# A REPRESENTATION SHIFT MODEL OF CLASSIFICATION AND ABSOLUTE IDENTIFICATION

Vlad Zotov and D. J. K. Mewhort
*Department of Psychology, Queens University*

**Abstract**

*We studied shifts of representation using production trials to measure the representation following classification and absolute identification. Identification induced a shift of representation away from the preceding exemplar. Shifts in classification depended on the category of the preceding trial. When the exemplars were from different categories, the category representation was pushed away from the preceding category. When the exemplars were from the same category, the representation of the current category was pulled from its centre in the direction of the preceding exemplar.*

*We propose a representation shift model to account for trial-to-trial errors in both tasks. The model treats the subject's representation of category and item structure as points in a space. The points shift dynamically from trial to trial as a function of the properties of the preceding exemplar.*

*The RSM captures overall accuracy in both classification and absolute identification, replicates benchmark characteristics of performance in both absolute identification, and classification.*

To investigate the source of systematic trial-to-trial errors in classification, Jones, Zotov, & Mewhort (2003) used a production technique to measure trial-to-trial changes in category representation. There were two systematic trial-to-trial shifts: When a second example from the same category was produced, it was pulled towards the exemplar on the preceding trial. When the second example was from a different category, produced exemplars were pushed away from the preceding category.

To tie production and performance data we combined the production measurement with classification and identification techniques in a series of three-trial tests—production-production-classification and production-production-identification trials. The three tests allows us to consider sequence biases not only from one classification or identification trial on another but also from a classification or identification trial to a production trial. In the identification-production task (Task 1, test 3), we found that the subject's representation—measured using his own production responses—was shifted away from the stimulus of the preceding trial. The magnitude of push increase as the distance between current and preceding stimulus increased. The results are consistent with the pattern of trial-to-trial errors documented in the absolute identification literature (Garner, 1953; Holland & Lockhead, 1968; Ward & Lockhead, 1970) and by Jones, et al.'s (2003) production task.

In the classification-production task (Task 2, test 3), we found that the subject's representation was shifted away from the stimulus of the preceding trial, but the direction of the shift depended on the category of the preceding trial. Figure 1 illustrates the shift after stimulus 5, 6, or 7 (top, middle, and bottom panels of Category B) was presented: the representation of Categories A and C was pushed away from the preceding trial. The magnitude of push increased as the distance between current and preceding stimulus increased. The representation of Category B itself was shifted from its centre towards the direction of the preceding exemplar.

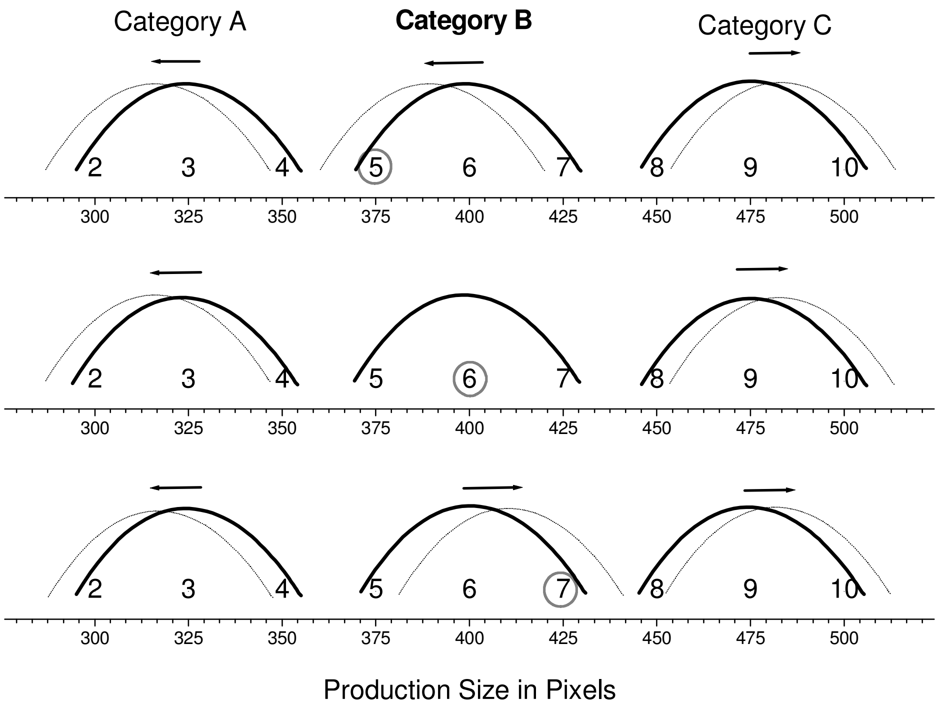


***Figure 1. A schematic diagram illustrating shifts in representation found in Task 2 . The black arcs show static representations of categories, and the dotted arcs show the shifted representation of categories. Circled are stimuli presented on the preceding trials.***

