Representational Shifts in a Multiple-Cue Judgment Task with Continuous Cues

Linnea Karlsson (linnea.karlsson@psy.umu.se)
Department of Psychology, Umeå University
SE-901 87 Umeå, Sweden

Peter Juslin (peter.juslin@psyk.uu.se)
Department of Psychology, Uppsala University
SE-751 42 Uppsala, Sweden

Henrik Olsson (henrik.olsson@psyk.uu.se)
Department of Psychology, Uppsala University
SE-751 42 Uppsala, Sweden

Abstract

Research on multiple cue judgment with continuous cues and a continuous criterion has been dominated by statistical modeling of the cue utilization with linear multiple regression. In this study we apply two cognitive process models to investigate the relative contributions of explicit abstraction of the cue-criterion relations and memory for concrete exemplars in a multiple-cue judgment task. The task was an extension of a previous task with binary cues (P. Juslin, H. Olsson., A-C. Olsson, 2003) and involved multiple continuous cues that either combined by addition or multiplication. As predicted by the process model ∑ (P. Juslin, L. Karlsson, & H. Olsson, manuscript) explicit abstraction of cue-criterion relations were induced in the additive task, while exemplar memory was induced in the multiplicative task.

Introduction

Multiple-cue judgment research has traditionally been concerned with statistical modeling of judgment data. Rather exquisite regression models have been developed that describe multiple-cue judgment as *a*) well fitted by a linear additive model; *b*) only taking a few cues into account; *c*) hard to report on subjectively; *d*) characterized by cue weightings that differ greatly between individuals; and *e*) plagued by considerable inconsistency in the weighting of the cues (see Brehmer, 1994; Cooksey, 1996; Hammond & Stewart, 2001).

In the light of the cognitive revolution it might seem puzzling that this field of research has not benefited from the growth of cognitive modeling as a means to track the underlying cognitive representation and process of judgment, a growth seen in related fields like categorization learning (but see for example Bott & Heit, 2004; Busemeyer, Byun, DeLosh, & McDaniel, 1997; or DeLosh, Busemeyer & McDaniel, 1997, single-cue learning). Categorization – which is in many ways similar to multiple-cue judgment (see Juslin, Olsson, & Olsson, 2003) – has invited extensive investigation of the cognitive representations and processes that underlie behavior. A plethora of models, ranging from an emphasis on how abstract rules or prototypes guide category decisions to a domination of memory for category exemplars are thus available in cognitive science today. In this study we apply the methods of cognitive modeling to a typical multiple-cue judgment task. By connecting research on cognitive science to judgment and decision making research, we can gain an understanding of what cognitive representations and processes guide the judgments, and how this is manifested in the results of the traditional statistical modeling (e.g., Cooksey, 1996).

Arguably, it is not mere coincidence that linear, additive models fit multiple cue judgment data well and that categorization is often well captured by exemplar models that entail a linear additive combination of retrieved exemplars (Juslin, Karlsson, & Olsson, manuscript). Imagine how you sequentially consider and weigh the pros and cons of different aspects of a car before you purchase it (its looks, reliability, etc). You may weigh them differently but positive qualities *add* to and negative qualities *subtract* from your overall opinion. Likewise, you may sequentially consider exemplars of similar cars that you are aware of: similar cars (e.g., same model) that have worked properly *add* to the appeal of the car and cars that that have been frustrating *subtract* from it.

We have proposed a general process model, ∑, that captures the essentials of multiple-cue judgment, both when it is driven by consideration of cue-criterion relations and exemplar retrieval (Juslin et al., manuscript). The assumptions in ∑ are that our controlled and explicit thought processes have an architectural constraint enhancing sequential, real-time consideration of multiple pieces of evidence (cues or exemplars). The process involves successive adjustment of an estimate, a process structurally compatible with linear, additive cue integration (Einhorn, Kleinmuntz, & Kleinmuntz, 1979) and exemplar models (Nosofsky & Johansen, 2000).

