

Objective vs. Subjective Scales: The Challenge That the Scale Type Poses to the JUDGEMAP Model of Context Sensitive Judgment

Penka Hristova, Georgi Petkov, and Boicho Kokinov

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

Abstract. The paper presents a computational model of context sensitive judgment, called JUDGEMAP, which has been developed for modeling *judgment on a subjective scale*. This paper presents an attempt to apply the same model to the case of *judgment on an objective scale*. This is a big challenge since the behavioral data are showing the opposite type of effects. Thus we have repeatedly obtained *contrast effects of irrelevant information*, when judging on a subjective scale. In the experiment described here we obtained an *assimilation effect* in exactly the same conditions except that the scale was objective. Without any changes of the model we run the corresponding simulations and there are a good and bad news. The bad news is that we did not obtain the assimilation effect, but the good news is that the contrast effect disappeared. The paper discusses possible reasons for these results and possible ways to improve the model.

Keywords: judgment, context effects, assimilation, contrast, analogy-making, cognitive modeling, psychological experimentation, cognitive architectures.

1 Introduction

The present research focuses on the mechanisms behind contextually sensitive judgment. Many experimental results in the field of judgment demonstrate a particular shift in ratings depending on the context. The two main contextual effects usually found in such studies are called contrast (a shift in judgment away from the context) and assimilation (a shift in judgment toward the context).

Unfortunately, the existing empirical studies could not unambiguously determine the conditions under which contrast and assimilation appear. This uncertainty calls into question the contextual nature of judgment in general, that is, whether the two contextual effects are systematic or can be considered as a noise.

Definitely, there is no consensus among the researchers in the field about the causes of contextual effects in judgment. Is there one particular factor that produces contrast and assimilation under specific circumstances or, probably, there are several different factors that contribute to the contextual effects? Finally, it is also possible that these factors compete with each other and the outcome becomes not easily predictable.

A profitable step to overcome the current state of affairs may be an explanation of contextual influences in terms of the cognitive mechanisms that cause them. This was exactly the view-point taken by the JUDGEMAP project. Its aim was to suggest several mechanisms that may account for contextual effects in judgment on a subjective scale.

2 JUDGEMAP Model

Instead of modeling explicitly all effects in a judgment task, the JUDGEMAP project starts from a different perspective. We defined several principles that we assume to be central for the cognitive system, and then the attempt was to model the specific task for judgment on the basis of these principles.

We assume: First, that human memory is associative; second, that context is not just a source of noise, but is necessary for a flexible and at the same time effective reasoning; and third, that the ability for mapping (in particular, analogy-making) is not just a specific human capability, but is essential for cognition.

All these principles are integrated and implemented in the cognitive architecture DUAL [11, 12]. It is a multi-agent system that combines connectionist and symbolic mechanisms. The memory is represented with a localist semantic network. Each DUAL-agent 'stands for' something, but even very small pieces of knowledge are represented with a huge number of interconnected DUAL-agents. The sources of activation are two special nodes, called INPUT and GOAL, which represent respectively the environment and the goals of the system. The activation level of the agents, however, does not represent in any way the meaning of the agents, but their relevance to the current context. Thus, the context is defined as the overall pattern of activation that dynamically and continuously changes in response to the environment.

Each DUAL-agent can perform restricted number of symbolic operations. In particular, it can exchange messages with its neighbors; can modify its links; and can create new DUAL-agents. All symbolic operations, however, have a 'price' that should be paid by activation. Thus, the higher the relevance of a certain agent is, the faster it works [11, 12].

The AMBR model [9, 10, 13] is a DUAL-based model for analogy-making. After attachment of the description of the target situation on the GOAL node, the activation spreads from the particular instances to the corresponding concepts and close associations, and then back to some past instances, stored in memory. All relevant instances send markers, which spread upwards in the class hierarchy. If two markers cross somewhere, a hypothesis for correspondence between their origins emerges, i.e. the respective marker-origins are assumed to be analogous, because they have common super-class. The speed of marker spreading reflects the relevance of the respective agents and thus only small number of relevant hypotheses is created. There are other mechanisms for structural correspondence that ensure the systematicity of the mapping. On one hand, there are mechanisms for creation of other hypotheses. For example, if two relations are analogous, than their respective arguments should also be analogous; if two instances are analogous, then their respective concepts should be analogous, etc. On the other hand, inhibitory links between the competing hypotheses emerge. If one and the same element has two or more hypotheses, these

hypotheses are assumed to be inconsistent with each other because of the constraint for one-to-one mapping.

