### Title
JUDGEMAP - Integration of Analogy-Making, Judgment, and Choice




### Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

### ISSN
1069-7977



### Authors
Kokinov, Boicho
Petkov, Georgi



### Publication Date
2006

Peer reviewed

# JUDGEMAP – Integration of Analogy-Making, Judgment, and Choice

**Georgi Petkov (gpetkov@cogs.nbu.bg)**
**Boicho Kokinov (bkokinov@nbu.bg)**

Central and East European Center for Cognitive Science, Department of Cognitive Science and Psychology,
New Bulgarian University, 21 Montevideo Street, Sofia 1618, Bulgaria

**Abstract**

This paper illustrates how mechanisms initially designed for explaining analogy-making can also model judment and choice and account for contextual effects on these processes. The JUDGEMAP model is presented as well as three simulations that replicate some well known contextual effects in judgment and choice. It is demonstrated how the same basic mechanisms, designed for analogy-making can be responsible for seemingly unrelated phenomena, like the frequency effect in judgment and the concave form of the utility function; the preference for the middle ratings in judgment, the nonlinear form of subjective probability; and the effect of preference reversal.

## Introduction

JUDGEMAP is a model of judgment and choice that assumes that the structure mapping ability is fundamental for human cognition. It successfully replicates many contextual effects that are reported in the literature. Some of these effects seem unrelated but it turns out that the same basic mechanisms can produce them. The phenomena, which are simulated with JUDGEMAP, are briefly presented below.

Kahneman & Tversky (1979, 1983) demonstrated that the subjective value (utility) is a concave function of money. On the other hand, Parducci & Perret, (1971) demonstrated the frequency principle in human judgments, i.e., the tendency people to use all available categories almost equal times. These two facts seem unrelated and there is no model that accounts for both of them. However, we assume that these phenomena may result from one and the same mechanism, that is: the pressure for one to one mapping which is fundamental for analogy-making. At the same time, we assume that this mechanism has evolved for analogy, not for the processes of judgment and choice.

Kahneman & Tversky (1983) demonstrated that the value of a gamble is not a linear function of the probability of winning. More precisely, an increase in the range of small and large probabilities appears to have a larger effect than the equal increase in the middle-range probabilities. On the other hand, Petrov & Anderson (2005) demonstrated that when choosing, people prefer to use the middle ratings more often than the extreme ones. In other words, the overall distribution of the ratings is not uniform but with a peak in the middle even when the distribution of the stimuli magnitudes is uniform.

Sharif, Simonson, & Tversky (1993) demonstrated preference reversals when people chose among the same alternatives under different circumstances. They proposed the idea that people just count the number of justifications for each alternative. However, people use in their choices not only the ordering relations between the stimuli but also the exact values of those stimuli. This cannot be explained by simple enumeration of the justifications, but requires a kind of weighting.

There are many models that describe these effects but few are able to explain *why* they appear. Some models postulate a formula that reproduces one or another contextual effect (Parducci, 1965); other models look for such mechanisms at the sub-symbolic level (Busemeyer, Johnson, 2004). Our goal was not to create special mechanisms that can reproduce all these effects. Instead, we began with several general assumptions about cognition, in particular the ideas that memory is decentralized and associative; that analogy-making is at the core of many cognitive processes; and that context is not just a source of noise but is a necessary condition for effective and flexible cognition. After that we designed the JUDGEMAP model on the basis of these principle and the mechanisms already developed for modeling analogy-making and at the end we demonstrated that the model is able to replicate a number of seemingly unrelated phenomena.

### The DUAL architecture and the AMBR model

The principles that we assume to be basic for human cognition are organized in the DUAL architecture (Kokinov, 1994b, c). It is based on decentralized associative memory, and is designed to model the context sensitivity of human cognition. The architecture combines symbolic and connectionist representations and processes. From the symbolic point of view, DUAL is an associative network of huge number of micro-agents running in parallel. The situations and the concepts are represented in a decentralized way. Even relatively simple concepts and instances are represented by a large coalition of interconnected DUAL-agents. From the connectionist point of view, the same network is a connectionist network. Unlike many connectionist models, however, the activation does not represent the meaning but the *relevance* of the respective agents to the current context. The activation spreads from two special nodes INPUT and GOAL, representing the sensory input and the current goals, respectively. The overall pattern of activation changes continuously in response to the changes in the environment and in the goals and thus, it represents the current context.