The key assumption is that, in effect, all integration of information involves addition (or subtraction). This hypothesis suggests that explicit and controlled thought processes are apt at performing cue-integration only in tasks where the cue-criterion relations in the task indeed combine by addition. By contrast, a task that involves non-linear or multiplicative cue combination requires capitalization on exemplar memory (Medin & Schaffer, 1978; Nosofsky & Johansen, 2000). Exemplar memory involves no strong computational commitments to particular task structures. With a division of labor between distinct representations we are better equipped to adapt to different task structures. We propose that the judgment process adapts to specific task environments and predict that in a multiple-cue judgment task with continuous cues we can induce a shift between qualitatively distinct processes by manipulating the structural properties of the environment: additive cue combination should promote cue abstraction and multiplicative cue combination should promote exemplar memory.

Judgment Task and Cognitive Models

The judgment task involves judgment of a continuous criterion based on continuous cues. The task concerns judgments of the effectiveness of different species of herbs as medical treatments to a lethal virus. The effectiveness is measured as the *maximal amount of a chemical substance* (mg) that can be extracted from the species. The species have four continuous dimensions (C_1, C_2, C_3, C_4), and each cue dimension can take a value between 0 and 10. The judgment of effectiveness requires inference from these dimensions, which are presented as verbal statements (e.g., weeks of bloom per year, geographic place of growth).

The tasks involve two manipulations. First, there is one condition in which all cues are related to the criteria positively and linearly, and one condition in which two cues are positively and two cues are negatively linearly related to the criterion. This manipulation makes the task a matter of function learning. Second, there is a manipulation of whether the effects of the four cues on the criterion combine by *addition* or *multiplication*.

In the additive condition the criterion is a linear, additive function of the continuous cues:

$$c = 500 + 4 \cdot C_1 + 3 \cdot C_2 + 2 \cdot C_3 + 1 \cdot C_4 + \varepsilon \cdot \qquad (1)$$

C_1 is the most important cue with *coefficient* 4 (i.e., a relative weight .4), C_2 is the second to most important with coefficient 3, and so forth. The cues are uncorrelated. ε is a normally and independently distributed random error with a standard deviation that produces a multiple correlation R between cues and criterion of .9 (i.e., defining the ecological validity of the cues).

In the multiplicative condition the criterion c is a multiplicative function of the four cues:

$$c = 509.05 + 0.54545 \cdot e^{(C_1 \cdot 4 + C_2 \cdot 3 + C_3 \cdot 2 + C_4 \cdot 1)/18} + \varepsilon, \quad (2)$$

with the same coefficients as in the additive task (Eq. 1). The effectiveness varies between 500 and 600 mg of chemical substance in the additive task and 509 to 650 mg in the multiplicative task. However, the training ranges are hold equal for the two conditions. The range of cue values observed in the two tasks is therefore the same. Moreover, the criterion in the multiplicative condition is an exponential function of the criterion presented in the additive condition.

We use two structural models to derive predictions, a cue-abstraction (CAM) and an exemplar model (EBM). $\sum$ implies that in the additive task CAM should be the correct structural description of the process, whereas in the multiplicative condition EBM should be the appropriate description[1]. The CAM assumes that participants abstract explicit cue-criterion relations in training that are mentally integrated at the time of judgment. When presented with a probe the participants retrieve rules connecting cues to criterion (e.g., "More weeks in bloom gives more effectiveness"). The rules specify the sign and importance of each cue with a cue weight. For example, after training the rule for C_1 may specify that high C_1 goes with an increase in the criterion.