Thus, locally and asynchronously many interconnected hypotheses emerge. The result of the relaxation of the network of hypotheses is considered as a final analogy.

JUDGEMAP model [14, 23, 24] is a DUAL-based model of judgment on a scale, highly integrated with AMBR. It is designed under the assumption that the process of judgment on a scale is a process of mapping between a set of stimuli and the set of the available ratings. During this mapping the main constraint is that stimulus with higher magnitude should receive higher rating.

The representation of the stimulus that should be judged is attached to the GOAL node. The scale is represented with a chain of interconnected agents standing for the respective ratings and is attached to the INPUT node. The task of judgment is represented with several already created correspondences, attached to the GOAL node: 1) On one hand, the concept of the stimuli should correspond to the concept of the ratings; 2) On the other hand, the relation for higher magnitude should correspond to the relation for higher rating. Moreover, in order to ensure that the scale will have interval (not only range) properties, higher order relations that compare differences between magnitudes and differences between ratings are also built.

JUDGEMAP models judgment using some mechanisms inherited from the AMBR model and a few new ones. The model treats the judgment process as a result of a structural mapping between two sets. One of the sets is called comparison set. It consists of the target stimulus together with some similar memorized stimuli and the recently judged ones. The main mechanism for construction of the comparison set is the spreading activation mechanism. The second set consists of the available ratings. Note, however, that one and the same stimulus could be judged within different comparison sets, depending on the context, and thus it can receive different rating. Moreover, the context influences the set of the rating too. All ratings have their activation level that changes dynamically during the process of judgment. For example, the ratings of the recently judged stimulus would be more active than the rest. Because of the chain-like organization of the scale, this higher activation would spread to the neighbors of the recently used rating. In general, one of the important characteristics of the JUDGMAP model is its context sensitivity.

Thus, when the target stimulus is presented to the model to be judged, it serves as a source of activation that spreads through the conceptual system and extracts similar exemplars from memory, together with their ratings. The relevant to the task magnitude of the target stimulus is compared to the magnitudes of the elements from the comparison set and thus new agents, who represent the differences between stimuli, emerge. Due to the mechanisms for structural correspondences, the comparisons between the target stimulus and the memorized ones serve as justifications for creation of hypotheses for correspondence between the target stimulus and various ratings. The justifications for a certain hypothesis support it, whereas the inconsistent hypotheses inhibit each other. As a result of the relaxation of this constraint-satisfaction network, one of the hypotheses wins against its competitors and the respective rating is interpreted as the response of the model.

The JUDGEMAP model successfully simulated many of the phenomena in human judgment, found in psychological experiments. In particular, it captures the range and the frequency effects, i.e., the tendency to use all available ratings and to use them

approximately equal times [24]. The model also simulates the sequential assimilation effect, and replicates these results with two-dimensional stimuli [24]. In addition, JUDGEMAP was also used for modeling choice between alternatives and successfully replicates some psychologically validated phenomena [23].

3 JUDGEMAP’s Prediction and Its Testing

When people judge skewed set of stimuli, they tend to shift their ratings in the direction, opposite to the skew. This is the well-known frequency effect [21, 22]. With other words, suppose that people should judge the lengths of lines, but the short lines dominate. In this case people would overestimate all lengths, thus tending to use all ratings almost equal times.

JUDGEMAP successfully simulated this result because of the pressure for one-to-one mapping (Fig.1). If there are numerous short lines, it is more probably the short lines to dominate in the comparison set. Thus, the small ratings would be more frequently used. Now, suppose that there are competing hypotheses for judging a certain target stimulus with the ratings 3 and 4. Keeping everything else equal, suppose that there are more lines in the comparison set that were already judged with 3 than with 4. Then the target stimulus would be judged with 4, because of the inhibitory links between hypotheses that connect one and the same rating with different stimuli. Note, however, that this mechanism is inherited from the AMBR model for analogy-making and is not specifically designed for simulating the frequency effect. As a result, the JUDGEMAP model was able to replicate the frequency effect, when the judged stimulus set is a skewed one.