In order a system to be flexible, it should be potentially able to search many possible different paths. In order to be effective, however, in any certain moment it should explore relatively small number of paths. The DUAL answer to this trade-off is that the context determines which paths are relevant and should be explored in any given moment. Actually, the relevance of the DUAL-agents, represented by their activation level, controls the speed of symbolic processing. The most active agents work faster, the less active ones slower, and the inactive ones do not work at all. There is a certain threshold and all agents, which activation exceeds this threshold form the Working Memory (WM) of the model. The agents in the WM exchange symbols among each other, create new agents and new connections. However, all symbolic operations are performed locally, without any central executor and the global behavior of the system emerges from these local interactions.

The AMBR model (Kokinov, 1994a; Kokinov, Petrov, 2001) is a DUAL-based model for analogy-making. Its main mechanisms involve detecting local similarities between entities and these similarities serve as justifications for creating hypotheses for possible correspondences between the respective entities; the local mappings grow up, involving their neighbors, keeping the structural correspondence between the relations.

A marker-passing mechanism is responsible for finding semantic similarities between objects in Working memory. The markers originate from objects and relations, spread up through the conceptual class hierarchy, meet each other, and thus justify creation of hypotheses for correspondence between the elements of the target situation and memorized elements from past situations. However, it is important to stress that the markers are passed with speed that reflects the relevance of the respective elements and thus the system is fully prevented from the so-called 'combinatory explosion'.

The local correspondences 'grow up' under the pressure for structural consistency. A set of mechanisms are responsible for finding such structural justifications for novel hypotheses. For example, the hypotheses for correspondence between two relations create hypotheses for correspondence between their respective arguments; the correspondences between instances create correspondences between their respective concepts, etc.

The consistent hypotheses support each other; the inconsistent ones compete with each other. The pressure for one-to-one mapping results in building inhibitory links between hypotheses that relate one and the same agent with several possible partners. Thus, dynamically, a constraint satisfaction network of hypotheses, interconnected with the network of concepts and events emerges.

As a final result of all these mechanisms, the most consistent analogy wins against its competitors.

## JUDGEMAP model

**Judgment on a scale:** JUDGEMAP considers the process of judgment on a scale as an emerging result from the work of several overlapping processes. First, a comparison set is formed via the spreading activation mechanism and the target stimulus (or stimuli) is *included* in this set. Second, some ordering relations between the elements in the comparison set are recognized. Third, a mapping between the comparison set and available scale labels is established.

The simulation begins when the representation of the target stimulus (stimuli), possibly together with some contextual elements are attached to the INPUT node, and the representation of the instruction, namely the correspondence "higher magnitude corresponds to higher rating" is attached to the GOAL node. Then the activation spreads through the network, it activates the properties of the stimuli and their respective concepts. Possibly, some prototypes of these concepts, and some recently used instances are also activated. Thus, gradually, many exemplars and prototypes enter the WM and form the comparison set.

There are comparison relations that are activated by the GOAL node and they have to find any manifestations of these relations in the environment and to create new temporary instances. For example, if the task is to judge the overall utility of gambles, represented with their probability to win and their profit, the comparison relation 'higher profit' compares the profits of the gambles in the comparison set, and creates new relations of the following form: 'The profit of gamble A is higher than the profit of gamble B'. Independently, the comparison relation 'higher probability' does the same with the probabilities of the gambles. Following the main DUAL principles, all these operations are performed locally on the basis of exchange of symbolic messages between neighboring agents.