The CAM implies that participants compute an estimate of the criterion c based on sequential consideration of cues. For each cue, the estimate of c is adjusted according to the cue weight ω_{iA} (i=1…4). The final estimate $\hat{c}_{CA}$ is a linear additive function of the cues C_i,

$$\hat{c}_{CA} = k + \sum_{i=1}^{4} \omega_{iA} \cdot C_i, \qquad (3)$$

where $k = 500 + .5 \cdot (100 - 10 \cdot \sum \omega_{iA})$. If ω_{1A}=4, ω_{2A}=3, ω_{3A}=2, and ω_{4A}=1, Eq's 1 and 3 are identical and the model produce perfect judgments. The intercept k constrains the function relating judgments to criteria to be regressive around the midpoint (550) of the interval [500, 600] (Juslin et al., manuscript).

Although ruled out by the predictions of $\sum$, we also consider the possibility that participants have correctly abstracted the multiplicative cue-criterion relations by fitting a multiplicative cue-abstraction model to the data:

$$c = 509.05 + 0.54545 \cdot e^{(\sum_{i=1}^{4} \omega_{iM} \cdot C_i)/18} \qquad (4)$$

where ω_{1M} are the best fitting subjective cue weights in the multiplicative cue abstraction model.

EBM is commonly applied to classification, but here we apply it to a continuous criterion. EBM implies that participants make judgments by retrieving similar exemplars (herb species) from memory. When the exemplar model is applied to judgments of a continuous criterion variable, the estimate $\hat{c}_E$ of the criterion c is a weighted average of the criteria c_j stored for the J exemplars, where the similarities $S(p,x_j)$ are the weights:

$$\hat{c}_E = \frac{\sum_{j=1}^{J} S(p, x_j) \cdot c_j}{\sum_{j=1}^{J} S(p, x_j)}. \qquad (5)$$

[1] $\sum$ is a model of the real-time process of judgment that becomes structurally identical with a CAM when the representations fed to the process are abstracted cues and structurally identical to an EBM when the process is fed by exemplars. The structural description refers to the relationships between stimulus features and the response (Juslin et al., manuscript).

p is the probe to be judged, x_j is exemplar j (j= 1…J), and $S(p,x_j)$ is the similiarity between probe p and exemplar x_j. Eq. 5 is the *generalized context model* (GCM: Nosofsky, 1984; 1986), which generalizes the original version of the context model (Medin & Schaffer, 1978). The similarity $S(p,x_j)$ between exemplars is found by transforming the distance between them. The distance between a probe p and an exemplar j is,

$$d_{pj} = h\left[\sum_{M=1}^{M} w_m \left|x_{pm} - x_{jm}\right|\right], \qquad (6)$$

where x_{pm} and x_{jm}, respectively, are the values of the probe and an exemplar on cue dimension m, the parameters w_m are the attention weights associated with cue dimension m, and h is a sensitivity parameter that reflects overall discriminability in the psychological space (the sensitivity parameter is usually denoted c, but we changed that to avoid confusion with the criterion c). Attentional weights vary between 0 and 1 and are constrained to sum to 1. The similarity $S(p,x_j)$ between a probe p and an exemplar j is assumed to be a nonlinearly decreasing function of their distance (d_{pj}),

$$S(p, x_j) = e^{-d_{pj}}. \qquad (7)$$

In the experiment, herb species with a criterion above 590 and below 510 are not included in the training phase. This makes it possible to distinguish between the models as they provide different predictions (Figure 1). In the training phase, all exemplars have effectiveness between 510 and 590. If participants have estimated the correct cue weight for each cue they should compute the most extreme judgments for the extreme exemplars that are left out in the training phase. More specifically, whenever participants have correctly identified the sign of each cue (i.e., whether it increases or decreases the criterion) they should make more extreme judgments for the exemplars with all cues at their maximum and the exemplars with all cues at their minimum, as illustrated on the left-side of Figure 1. By contrast, the exemplar model computes a weighted average of the criteria between 510 and 590 stored with the exemplars and this can never produce a value outside of this observed range (Erickson & Krusckhe, 1998; but see DeLosh et al., 1997). Moreover, because of the nonlinear similarity function of the GCM the most extreme judgments tend to be made for the second to most extreme exemplars. For these exemplars the judgment is dominated by retrieval of the identical stored exemplars and these identical exemplars are the most extreme that were encountered in the training phase. These predictions are illustrated on the right side of Figure 1.

