal components. For each of the categories with multiple item organizations, there was at least one component that had a similar effect on every item organization. Common components indicate agreement among groups on what it means to be a category member. This is required for members of different groups to succesfully communicate with one another using the studied natural language terms. The item organizations could also be distinguished on the basis of other components that were of importance to single subgroups only. These distinct components indicate disagreement among groups on what it means to be a category member but do not appear to hamper communication between members of different groups.

Acknowledgments

Steven Verheyen is a postdoctoral fellow at the Research Foundation-Flanders.

References

Black, M. (1937). Vagueness: An exercise in logical analysis. *Philosophy of Science*, *4*, 427-455.

Hampton, J. A., Dubois, D., & Yeh, W. (2006). Effects of classification context on categorization in natural categories. *Memory & Cognition*, *34*, 1431-1443.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, *6*, 462-472.

Mislevy, R. J., & Verhelst, N. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika*, *55*, 195-215.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, *14*, 271-282.

Verguts, T., De Boeck, P., & Storms, G. (1998). Analyzing experimental data using the Rasch model. *Behavior Research Methods*, *30*, 501-505.

Verheyen, S., Hampton, J. A., & Storms, G. (2010). A probabilistic threshold model: Analyzing semantic categorization data with the Rasch model. *Acta Psychologica*, *135*, 216-225.