However, in order to judge the absolute magnitudes, it is not enough to know only the relative ordering relations between the stimuli. The so-called second order comparison relations compare the differences. For example, if the system has already recognized that the profit of 100 is higher than the profit of 50, and that the profit of 20 is higher than the profit of 10, than a second-order comparison relation could be produced that will represent the fact that the first *difference* is larger than the second one. The second order comparison relations allow the properties of the interval scales to be computed using only local computations. Without the higher order relations the model would not be able to differentiate, for example, the set of magnitudes 10, 20, and 100 from the set of magnitudes 10, 50, and 100.

All these comparisons between magnitudes serve for justifications to create hypotheses for possible correspondences between the stimuli in the comparison set and the available scale labels, keeping in correspondence the structure of the ordering relations among the stimuli and the respective structure of the ordering relations among the scale values. Thus, the work of JUDGEMAP can be viewed as a process of making *forced analogy*.

Since the scale labels are instances of the respective integer numbers, the neighboring ratings are interconnected with relations 'next' and 'previous'. The activation spreads over this chain in both directions. Thus if the stimulus A is already mapped onto the rating "3", and stimulus B is greater than A, then under otherwise equal conditions the stimulus B will be mapped onto 4 rather than 5 or 6, since the label "4" is already active from the previous stimulus (A) mapping. This produces an assimilation effect towards the last rating.

All mapping principles are inherited from the AMBR model, including the pressure for one-to-one mapping.

**Choice:** JUDGEMAP assumes that the same basic mechanisms underlie judgment and choice (see Medin, Goldstone, Markman, 1995, for a related view),. Actually, the process of choice-making is considered as a process of judgment on a two-point scale. Several alternatives are attached to the INPUT in order to be judged on a two-point scale. The only difference with the judgment task is that the system does not report the winners for all target alternatives, but the winner for the higher rating, i.e. the rating-label 'chosen' becomes a driver.

The alternatives can be attached to the INPUT sequentially or simultaneously. For example, let two gambles, named A and B are attached to the INPUT simultaneously. The relevant to the task dimensions are then activated through the spreading activation mechanism. The irrelevant dimensions, of course, would also be activated and as a consequence, under certain circumstances the irrelevant dimensions can also play role in the process.

The two-point scale consists of two 'grades' – 'chosen' and 'rejected', and the relation that the first rating is 'higher' than the second one. This 'scale' is also attached to the INPUT. The corresponding relation 'better stimulus corresponds to higher rating' is attached to the GOAL. The understanding of the term 'better stimulus' is coded with links from this relation to more specific ones, e.g., 'higher profit' or 'higher probability'. The first hypotheses between gambles and ratings emerge just after the first comparisons; then the structure mapping mechanism creates more hypotheses. Inhibitory and excitatory links are created between the inconsistent and the consistent hypotheses, respectively.

The hypotheses that connect the grade 'chosen' with different gambles compete with each other. However, they do not wait until all possible hypotheses emerge. Instead, if one of these hypotheses keeps being the leader for sufficiently long period of time, it is promoted to become a winner, and is interpreted as the response of the model.

## Simulations

In previuos studies (Petkov, 2005) it was demonstrated that JUDGEMAP replicates many of the empirical phenomena in human judgments, for example, the sequential assimilation effect, the range and frequency effects. JUDGEMAP produced also some new predictions about the role of the irrelevant dimensions in judgment that were empirically tested and comfirmed (Kokinov, Hristova, Petkov, 2004). In the current study the role of the retrieval and mapping mechanisms in judgment and choice are explored. More preciesly, it is demonstrated how these mechanisms that were designed for analogy-making may influence judgment and choice as well. Moreover, these mechanisms can reproduce contextual effects in seemingly unrelated to analogy-making tasks.

### Simulation 1

The first simulation shows that a mechanism for one-to-one mapping designed explicitly for analogy-making may model successfully two additional, seemingly unrelated phenomena in human judgment and choice, namely that:

- people use all available ratings almost equal number of times in their judgments, a phenomena known in the field of judgment as the *frequency principle* proposed by Parducci & Perret (1971);
- subjective value (utility) is a *concave function* of money, a phenomena well known in the field of choice (Kahneman & Tversky, 1979, 1983).