For new exemplars in the mid range of the criterion cue abstraction suggests no systematic difference between new exemplars and old exemplars matched on the criterion: the cognitive process is the same regardless of whether a specific exemplar has been encountered before or not. The exemplar model predicts more precise judgments for the old exemplars because for these exemplars the participants can benefit from previous identical exemplars with the correct criterion c.

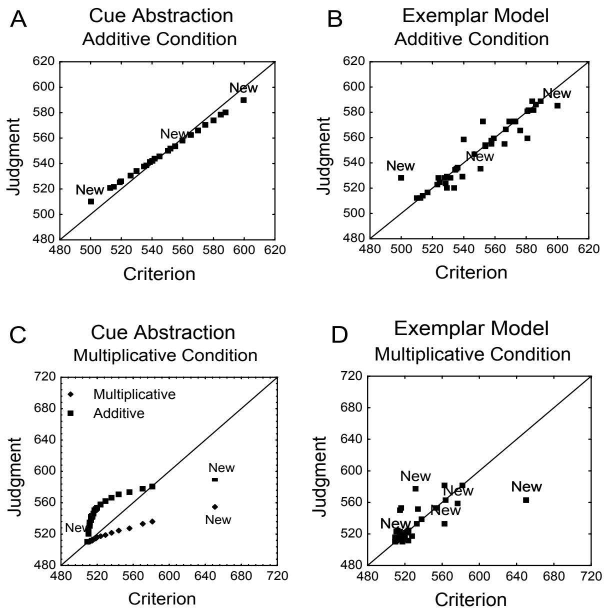

Figure 1: Predictions by cue-abstraction model (CAM) and exemplar model (EBM) in additive and multiplicative task environments. Panel A: CAM in additive task environment (with slightly regressive weights, 3.2, 2.4, 1.6, & .8). Panel B: EBM in additive task environment (s = .25 and h = 10). Panel C: Additive [CAM(A)] and multiplicative [CAM(M)] cue-abstraction models in multiplicative task environment (with weights, 3.2, 2.4, 1.6, & .8). Panel D: EBM in multiplicative task environment (s = .25 and h = 10). The choice of values for the parameters is arbitrary and only used for illustrative purposes.

The Experiment

In the experiment we manipulated whether participants were confronted with a task that involved additive or multiplicative cue combination. For the reasons outlined in the introduction, we predicted that the additive task (Eq.1) should promote explicit cue abstraction with additive cue integration (Eq. 3). A multiplicative task (Eq. 2) should cause a shift to a qualitatively different process, that is, to exemplar memory (Eq. 5).

The sign of the linear relations between cues and criterion was also manipulated. For half of the participants all four cues were positively related to the criterion and for half of the participants two cues were positively and two cues were negatively related to the criterion. In line with the assumptions of ∑, we predict that in an additive task, whether cue directions are negative or positive should not affect the ability to perform cue abstraction. In a multiplicative task, both with homogeneous and heterogeneous cue directions, exemplar memory is predicted to prevail over cue abstraction.

Method

Participants

Thirty two undergraduate students volunteered, receiving a payment of 60-99 SKr, depending on their performance. Twenty participants were male and 12 were female, all in the age between 20 and 32.

Materials and Procedure

The experiment consisted of a training phase and a test phase. In the training phase, the participant learned to judge the effectiveness of each species of the herb by means of outcome feedback. The effectiveness was measured as the amount (mg) of the fictitious chemical substance *Ranulin*. In the training phase, the effectiveness varied between 510 and 590 mg. The species were shown as four written propositions on a computer screen. At each trial in the training phase, the participant was to answer the question “How many milligrams of Ranulin does this specie contain?”. After giving a response they received the correct answer: “This specie contain 540 milligrams of Ranulin”. The four dimensions were: number of weeks in bloom, the optimal amount of iron in the ground, the degrees of latitude where it does well, and the amount of water it emits per leaf area. Each dimension varied “pseudo-continuously” in 11 equidistant steps that ranged between 0 and 10, yielding a total of 11^4 different exemplars. In the training phase, a random sample of 300 exemplars was drawn from this distribution and shown to the participant. A pause of two minutes was given to the participant after the first 150 trials.