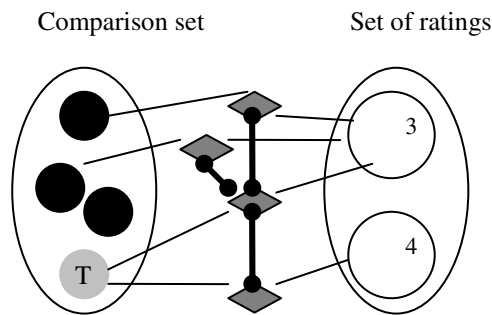


Fig. 1. The mechanism for one-to-one mapping, inherited from the AMBR model causes the hypotheses for the more often used ratings to receive more inhibition and hence, if a positively skewed set is presented, it is a tendency the stimuli to be overestimated

Suppose, however, that the whole set of stimuli is uniformly distributed, but it can be separated into two skewed sets according to a dimension that is irrelevant to the task. For example, let people should judge the lengths of uniformly distributed lines, but let the lines differ in their color. Half of the lines are positively skewed and are green, whereas the other half are negatively skewed and are red (Fig. 2).

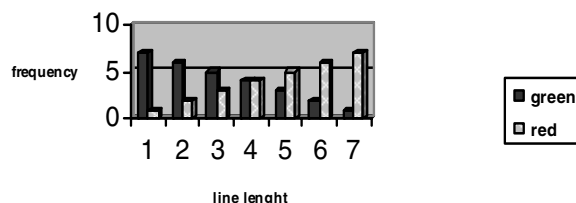


Fig. 2. An example of uniformly distributed set of lines, where each line length is presented for judgment equal number of times. The frequency of the line presentation, however, depends on the line's color. Short lines were presented more often in green, depicted with black lines on the figure, while long lines – more often in red, depicted with the gray textured bars on the figure.

According to the JUDGEMAP model when a green line is to be judged, the activation would spread to the concept 'green' and then back to some of the green instances. Thus, there would be more green lines in the comparison set. However, because the subset of the green lines is positively skewed, it is more probable the short lines to dominate in the comparison set. Thus, because of the frequency principle, the target line would be overestimated. On the contrary, if a red line should be judged, it is more probably the long red lines to dominate the comparison set, and hence, the target red line would be underestimated.

Thus, the JUDGEMAP's prediction is that the green lines would be judged with higher ratings that the red ones with exactly the same length. This particular prediction of the model was successfully simulated and replicated with a psychological experiment [14]. Participants and the JUDGEMAP rated the length of lines on a 7-point scale. The shorter lines were presented more often in green color, while the longer lines were presented more often in red color, thus forming respectively, positively and negatively skewed distributions of lines.

Surprisingly enough we obtained a similar in size difference in judgments of the target green and red lines from people (0.046) and from the model's simulation (0.053). Although this difference was small in size it was significant for both the simulation and the psychological experiment. Thus, the prediction of the JUDGEMAP model was considered to be experimentally confirmed.

The same effect of the irrelevant color on judgment of line length was robustly replicated in several subsequent experiments. We found the same small though significant effect of color on judgment of line length, when the lines were presented at random positions on the screen [6] and also for a very short time [7]. Moreover, in all subsequent experiments the color was counterbalanced across the groups in order to be sure that the effect of irrelevant line color is not due to a specific perceptual length illusion of the green and red colors used in all of our experiments. We didn't find any evidence for such perceptual distortions in our experiments.

In addition, the same effect of irrelevant-to-the-task dimension was found in judgment of abstract stimuli [7, 8]. We used the same design but changed the stimuli from simple green and red lines to much more abstract green and red digits, standing for a particular target characteristic. For example, in one of the experiments participants were asked to judge the age of a hypothetical man on the bases of a number from 10 to 75 that appeared on the screen [7]. Each number represented the

age of a person. Participants were instructed to rate, on a subjective 7-point scale, how old a man of such age is. The numbers were green or red and were skewed in such a way that digits with a particular color formed positive or negative stimulus distribution. As a result, participants judged the same number with a higher rank (i.e. as standing for the age of an older man) if the number shared the color of the positively skewed stimuli than if the digit shared the color of the negatively skewed ones. The same effect was found in price judgments [8].

In sum, the prediction of the JUDGMAP model was repeatedly tested and confirmed within a series of experiments. In all of them we obtain the same effect of irrelevant dimension in judgment on a subjective scale, i.e. stimuli ratings were pushed away from the context of irrelevant-to-the-task information.

4 The Importance of the Type of Scale

This result was quite successful in a way, since the same contrast effect of the context of irrelevant information was repeatedly demonstrated within a series of experiment with the same design. Unfortunately, Goldstone [5] reported an effect in opposite direction (assimilation) of an irrelevant stimulus dimension on judgment of object's color in an experiment with comparable design. Participants in Goldstone's experiment were asked to reproduce the color of the object on the screen and were influenced by the irrelevant-to-the-task shape of the objects. Basically, their color judgments were assimilated toward the prototype of the category to which the objects belong, depending on their shape. For example, if the object's shape belongs to the category of more reddish objects, the reproduced color was more reddish than the reproduced color of an identically colored object that belongs to a different shape, and hence, color category. Goldstone [4, 5] assumes that irrelevant information influences the judgment process relatively early in information processing and discusses the possibility for this effect to be a form of perceptual learning phenomena. Goldstone [4] argues that contextual manipulation of the irrelevant stimulus dimension may cause on-line detectors build up, responsible for the effect interest.