## Representation Shift Model (RSM)

***Mechanism***. In the RSM, categories reside in a multidimensional psychological space formed by encountering exemplars. The RSM stores the borderline exemplars that define category space and serve as anchors, defining category's boundaries. Hence, for uni-dimensional stimuli, each category's space is represented by the location of two peripheral exemplars $x_{min}$ and $x_{max}$. To account for perceptual noise associated with each of the exemplars, the category space is extended beyond $x_{min}$ and $x_{max}$ values by constant value $\gamma$, becoming ($x_{min} - \gamma$) and ($x_{max+}\ \gamma$). In identification, the RSM stores items as categories; each item has its own boundaries defined by the location of the item $x_i$. Similar to classification, the space of $x_i$ item is extended by a constant value $\gamma$, extending it from ($x_{i-}\ \gamma$) to ($x_{i\,+}\ \gamma$).

The $\gamma$-parameter used to identify the width of a category in the RSM is similar to the decision bound in the decision-boundary theory (DBT). The DBT relies on anchors that define category boundary. In DBT, subjects learn category structure in the form of a multidimensional cognitive space (Ashby & Maddox, 1993).

The representation in the RSM is dynamic. Each stimulus classification causes two shifts in representation of the tested categories—a between-category shift that pushes the representation of neighbouring categories *away* from the category of the preceding trial (***α*-shift**) and a within-category shift that pulls the representation of the category *toward* the item presented on the preceding trial (***β*-shift**). The magnitude of the $\alpha$-shift increases with the number of intervening categories by a constant ***Incr*** ; *Incr* is associated with an increased push as the distance in category steps between current and preceding category increases. When the location of the presented stimulus coincides with the centre of the category, the within-category shift sets category to its default location, that is, $\beta$=0. In identification, there are no within-category variations in location of items because each item is its own category; thus, there is only a between-item $\alpha$-shift in identification.

After Stimulus *x* has been classified as an exemplar of Category *X*, it caused changes in the positions of categories in psychological space, corresponding to two shifts in representation. The location of Stimulus *x* is stored in temporary memory serving as a reference point for shifts in representation. For the between-category case, the position of Category *A* is pushed away from Category *X*:

$$\begin{aligned} A_{\min} &= A_{default\,\min} + \alpha + (N_{steps} \times Incr); \\ A_{\max} &= A_{default\,\max} + \alpha + (N_{steps} \times Incr), \end{aligned} \tag{1}$$

where $N_{steps}$ is a number of categories separating categories *X* and *A*, $A_{default}$ is a default location of the boundaries of Category *A*, and *Incr* is a constant associated with increased push in representation. The same mechanism of the between-item shift is used in identification, where the default boundaries of each item shift away from the stimulus.

For the within-category case, (Category X), the position of Category X is pulled toward location of Stimulus *x*:

$$\begin{aligned} X_{\min} &= X_{default\,\min} + \beta; \\ X_{\max} &= X_{default\,\max} + \beta \end{aligned} \tag{2}$$

When the next stimulus is presented, it is classified according to the shifted representation induced by N-1 trial. After the stimulus is classified, categories are reset to their default locations, the new location of the latest stimulus is stored in memory, and representation of categories is shifted according to Equations 1 and 2.

In the RSM, classification depends on similarity of the tested categories. Unlike exemplar accounts that rely on distances between individual exemplars, the RSM relies on distances between the target stimulus and the boundary of the category (item in identification). The distance in the psychological space, *d*, is converted to a similarity measure $\eta$, by using a decreasing exponential function (Shepard's generalization law, 1987).