In the test phase, participants were to judge the effectiveness of the species of the herbs but now *without* outcome feed-back. In the test phase, new exemplars were included. The test phase consisted of 44 judgments of *a*) 20 randomly chosen old exemplars shown in training, *b*) 20 randomly chosen new exemplars, drawn from the training distribution and *c)* 4 extreme exemplars, with criterion values outside the training range (eg. the exemplars with cue values [0,0,0,0] and [10,10,10,10])

In the condition with heterogeneous cue directions, for half of the participants, negative sign was assigned to the cues with objective weight 4 and 2, and for half of the participants to the cues with weights 3 and 1.

Dependent Measures

The measure of performance is *Root Mean Square Error* (*RMSE*) of judgment (i.e., between judgment and criterion). Measures of model fit are *the coefficient of determination* (r^2) and *Root Mean Square Deviation* (*RMSD*) between predictions and data from the test phase.

Results

A two-way ANOVA with environment (additive vs. multiplicative) and cue directions (homogeneous vs. heterogeneous) as between-subject factors shows two main effects on RMSE (Table 1), but no interaction. In the additive condition performance is significantly better ($F(1.30) = 20.36$; $MSE = 32.36$; $p = 0.000$). Also, when the cue directions are homogeneous RMSE is lower compared to when the cue directions are heterogeneous ($F(1.30) = 20.36$; $MSE = 6.37$; $p = 0.018$)

Table 1: Judgment performance in the experiment as measured by the Root Mean Square Error (RMSE) between judgment and criterion.

		Condition		
Cue directions	Index	Add.	Mult.	Mean
Homogeneous	RMSE	11.69	21.23	*16.46*
Heterogeneous	RMSE	17.21	25.90	*21.56*
Mean	RMSE	*14.45*	*23.56*	

Mean judgments are shown in Figure 2. In the additive homogeneous condition the judgments are a linear function of the criterion and no extra- or interpolation effects are visible. The best fitting regression lines for old and new judgments coincides.

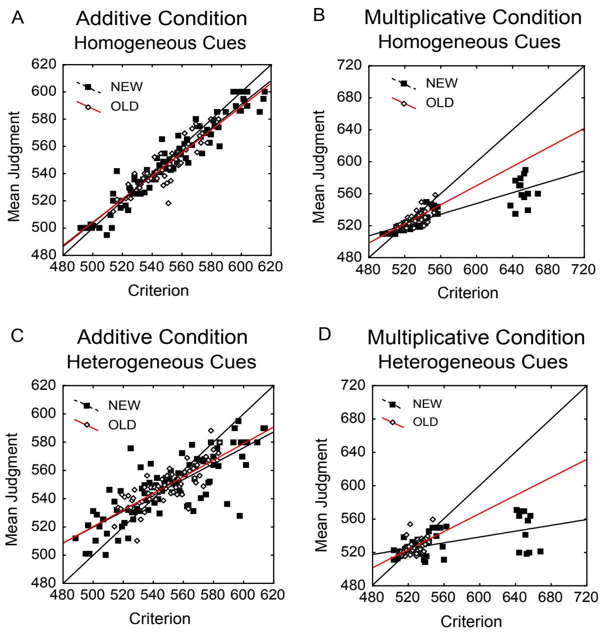

Figure 2: Mean judgments for the different conditions. Panel A: additive, homogeneous. Panel B: multiplicative, homogeneous. Panel C: additive, heterogeneous. Panel D: multiplicative, heterogeneous. Best-fitting regression lines are based on a) the old exemplars seen in training or b) the new exemplars introduced at test.