There are, however, several crucial differences in the experimental designs used by Goldstone [5] and in our case [7, 8, 14]. First, Goldstone manipulated only the range of the stimulus distribution, while we usually manipulate the frequency of the stimulus distribution with respect to irrelevant-to-the-task stimulus dimension. This difference, however, could not reverse the contextual effect of interest, since both stimulus range and stimulus frequency result in contrast from the context of stimulus set [21, 22]. Such contrast from the range and/or the frequency of the stimulus set was reported many times in experiments, manipulating the range and/or the frequency of the relevant stimulus dimension [21, 22]. Moreover, there is enough empirical evidence, showing that manipulation of the range of the stimulus set with respect to irrelevant stimulus dimension also results in contrast effect [1, 17, 18, 19, 20, 26]. Hence, manipulation of the range rather than the frequency of the stimuli do not seem to be able to reverse the effect of irrelevant stimulus dimension. The second crucial distinction in the experimental designs consists in the type of scale (objective vs. subjective scale), used by Goldstone [5] and in our case [7, 8, 14] This second difference between the designs should be considered much more carefully, since there

are researchers, arguing that subjective and objective scales may cause opposite contextual effects. Objective scales are considered to be more likely to produce assimilation effects than subjective scales [3, 5, 15, 16, 27, 28]. Although this claim is based on experiments that manipulate the context of the stimulus set with respect to a stimulus relevant (i.e., judged) dimension and there is no theoretical explicit explanation of the way the scale may influence judgment in such a crucial manner, it seems important to test the possible influence of the scale on the direction of contextual shift.

4.1 Psychological Experiment Judgment on an Objective Scale

We decided to test how the scale influences the effect of irrelevant-to-the-task dimension on judgment of line length. In this manner we hope to be able to isolate only the effect of the objective scale judgment, since we have enough data on judgment of line length on a 7-point scale. That is why, in this experiment we asked participants to judge line length in millimeters.

Method

Design. The within subject independent variable were color (varying at 2 levels). The group counterbalanced the experimental design so that the positively and the negatively skewed stimuli to be presented were either in green or in red. The dependent variable was the mean rating of line lengths in millimeters.

Stimuli. 14 color lines that vary from 180 pixels to 505 pixels with an increment of 25 pixels were presented 8 times each forming a basic set of 112 trials. Each line was presented either in red or in green. The frequency distribution of green lines in the first experimental group was positively skewed, while of the red lines – negatively skewed. In the second experimental group the presentation of lines was just on the opposite, i.e., red lines formed a positively skewed distribution and green lines formed a negatively skewed one. The frequency of the positively and negatively skewed lines is presented in Table 1.

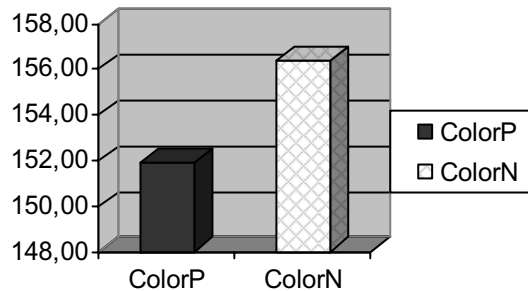
Procedure. Each line was presented horizontally on a gray background in a random position on the screen. The procedure for randomization of the position of each line on an iMac screen was programmed in PsyScope [2]. The participants were instructed to rate each line in millimeters. The experimenter writes down participant's answer and changes the slight manually. The experiment was conducted in sound-proved booths and lasted around 15 minutes for each participant.

Participants. 39 students (24 female and 15 male) from New Bulgarian University participated in the experiment. Participants' age varied between 19 and 36 years. All participants were paid 0.5 euro for taking part in the experiment. There were 19 students in group 1 and 20 students in group 2.