The probability of Stimulus *x* being classified to any of the tested categories depends on how similar Stimulus *x* is to its own category in comparison to its similarity to the alternative category. The conditional probability with which Stimulus *x* is classified into Category *A* is found by dividing the similarity of Stimulus *x* to Category *A* by the summed similarity of Stimulus *x* to all other categories:

$$P(A \mid x_i) = \frac{\sum \eta_{x_i, A}}{\sum_{K} \sum_{k \in K} \eta_{x_i K}}, \tag{3}$$

where K is a number of categories. Likewise, the conditional probability with which Stimulus *x* is identified as Item *x* is found by dividing the similarity of Stimulus *x* to Item *x* by the summed similarity of Stimulus *x* to all other items:

$$P(X \mid x_i) = \frac{\sum \eta_{x_i,X}}{\sum_{K} \sum_{k \in K} \eta_{x_i K}}, \tag{4}$$

where K is a number of items. Equations 3 and 4 provide the mechanism that is used to predict accuracy of responding to each category.

To fit the data, we used a downhill simplex algorithm (Press, Flannery, Teukolsky, & Vitterling, 1986). Each subject's data were fitted individually.

***Assimilation in identification.*** To analyze the effect of the preceding item on the subject's evaluation of the current trial in Task 1 (test 2), we tallied the number of responses to each item, conditional on the item of the preceding trial. Hence, the basic data for each item were contained in a 5x5 frequency matrix. We calculated a bias index associated each preceding item, and we weighted the frequency by its ordinal position in the matrix, summed the weighted frequency values, and divided the sum by the total number of responses. The bias score reflects the direction of the error responses. If the errors were distributed symmetrically around the correct answer, the adjusted bias score would be zero.

Figure 2 presents bias scores as a function of the item of the preceding trial observed in the Task 1. As is shown in the figure, the bias increased monotonically across the items as a function of the distance between the current and the preceding items. The increasing function shows that the subjects responded to the items as if they were closer to the preceding item than they actually were.

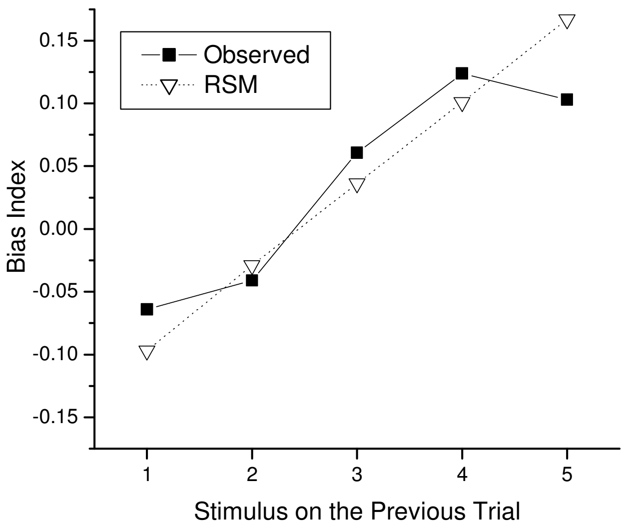


***Figure 2. The mean response bias score, as a function of the category on the preceding trial observed and predicted by the RSM.***

***Assimilation and contrast in classification.*** Figure 3 (left panel) presents bias scores as a function of the category of the preceding trial from the Task 2 (test 2). As shown, bias increased monotonically across the items as a function of the distance between the current and the preceding categories. The function implies that the subjects responded to the items as if they were closer to the preceding category than they actually were.

For within-category comparisons we report accuracy as a function of distance between preceding and current stimulus. With the same-category pairs in production trials of Task 2, produced exemplars shifted away from the centre of the category in the direction of the preceding stimulus. In classification trials of Task 2, the accuracy for same-category pairs reflected the shift: as category pulled away from the presented item, its accuracy decreased. Figure 3 (right panel) shows the proportion correct as a function of distance between successive within-category comparisons.