In the multiplicative homogeneous condition the judgments clearly deviate both from the identity line and the best fitting regression line based on old exemplars. Although the judgments are a positive function of the criteria in the training range there is evidence for an inability to extrapolate. The judgments are not extrapolated beyond the range of training.

In the additive heterogeneous condition (Figure 2C) the judgments are still close to the optimal judgment line, although there are signs of extra- and interpolation effects. In the multiplicative heterogeneous condition (Panel D) the mean judgments are a positive function of the criterion, but the inability to extrapolate is obvious.

Quantitative model predictions were obtained by fitting the models in the introduction (Eq. 3, 4 & 5) to the latter half of the *training phase* with Mean Square Error between predictions and data as the error function (Juslin et al., 2003; manuscript).

Table 2: Model fit: Root Mean Square Deviations (RMSD) and r^2 for the additive cue-abstraction model (CAM(A)) the exemplar model (EBM) and the multiplicative cue-abstraction model (CAM(M)) in the four conditions.

	CAM(A)		EBM		CAM(M)	
Cond.	r^2	RMSD	r^2	RMSD	r^2	RMSD
Add:						
Homogen.	.77	10.83	.75	13.73	-	-
Heterogen.	.53	12.67	.51	13.40	-	-
Mult:						
Homogen.	.21	28.26	.74	8.50	.70	12.20
Heterogen.	.18	48.62	.34	19.65	.32	18.97

The models were thus fitted to data from the training phase and applied with these parameters to the wider range of herb species in the *test phase*. This implies cross-validation for exemplars presented in training and genuine predictions for new exemplars. To capture individual differences, the models were applied to individual data. Table 2 shows the mean fit for the three models. In the additive condition cue abstraction is the overall dominant model, regardless of the cue directions. In the multiplicative condition exemplar memory describes the data best with regard to r^2 and the multiplicative cue-abstraction model yields a smaller mean *RMSD*. The rather low fit of all three models in the heterogeneous conditions may be explained by larger noise in these data since this task is presumably more difficult than the homogeneous task. Figure 3 shows the proportion of participants best accounted for by each model in terms of *RMSD*. In the additive homogeneous condition most of the participants are accounted for by the cue-abstraction model. In the multiplicative homogeneous condition the reverse is true, namely that the exemplar model produces the best explanation. In the additive heterogeneous condition the proportion of participants explained by the cue abstraction model decreases. In the multiplicative heterogeneous condition, as hypothesized the exemplar-based model continues to provide the best explanation of data for most of the participants. The multiplicative cue-abstraction model describes some of the participants in both the homogeneous and the heterogeneous multiplicative tasks.

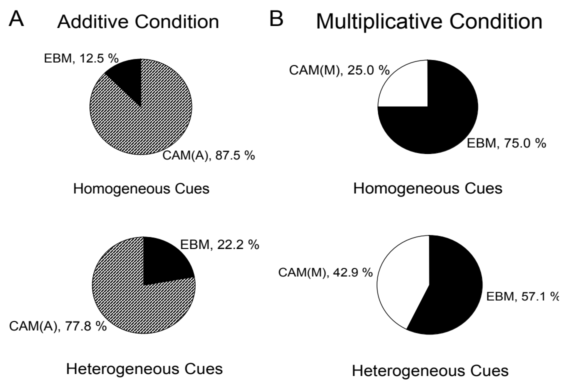

Figure 3: The proportion of participants accounted for by any of the three models in the additive and the multiplicative conditions in terms of *RMSD*. Panel A: additive condition. Panel B: multiplicative condition.

Discussion

The results reported in this paper support the assumptions made by ∑ that multiple-cue judgment processes conceal an effective division of labor between qualitatively distinct cognitive processes (Juslin et al., manuscript). Cognitive modeling supports the hypothesis that in a multiple-cue judgment task where the cues combine by addition, ∑ is fed with representations in form of abstracted knowledge of the relations between cues and criterion. On the other hand, in an environment where the cues relate to the criteria by a multiplicative function we seem to be equipped with no means to explicitly abstract the underlying structure. In such tasks, people seem to resort to the back-up process of exemplar-memory.