Results and Discussion. The data was averaged by item (14 lengths). Each participant had 28 mean judgments (14 lines*2 colors). The color was analyzed as a within-subject factor, while the group was a between-subject factor. As in all of our previous experiments, the Repeated Measurement Analyses showed a non-significant main effect of the group: $F(1, 37) = 1.166, p=0.287$ which means that it does not matter whether the red or the green color is positively skewed. Thus, the results from the two

Table 1. Frequency and color of the lines for a block of 112 trials, where lines with *color P* were positively skewed and lines with *color N* were negatively skewed

Lines	<i>Length in pixels</i>	Number of the lines with <i>color P</i> (Positively skewed distribution)	Number of the lines with <i>color N</i> (Negatively skewed distribution)
1;2	180;205	7	1
3;4	230;255	6	2
5;6	280;305	5	3
7;8	330;355	4	4
9;10	380;405	3	5
11;12	430;455	2	6
13;14	480;505	1	7

**Fig. 3.** Mean ratings of the *lines* for each color in millimeters. The black bar stands for ratings of the positively skewed lines, while gray-textured bar – for negatively skewed lines.

groups were accumulated and we use *color P* to indicate a positively skewed distribution and *color N* to indicate a negatively skewed distribution in all further analyses.

The main effect of the irrelevant dimension (*color P* vs. *color N*) on rating of the 14 lines was significant, as estimated with the Repeated Measurement Analysis: $F(1, 37) = 7.569$, $p = 0.009$, the effect size (ES) = 0.170. The difference between the mean judgment of positively skewed lines (151.906) and the mean judgment of negatively skewed lines (156.396) was 4.490 mm. Negatively skewed lines were rated higher than positively skewed lines despite the fact that they were equal in length (Fig. 3).

In conclusion, obviously, the scale reverses the effect of irrelevant dimension. The context of irrelevant-to-the-task dimension assimilated the estimated line lengths in millimeters in contrast to the effect of the irrelevant dimension in judgment of line length on a 7-point scale. This result could hardly be described by the JUDGE MAP model, since the model presupposes that judgment is a process of mapping and hence, assumes that the scale is predefined.

If we assume, however, that there is a qualitative difference between the objective and subjective scales, some empirical results cannot be explained. In particular, Wedell, Parducci & Geiselman [29] found that when people judge on a 100-point but subjective scale, the contrastive frequency effect disappeared.

Thus, an alternative possibility is to think about the objective scales as subjective ones, but having much more ratings. Moreover, in the case of length judgments of lines shown on a computer screen, we have an implicit predefined limit of the possible lengths in millimeters, since the screen has a constant width. Thus, a speculative prediction of the JUDGEMAP model is that if the scale consists of too many ratings, because of the limitations of spreading activation, only small part of the scale would be active in any particular moment. Hence, if a certain stimulus should be judged, and if the stimuli with high magnitudes dominate in the comparison set, then only the high ratings of the scale would be active. As a result, the well-simulated contrast effect would appear, but only relative to the active part of the scale. Considered with respect to the whole scale, however, the effect can be exactly the opposite one, namely assimilation.

4.2 Simulation with JUDGEMAP Judgment on a Hundred-Point Scale

This simulation tested the specific prediction of the JUDGEMAP model that the irrelevant dimension will influence the judgment on a hundred point scale. A set of 112 lines was designed. Each line was represented with a coalition of three DUAL-agents – one for the line itself, one for its length, and one for its color. The whole set consists of seven groups of lines. There were eight lines with an equal length in each group. Thus, in the first group there were eight lines with a length 100; in the second one – eight lines with a length 200, etc. In the last group there were eight lines with a length 1400. According to the line's color, however, the distribution was more complex. In the first 2 groups there were seven green and one red line in each group; in the second 2 groups – six red and two green lines in each group; etc. In the last two groups there were one green and seven red lines in each group (see Table 1).

Thus, the overall set was uniformly distributed, but it consists of two skewed subsets according to the line's color, which is completely irrelevant to the task property. In summary we used the same set as in the subjective scale simulations.

The scale was represented with a set of hundred ratings, interconnected each one with its neighbors. There were also associative links between the neighbor ratings that can be divided to 10, i.e., between 10th and 20th, between 20th and 30th, etc. All 100 ratings were instances of one concept of scale itself. There were, however, only few opposite links from the concept to the ratings of 10, 50, and 90, thus simulating several 'favorite' numbers. The same method for activating the scale was used in all JUDGEMAP simulations, since without the opposite links from the scale concept to some specific scale values no ratings could be activated at all.

The task was represented with a correspondence between the concepts of longer line and of higher rating. The whole set of 112 lines was judged 30 times, each time all of the lines were judged sequentially in a random order.

Results and Discussion. The data was analyzed in the same way as in the psychological experiment. The difference between the mean judgment of positively skewed lines (48.907) and the mean judgment of negatively skewed lines (50.359) was 1.452 on 100-hundred scale, but the main effect of the color on rating of the 14 lines was non-significant, as estimated with the Repeated Measurement Analysis: $F(1, 29) = 1.686, p = 0.204$.