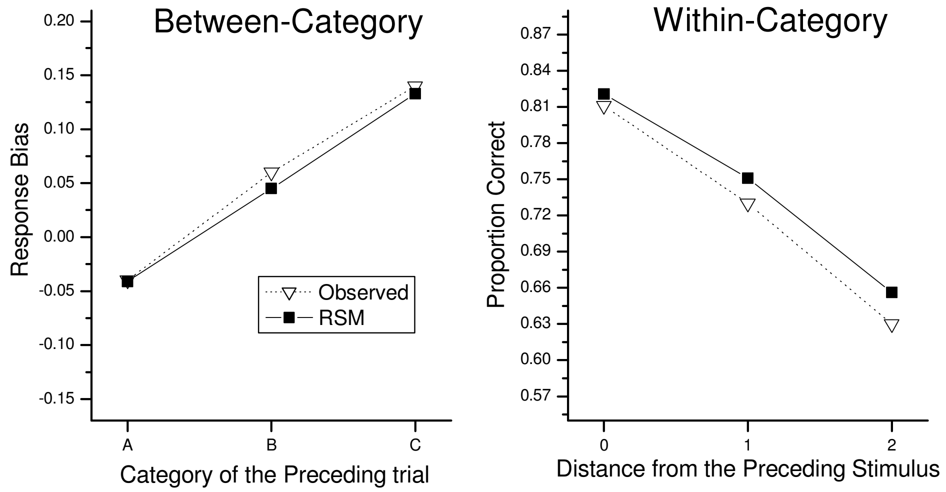


***Figure 3. The left panel shows the mean response error score, as a function of the category on the preceding trial observed and predicted by the RSM. The right panel shows proportion correct, as a function of the stimulus on the preceding trial observed and predicted by the RSM.***

***Bow effect.*** Figure 4 shows the accuracy of individual stimuli observed in Task 1.

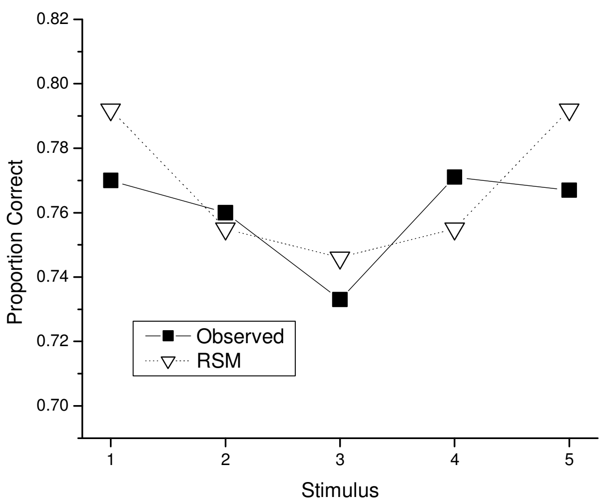


***Figure 4. Accuracy observed in identification-production task and predicted by the RSM for each stimulus.***

***Distance-to-boundary effect in classification.*** Figure 5 shows a typical boundary effect, where accuracy for borderline exemplars is lower than those located more centrally.

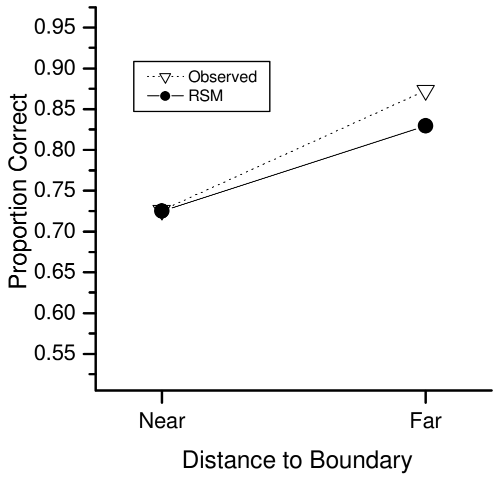


***Figure 5. The mean response error score, as a function of the category on the preceding trial observed and predicted by the RSM.***

## Conclusions

As it is clear from Figures 2 to 5, the RSM was able to account for both sequence-dependent and sequence-independent effects in both tasks: (a) It captured the well-known distance-to-boundary effect in classification and the bow effect in identification. (b) It also captured assimilation and contrast in classification and assimilation in identification. The RSM applies the same mechanism to both tasks and positions itself as a unified account of classification and identification.


### Acknowledgements

The research was supported by grants from Natural Science and Engineering Research Council of Canada (AP-318) and by AEG grant from Sun Microsystems of Canada to D. J. K. Mewhort.


## References


Ashby, F. G. & Maddox, W. T. (1993). Relations between prototype. Exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37,* 372-400.

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46,* 373-380.

Holland, M. K., & Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception & Psychophysics, 3,* 409-414.

Jones, M. N., Zotov, V., & Mewhort, D. J. K. (2003). *Follow the floating representation: Category sequence effects in an exemplar-production task.* Paper presented at the 13-th Annual Meeting of the CSBBCS, McMaster University.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vitterling, W. T. (1986). *Numerical Recipes: The art of scientific computing*. New York: Cambridge University Press.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*, 1317-1323.

Ward, L. M., & Lockhead, G. R. (1970). Sequential effects and memory in category judgments. *Journal of Experimental Psychology, 84,* 27-34.