These two phenomena may resemble each other if they are reduced to judgment of skewed sets. It could be speculated that low prices dominate in the environment and as a consequence the comparison set formed when judging prices would be positively skewed. Since JUDGEMAP replicates the frequency effect successfully (Petkov, 2005) it could be argued that the model will manage to replicate the concave utility function as well.

For the purpose of the current simulation, a set of prices was randomly extracted from Internet advertising sites and were given to the model to be judged on a seven-point scale. Six categories (home & garden, toys & baby, office, books, clothing, sports) and 15 products from each category were randomly chosen, thus forming a set of 90 prices (47 different price-values). The distribution of this set turned out to be the following: 32 prices less than \$10, 38 prices between \$10 and \$50, 11 prices between \$50 and &200, and 9 prices were larger than \$200. All 90 prices were judged sequentially. The procedure was repeated 100 times, varying the order of stimuli presentation in a random way.

The initial moment – judgment of the first stimulus – needs a special comment. The system must judge the stimulus in a 'vacuum' – without anything to compare with. Our belief is that human beings do not face such a situation since they have extremely huge knowledge bases and in all cases, they are able to retrieve (or to construct) something similar to the target. However, if (and only if) such a case appears the system creates the first hypothesis without any reason (justification). It just takes the most active rating and maps the target onto it. Suppose, for example, that the first price is judged with 4 and let the next price be lower than the first one. JUDGEMAP compares it with the previous one and creates a new comparison relation that in turn serves as a justification for new hypotheses (Fig. 1). The order of creation of these new hypotheses reflects the activation of the scale elements. Since the previously used rating was 4 it would be the most active rating and because of the chain-like organization of the scale representation the rating 3 would be more active than 2 and 1. Thus, the first new hypothesis would connect the second target stimulus with the rating 3. It could win the competition even before the creation of any new hypothesis.

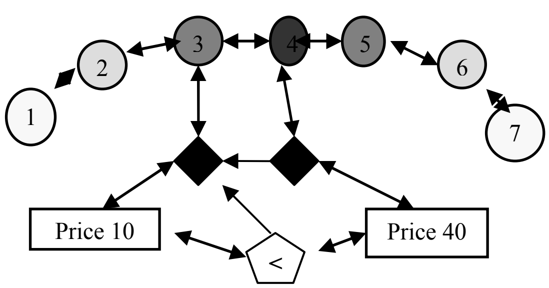



Figure 1: The comparison relations are justifications for new hypotheses for correspondence.

Since the overall set of prices is skewed (more low prices than high prices), the competition for the low ratings would be higher (Fig.2). Suppose, for example, that there are two hypotheses for the target stimulus 'price 50' – to judge it with 3 or with 4. If there are many memorized stimuli that were judged with 3, being everything else equal, the rating 4 would be the winner, because of the inhibition received from its rivals. The overall result of this mechanism would be an overestimation of all stimuli from a positively skewed set.

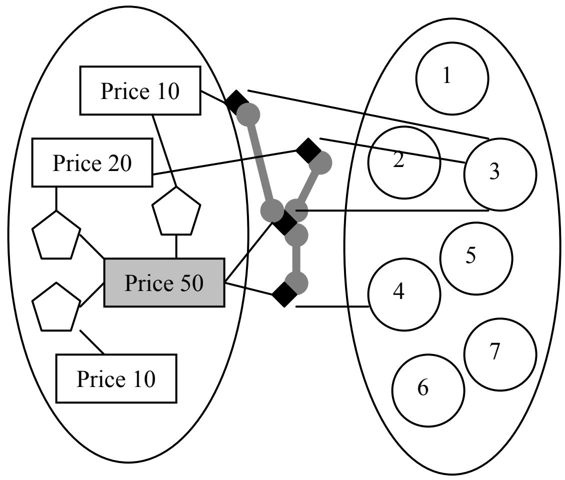



Figure 2: The constraint for one-to-one mapping creates inhibitory links (with gray) between the hypotheses that connect different stimuli with one and the same rating as well as different ratings with one and the sane stimulus.