The significant assimilative effect to the direction of the skew, received in the psychological experiment, was not replicated. The contrast effect, however, simulated with a seven-point scale, also disappeared. This result is in line with the experiment of Wedell Parducci & Geiselman [29], which demonstrates that, the role of the stimulus skew (i.e., the pressure of frequency principle) decreases when the number of the available rating increases.

Obviously, however, there was an extremely high bias of the opposite links from the scale concept to the ratings 10, 50, and 90. It can be noticed in Figure 4 that JUDGEMAP prefers these ratings and their neighbors much more often than the other ratings. In other words, the main problem of the simulation was that JUDGEMAP could not successfully support the appropriate parts of the scale active enough. The main difference between the results of the simulations with 7-point and 100-point scales was that in the first case all ratings were active enough, whereas in the second case they are not. When judging on a 7-point scale most of the ratings participate in the competition and the small difference in their activation level causes small effects like the sequential assimilation. When judging on a 100-point scale, however, the activation was not enough to support all ratings (and should not be!). The main source of activation in this case was the opposite links, thus extremely increasing the bias that they press.

Thus, one possibility for JUDGEMAP to simulate assimilative effect when judging on an objective scale, using the same mechanisms as when judging on a subjective one, is to find another representation of the scale. For example, we may assign randomly the opposite connections from the scale concept to the scale ratings, avoiding in such a way the observed high bias in the 100-point scale simulation. In this manner we may overcome the JUDGEMAP's preference to use mainly the rating around the scale values (i.e., rating 10, 50 and 100 in the reported simulation) that receive activation from the 100-scale concept. However, even random, the opposite links would press too high bias.

Another possibility for JUDGEMAP model to simulate the observed assimilation in judgment on objective scale is to think that when people judge on an objective scale, they do not use any predefined "chain"¹ of possible answers, but rather generate the ratings. Thus, if each time JUDGEMAP generates a new rating, there would not be a need for inhibitory links between hypotheses that connect one and the same rating with different stimuli. Thus, the pressure for one-to-one mapping, which is the main source of the contrast effects in judgment on a seven-point scale, would be eliminated. How exactly to generate these new ratings, however, is still an open question for us. Although the idea that people use different strategies when judging on a subjective or on an objective scale seems reasonable the main idea of the JUDGEMAP model is to demonstrate how different cognitive phenomena can be explained as an emergent result of a limited number of basic mechanisms (more precisely, the mechanisms for analogical mapping).

¹ In all JUDGEMAP simulations the scale was represented as a "chain" of interconnected nodes, where each node, standing for a particular scale value is connected with associative links to its neighbors. For example, rating "3" is connected with associative links to rating "2" and respectively, to rating "4", rating "4" is connected to rating "3" and rating "5" etc.

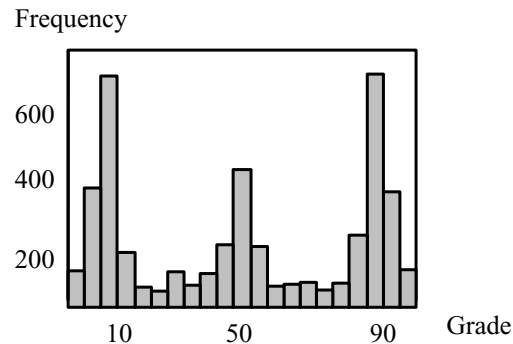


Fig. 4. The frequency distribution of the ratings, given by JUDGMAP

They are, however, other alternative possibilities. One of them is to think about additional mappings between the target stimulus and the elements of the comparison set. During such type of mappings, the ratings of the memorized stimuli would be transferred with appropriate modifications. Suppose that the target stimulus that a person has to judge reminds him/her of another stimulus that he/she has already judged or just remembers its absolute value. Then, a person may just transfer the rating of the known stimulus to the similar new one. Let's think for a while about our green positively skewed and red negatively skewed lines, although the possibility that a particular line may remind you about another one does not sound quite realistic. Suppose, however, that you judge the length in millimeters of a green target line. The target line may remind you about another green line of approximately the same length that you have already estimated as being 200 millimeters. Then you may just transfer the rating of the previously judged "similar" line to the new one, probably with slight modifications depending on the degree of similarity between the length of the old and the new line. Moreover, since the green lines were positively skewed, there is higher probability that the target green line may remind you about a short green line. Thus, the ratings of the green lines would be biased toward the lower part of the scale, while the rating of the red lines – toward the upper scale values. In this manner, we may increase the probability for assimilation toward the context of irrelevant dimension, but only for judgment on a scale that has more values than the system may simultaneously keep in the WM. Judgment on a 7-point scale would not be substantially affected by this transfer mechanism, since the pressure for one-to-one mapping would be still strong enough to overcome the possible assimilation introduced by this transfer mechanism. An attempt to implement transfer mechanisms in the AMBR model was already made [25], and these mechanisms can be used in the judgment task, namely to explain why the scale was able to reverse the effect of context in the psychological experiments.