The fact that exemplar-memory plays part also in additive tasks is not a coincidence, since both processes allows accurate performance in training. That the multiplicative cue-abstraction model provide an explanation for some of the participants in the multiplicative task is more surprising. This is probably an effect both of large noise in data and of its high correlations to the exemplar model. Figure 2 B & C yields no evidence for successful extrapolation beyond the range of training.

The bad performance in the multiplicative heterogeneous condition, together with the low fit of the models makes it unfair to draw conclusions regarding what cognitive process that has dominated the judgments in this condition. What makes this task difficult to learn? A tentative hypothesis would be that in training there is a bias towards the abstraction of specific rules (eg. *rule bias;* see for example Ashby et al., 1998; Juslin et al., 2003; manuscript). Presumably, a period of extensive hypothesis testing is therefore taking place at beginning of training. However, the multiplicative heterogeneous task may be inductive of more extensive hypothesis-testing procedures. The back-up of exemplar memory is thus postponed, and thereby learning may be impaired.

An interesting approach to the interpretation of the data would be to consider how an exemplar-model augmented with linear extrapolation would account for the results (see *EXAM*; DeLosh et al., 1997; Busemeyer et al., 1997, for results on single-cue learning). *EXAM* suggests that, although learning has been in the form of exemplar-memory, abstraction of cue-criterion relations is possible at test. When encountered with a new exemplar at test, familiar exemplars and their stored criterion are retrieved from memory. An extrapolated judgment for the new exemplar is then made possible through linear regression based on the old exemplars. How this model explains the data reported in this paper remains to be tested, although a first qualitative evaluation of the data in Figure 2 can be made. Given the data in the additive condition, *EXAM* is likely to produce the same fit as the cue-abstraction model. In the multiplicative condition *EXAM* would predict no difference between the regression made on old exemplars and the regression made on new exemplars. This difference is however apparent in the data in Figure 2 (Panel B & C) and thus suggests the refutation of *EXAM* in favor of EBM.

The main interpretation to be drawn from the results reported in this paper is that the human judge, under the constraints imposed by ∑, adapt to different task structures by means of representational shifts. This highlights how the task is a powerful predictor of cognitive process in human multiple-cue judgment.

Acknowledgments

Bank of Sweden Tercentenary Foundation supported this research.

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105,* 442-481.

Bott, L. & Heit, E. (2004). Non-monotonic Extrapolation in Function Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30 (1)*, 38-50.

Brehmer, B. (1994). *The psychology of linear judgment models. Acta Psychologica, 87.* 137-154.

Busemeyer, J. R., Byun, E., DeLosh, E. L., & McDaniel, M. A. (1997). Learning Functional Relations Based on Experience With Input-Output Pairs by Human and Artificial Neural Networks. In K. Lamberts & D. Shanks (Eds.), *Knowledge, Concepts and Categories*. London: UCL Press.

Cooksey, R. W. (1996). *Judgment analysis. Theory, methods and applications.* San Diego. Academic Press, Inc.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 968-986.

Einhorn, J. H., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Regression models and process tracing analysis. *Psychological Review, 86*, 465-485.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107-140.

Hammond, K. R., & Stewart, T. R. (Eds.) (2001). *The essential Brunswik: Beginnings, explications, applications*. Oxford University Press: Oxford.

Juslin, P., Karlsson, L. & Olsson, H. (manuscript) *Exemplar Memory in Multiple-Cue Judgment: A Division of Labor Hypothesis.*

Juslin, P., Olsson, H., & Olsson, A-C. (2003) Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General132 (1),* 133-156.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning, *Psychological Review, 85*, 207-238.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 104-114.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin and Review, 7,* 375-402.

Smith, J. D, & Minda, J. P (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26, 3-27.*