The final result of the first simulation was that the ratings given by the model formed a concave function of the money (mean difference from the straight line for all 47 different price values is 2.71, $t(46) = 19.11$, $p<0.01$), in accordance with the empirical data.

## Simulation 2

The second simulation starts with the assumption that two seemingly different facts are actually manifestation of one and the same effect, that is:

- On one hand, Petrov & Anderson (2005) demonstrated that people prefer to use the middle ratings more often than the extreme ones even when the distribution of the stimuli magnitudes is uniform.
- On the other hand, Kahneman & Tversky (1983) demonstrated that the value of a gamble is not a linear function of the probability of winning. More precisely, an increase in small and in large probabilities appears to have a larger effect than the equal increase in middle-range probabilities. In order to test whether JUDGEMAP distributes ratings in a way similar to humans, the model was tested with a set of 112 lines, uniformly distributed over their length. Each line was represented with a line-agent and with an agent that represents its length, coded with a real number. The task of JUDGEMAP was to judge the length of the lines on a seven-point scale. All 112 lines were given to the system sequentially in a random order. As demonstrated in previous studies (Petkov, 2005) JUDGEMAP successfully simulates the sequential assimilation effect towards the recently used grade, and a smaller in size contrast effect with respect to the previous line length. One additional effect, however, was that the ratings produced by JUDGEMAP were

not uniformly distributed, but had a peak in the middle range (Figure 3). The standard deviation of all ratings, obtained from the simulation was 1.34, whereas for assumed uniform distribution it should be 2.01 ($\chi_2(111)=49.53$, $p<0.01$).

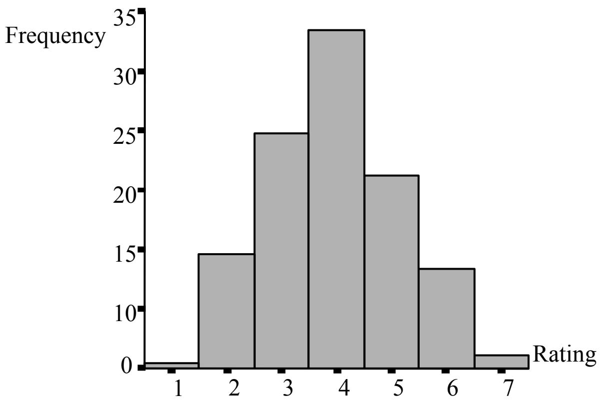



Figure 3: Distribution or ratings when a uniform set of stimuli was judged by JUDGEMAP.

An analogous simulation was performed using gambles as stimuli. A set of 100 gambles was designed. The first gamble proposed 1% chance to win 1000, the second one – 2% chance to win 1000/2, etc. The last gamble proposed 100% chance to win 10. Thus, the expected utility for all 100 gambles was equal. JUDGEMAP judged how large the probability for winning is on a seven-point scale (the profit was irrelevant to this task). All 100 stimuli were judged 100 times in a random order each time. Not surprisingly, the results again were analogous. (The standard deviation obtained from the simulation was 1.21, whereas if a uniform distribution is assumed it should be 2.01, $\chi_2(9999)=3660$, $p<0.01$).

The same data are illustrated in the fig.4 in the form of money/judgment function.

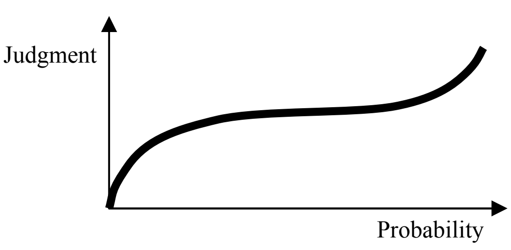
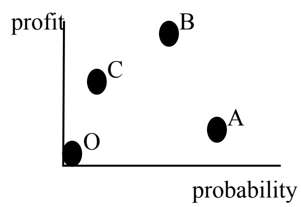



Figure 4: The concave form of the money/judgment function.