4 Conclusions

The idea that people perform judgments relative to a set of comparable elements is not new or surprising. According to the JUDGEMAP model, however, people use

even irrelevant dimensions for the construction of the comparison set. Furthermore, our main efforts in building the JUDGEMAP are not just to replicate some specific psychological phenomena, but to explain all these various phenomena in terms of the same basic mechanisms (such as spreading activation and structure mapping), which are also basic for a range of other cognitive processes like perception, analogy, memory.

Objective scales pose an interesting challenge for the JUDGEMAP model, although the model has never pretended to account for contextual effects in judgments on this kind of scales. JUDGEMAP uses predefined discrete subjective scales and successfully simulates a range of well-known context effects on judgment on subjective scale. In addition, JUDGEMAP's prediction about the role of irrelevant stimulus information in judgment was systematically tested and confirmed within a series of psychological experiments.

The idea, however, that the type of scale reverses the context effect because of different underlying mechanisms [5, 15, 16, 27, 28] challenges in an important way some basic assumptions of JUDGEMAP. The model assumes and successfully demonstrates that one and the same basic mechanisms may be used for judgment, analogy-making, choice-making [23] and even partially for perception [25]. Thus, it seems reasonable to suggest that the same mechanisms may also account for judgment on objective scale.

Our first attempt to model assimilation toward the context of irrelevant dimension, without any changes in the model, failed. In the previous section, we outlined several possible opportunities to extend the JUDGEMAP model in order it to be also able to explain judgment on objective scale without any essential changes in the model's mechanisms and assumptions. In summary the main idea would be to add a mechanism for rating transfer (based on direct comparison – analogical mapping between an already evaluated stimulus and the target one) and to solve the problem of realistic representation of large scales. The analogy transfer will be a new mechanism for hypothesis building in both the subjective and objective scales. It can potentially contribute to assimilation effects in both cases. The difference would be in the role of inhibitory links in both cases. In the case of subjective scales there is a pressure to use the whole scale and thus the role of competition between the hypotheses should be higher (and leads to contrast), while in the case of objective scales no such competition is required. The good news from the current simulation is that with large scales the contrast effect disappears. Thus if the two mechanisms (competition between hypotheses leading to contrast and hypothesis formation based on the ratings of similar stimuli leading to assimilation) are co-existing in both subjective and objective scales and working against each other, in large objective scales only the assimilation effect will persist and this would explain the overall assimilation effect. This has yet to be tested with a further development of the JUDGEMAP model.

Acknowledgments. This work is supported by the Project ANALOGY: Humans – the Analogy-Making Species, financed by the FP6 NEST Programme of the European Commission.(Contr. No 029088).