The preference for the middle ratings may result in the observed overestimation of the small probabilities and underestimation of large probabilities that Kahneman & Tversky (1979) called the $\pi$-function. This effect naturally emerges from the basic mechanisms, designed in DUAL: firstly, the activation level of the ratings controls the speed of the creation of the hypotheses about those ratings and secondly, the chain-like organization of the ratings that makes the middle ratings dominating in their activation level (see Fig.1). As a result, JUDGEMAP rarely gives extreme grades because hypotheses for them are usually created at later moments. For example, suppose that a certain probability was rated 4. If the next probability had to be higher, then the first created hypothesis would most probably be 5, the second one would be 6, etc. Being created first, these hypotheses had an advantage, and only a justification with very high activation level should be established, in order the extreme rating of 7 to win.

## Simulation 3

This simulation highlights the role of the justifications produced by JUDGEMAP for each hypothesis. It illustrates also the application of the model to making choice and demonstrates the preference reversal observed by Sarif, Simonson, & Tversky (1993). Two gambles, A and B, presented with their probability for winning and their profit,, were used for stimuli in all trials, while in half of the trials an additional gamble C was presented.




The exact values of the probabilities and profits were randomly chosen but the relative position of the stimuli in the probability/profit space remained the same (Figure 5).

Figure 5: Relative positions of the stimuli.

The probability of the gamble A was always larger than the probability of B. The profit of A was always smaller than B. JUDGEMAP is sensitive to the ordering relations between the stimuli and takes also into account the differences between the absolute magnitudes using second-order relations. In order to do so, however, at least three stimuli should enter in the comparison set. For this reason a third stimulus, named O, with probability and profit zero was added. This additional "gamble" retrieved (constructed) from memory equipped the model with ability to detect whether A and B differ much more along one of the dimensions in comparison to their difference on the other dimension. Thus, 100 different sets of gambles A and B were designed and given to the model for making a choice between them. In each task the two gambles were attached to the INPUT simultaneously. When JUDGEMAP chose between these stimuli, it preferred A – 49 times, and B – 51 times. After that, a third gamble, C, was added to each of the 100 tasks. Its profit was randomly chosen to be between the profits of A and B. Its probability was always smaller than the probability of A and B (Figure 5). As a result, JUDGEMAP preferred A – 37 times, B – 60 times, and C – 3 times. The existence of the alternative C became crucial for preferring B over A ($\chi_2$ (1, N = 97) =5.45, p=0.02). JUDGEMAP obtained this result because its decisions are always based on justifications, which property was inherited from the AMBR model. There were more reasons to prefer B than A (B was better than C on to two dimensions, A was better than C on one dimension only). Since the higher-order relations are sensitive to the absolute magnitudes as well, JUDGEMAP did not choose *always* the gamble B in accordance with the empirical data while the reason-based choice theory (Sarif, Siminson, & Tversky, 1993) could not explain why B is not always preferred to A in this case.

## Conclusions

The JUDGEMAP model of human judgment and choice is presented. It is based on the cognitive architecture DUAL and is integrated with the AMBR model for analogy- making. According to the model, the process of judgment is a process of mapping between a dynamically constructed comparison set and the set of the ratings. During the judgment process some ordering relations between the elements of the comparison set are recognized. These relations serve then as justifications for hypotheses formation for correspondence between the stimuli and the available ratings. The consistent hypotheses support each other; the inconsistent ones compete with each other. After the relaxation, one of the hypotheses about the target stimulus wins and is interpreted as a response of the system. According to the model, the process of choice making is a kind of judgment. However, when choosing, the driving force is to find the stimulus that corresponds to the highest rating. In the three simulations some contextual effects in human judgment and choice were replicated. It is demonstrated how the pressure for one-to-one mapping, designed for the AMBR model for analogy-making, can be responsible both for the frequency effect in judgment and for the non-linear form of the function between the subjective utility and the amount of money. The fact that people use the middle ratings more often than the extreme ones in judgment is explained in terms of the dynamic mechanism for hypothesis creation, which is highly sensitive to the relevance of the items. In addition, the same basic mechanism is assumed to be responsible for the fact that an equal increase in the probability has larger effect in the extreme probabilities than in the middle ones. Finally, JUDGEMAP also accounts for the preference reversal effect in choice. The mechanism for creating only justified hypotheses is responsible for this effect. All these mechanisms, however, are not created for the purpose to obtain these concrete results. Instead, they are inherited from the AMBR model. This supports the assumption that the process of structural mapping is at the core of human cognition and people adapt the available basic mapping mechanisms to other tasks, like judgment or choice.

## References


Busemeyer, J., Johnson, J. (2004) Computational models of decision making. *In: D. Koehler & N. Harvey (Eds.) Handbook of Judgment and Decision Making*. *Blackwell Publishing Co., pp. 133-154.*

Hofstater, D. (2001*)* Analogy as the Core of Cognition. *In Dedre Gentner, Keith J. Holyoak, and Boicho N. Kokinov (eds.)* The Analogical Mind: Perspectives from Cognitive Science. *Cambridge, MA: The MIT Press/Bradford Book.*

Kahneman, D., Tversky, A. (1979) Prospect Theory: An Analysis of decision under risk. *American Psychologist, vol. 39 (4), pp.341-350.*

Kahneman, D., Tversky, A. (1983) Choices, Values, and Frames. *Econometrica, vol. 47, pp. 263-291.*

Kokinov, B. (1994a) A Hybrid Model of Reasoning by Analogy. *In Holyoak, K., Barnden, J. (eds)* Advances in Connectionist and Mental Computation Theory, vol. 2: Analogical Connections. *Norwood, NY: Ablex.*

Kokinov, B. (1994b) The DUAL Cognitive Architecture. A Hybrid Multi-Agent Approach. *In A. Cohn (ed.) Proceedings of the Eleventh European Conference of Artificial Intelligence. London: John Wiley & Sons, Ltd.*

Kokinov, B. (1994c) The Context-Sensitive Architecture DUAL. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: LEA.*

Kokinov, B. & Petrov, A. (2001) Integration of Memory and Reasoning in Analogy-Making: The AMBR Model. *In: Gentner, D., Holyoak, K., Kokinov, B. (eds.) The Analogical Mind: Perspectives from Cognitive Science, Cambridge, MA: MIT Press.*

Kokinov, B., Hristova, P., Petkov, G. (2004) Does Irrelevant Information Play a Role in Judgment? *In: Proceedings of the 26th Annual Conference of the Cognitive Science Society,* (pp. 720-725) Erlbaum, Hillsdale, NJ.

Medin, D., Goldstone, R., Markman, A. (1995) Comparison and Choice: Relations between Similarity Processes and Decision Processes. *Psychonomic Bulletin & Review, vol. 2 (1), pp. 1 – 19*

Parducci, A. (1965) Category Judgment: A Range- Frequency Model. *Psychological Review, vol. 72 (6), pp 407-418.*

Parducci, A., Perret, L. (1971) Category-Rating Scales: Effects of Relative Spacing and Frequency of Stimulus Values. *Journal of Experimental Psychology, vol. 89 (2), pp 427-452.*

Petkov, G. (2005) Judgment as Mapping (JUDGEMAP2). *In: Proceedings of the Doctoral Consortium at the Fifth International and Interdisciplinary Conference on Modelling and Using* Context (pp. 95-104). Technical Report LIP 2005/007 of the Laboratoire d'Informatique de Paris 6, URL http://www.lip6.fr

Petrov, A., Anderson, J. (2005) The Dynamics of Scaling: A Memory-Based Anchor Model of Category Rating and Absolute Identification. *Psychological Review, vol. 112* (2), pp. 383-416.

Shafir, E., Simonson, I., Tversky, A. (1993) Reason – Based Choice. *Cognition, vol. 49, pp. 11-36.*