References

1. Arieh, Marks, L.: Context effects in visual length perception. Role of ocular, retinal, and spatial location. *Perception & Psychophysics* 64(3), 478–492 (2002)
2. Cohen, MacWhinney, Flatt, Provost: PsyScope: An Interactive Graphic System for designing and controlling experiments in Psychology Laboratory Using Macintosh Computers. *Behavior Research, Methods, Instruments and Computers* 25, 257–271 (1993)
3. Foti, Hauenstein: Processing Demands and the Effects of Prior Impression on Subsequent Judgments: Clarifying the Assimilation/Contrast Debate. *Organizational Behavior and Human Decision Processes* 56, 167–189 (1993)
4. Goldstone, R.: Perceptual Learning. *Annual Review of Psychology* 49, 585–612 (1998)
5. Goldstone, R.: Effects of Categorization on Color Perception. *Psychological Science* 6(5), 298–304 (1995)
6. Hristova, P., Kokinov, B.: Perceptual Learning vs. Context-Sensitive Retrieval: Why do people judge green lines to be shorter/longer than red lines of the same length? Do they perceive them differently or do they retrieve a biased set of alternatives in their comparison set? In: *Proceedings of the Second European Cognitive Science Conference* (in press)
7. Hristova, P., Kokinov, B.: A Common Mechanism Is Possibly Underlying the Shift in Perceptual and Conceptual Judgment Produced by Irrelevant Information. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Erlbaum, Hillsdale, NJ (2006)
8. Hristova, P., Petkov, G., Kokinov, B.: Influence of Irrelevant Information on Price Judgment. In: *Proceedings of the International Conference on Cognitive Economics*, NBU Press (2005)
9. Kokinov, B.: Associative Memory-Based Reasoning: How to Represent and Retrieve Cases. In: O’Shea, T., Sgurev, V. (eds.) *Artificial Intelligence III: Methodology, Systems, Applications*, Elsevier, Amsterdam (1988)
10. Kokinov, B.: A Hybrid Model of Reasoning by Analogy. In: Holyoak, K., Barnden, J. (eds.) *Advances in Connectionist and Mental Computation Theory*. Analogical Connections, Ablex, Norwood (1994a)
11. Kokinov, B.: The DUAL Cognitive Architecture. A Hybrid Multi-Agent Approach. In: Cohn, A. (ed.) *Proceedings of the Eleventh European Conference of Artificial Intelligence*, John Wiley & Sons, Ltd. London (1994b)
12. Kokinov, B.: The Context-Sensitive Architecture DUAL. In: *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. LEA, Hillsdale (1994c)
13. Kokinov, B., Petrov, A.: 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*, MIT Press, Cambridge, MA (2001)
14. Kokinov, B., Hristova, P., Petkov, G.: 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, Mahwah (2004)
15. Lockhead, King: A Memory Model of Sequential Effects in Scaling Tasks. *Journal of Experimental Psychology: Human Perception and Performance* 3, 461–473 (1983)
16. Manis, Nelson, Shedler: Stereotypes and Social Judgment: Extremity, Assimilation and Contrast. *Journal of Personality and Social Psychology* 55, 28–36 (1988)
17. Marks, L.: The slippery context effect in psychophysics: Intensive, extensive, and qualitative continua. *Perception and Psychophysics* 51, 187–198 (1992)
18. Marks, L.: Recalibrating the auditory system: The perception of loudness. *Journal of Experimental Psychology: Human Perception and Performance* 20, 382–396 (1994)

19. Marks, L., Armstrong: Visual and haptic representations of space. In: Inui, McClelland (eds.) *Attention and Performance XVI: Information Integration in perception and communication*, MIT Press, Cambridge (1996)
20. Marks, L., Warner: Slippery context effect and critical bands. *Journal of Experimental Psychology: Human Perception and Performance* 17, 986–996 (1991)
21. Parducci, A.: Category Judgment: A Range-Frequency model. *Psychological Review* 72(6), 407–418 (1965)
22. Parducci, A.: Contextual Effects: A Range-Frequency Analysis, *Handbook of Perception*, vol. 2. Academic Press, NY (1974)
23. Petkov, G.: Modeling Analogy-Making, Judgment, and Choice with Same Basic Mechanisms. In: Fum, D., Missier, F., Stocco, A., Goliardiche, E. (eds.) *Proceedings of the Seventh International Conference on Cognitive Modeling*, pp. 220–225 (2006)
24. Petkov, G.: Judgment as Mapping (JUDGEMAP2). In: *Proceedings of the Doctoral Consortium at the Fifth International and Interdisciplinary Conference on Modeling and Using Context*. Technical Report LIP 2005/007 of the Laboratoire d'Informatique de Paris 6, pp. 95–104 (2005), URL <http://www.lip6.fr>
25. Petkov, G., Kiryazov, K., Grinberg, M., Kokinov, B.: Modeling Top-Down Perception and Analogical Transfer with Single Anticipatory Mechanism. In: *Proceedings of the Second European Cognitive Science Conference* (in press)
26. Rankin, Marks, L.: Differential context effects in chemosensation: Role of perceptual similarity and neural communality. *Chemical Senses* 25, 747–759 (2000)
27. Ward: Critical Bands and Mixed-Frequency Scaling: Sequential Dependencies, Equal Loudness Contours, and Power Function Exponents. *Perception and Psychophysics* 47, 551–562 (1990)
28. Weddel: Methods for Determining the Locus of Context Effects in Judgment. In: Caverni, Fabre, Gonzales (eds.) *Cognitive Biases*, New York, Elsevier Science Publishers, Amsterdam (1990)
29. Wedell, Parducci, Geiselman: A Formal Analysis of Ratings of Physical Attractiveness: Successive Contrast and Simultaneous Assimilation. *Journal of Experimental Social Psychology* 23, 230–249 (1987)