NEW BULGARIAN UNIVERSITY

CENTRAL AND EAST EUROPEAN
CENTRE FOR COGNITIVE SCIENCE

Does Irrelevant Information Influence Judgment?

Penka Hristova

A dissertation submitted in partial fulfillment of the requirements for the degree of Ph.D. in Cognitive Science

Supervisor:
Assoc. Prof. Boicho Kokinov

SOFIA • 2005

Abstract

When people judge how tall a person is it is rational to assume that they ignore all irrelevant information about the color of his/her hair and eyes, the fact that he/she wears glasses, or the profession of that person, and focus on the relevant dimension "height" only.

This dissertation is based on a model of judgment, JUDGEMAP-1, that proposes certain mechanisms that underlie judgment. This model assumes that in order for judgment to happen people activate a number of exemplars from the same category in Working Memory and this active set is called the comparison set. The target stimulus is included in this comparison set and altogether they are mapped on the scale using the same mechanisms that underlie analogical mapping. Thus, key for the rating, which the target stimulus will obtain, is what the other members of the comparison set are (if smaller exemplars happened to be involved in the comparison, the target will most probably receive a higher rating as it will stand out in the comparison set). The basic mechanism responsible for retrieving exemplars in WM in JUDGEMAP-1 is *spreading activation* and thus any kind of shared features will make the retrieval of a particular exemplar more probable. This means that the color of the hair and eyes, the profession of the person may actually play a role, since the comparison set will consist of people with similar characteristics. This theoretical prediction was backed up by a particular simulation experiment with JUDGEMAP-1 and thus formed the main prediction to be tested within this dissertation.

A series of experiments has been performed that test this prediction under various circumstances. Even though the first experiment confirmed the prediction, , the size of the effect was very small (although significant), so there were serious doubts whether this result can be replicated. Thus, the series of experiments replicates the initial experiment varying some parameters in the hope to enlarge the size of the effect. The experiments varied the dimensionality of the stimuli, the presentation manner and presentation time of the stimuli, and the type of stimuli – from simple lines to rabbits to abstract numbers. Even though we could not increase the size of the effect, the effect was replicated quite robustly in almost all experiments.

Acknowledgments

I would like to express my special gratitude to my supervisor Boicho Kokinov who has guided patiently my work during the recent 2 years and has strongly influenced the way I look at cognition. I heartfelt acknowledge his intellectual inspiration and constant moral support along the way.

I would also like to thank my husband and collaborator Georgi Petkov for providing years of general support and encouragement. He has always asked me the most difficult and painful questions that turned to be, usually, the most important ones in my work.

Two professors inspired and influenced me in significant ways. Prof. Stefan Mateev pushed to the limit my naïve understanding on data analysis and interpretation. Alexander Petrov has provided insightful and constructive ideas on my research. In many ways, they paved the road for this dissertation and determined the way it looks like now.

Thanks to Armina Janjan for her patient assistance on my writing skills and for providing essential software and even hardware support.

This research has been influenced in various ways from other past and present members of AMBR research group and Cognitive Science Department, including Evgenia Hristova, Neda Zareva, Boika Bratanova, Velina Balkanska, Veselina Feldman, Krasimira Janeva.

Finally and most of all, I would like to thank my family for their love and support over the years and especially to my lovely little daughter for her patience and understanding.

Table of Content

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.

The presented research tests a particular prediction of the JUDGEMAP-1 (Judgment as Mapping) Model, namely that irrelevant information influences judgment. This prediction is based on the spreading activation mechanism underlying the retrieval of information in memory. In JUDGEMAP-1 spreading activation brings from Long Term Memory (LTM) into Working Memory (WM) information about stimuli that are similar to the target stimulus both in terms of relevant and irrelevant features. For example, if a person has to judge the height of a particular target person, his/her judgment will be affected by the extra knowledge about the target. If the target person is a teacher, the stimuli, which could be elicited in WM during judgment, could be predominantly teachers. Moreover, if the target person is blond, it is quite likely that the exemplars elicited in WM would also be predominantly of other blond people. This prediction was tested experimentally with simple stimuli, e.g. lines with different color and confirmed in an empirical study conducted by Kokinov, Hristova, and Petkov (2004). The series of experiments reported in this dissertation focus extensively on the boundary conditions of this effect. It investigates the size of the effect and its systematization varying the dimensionality of the stimuli, the presentation conditions, and the abstractness of the stimuli.

The thesis is organized in the following manner. Chapter 1 reviews empirical data and theories of contextual effects on judgment of both simple and complex stimuli. Chapter 2 presents the JUDGEMAP-1 Model, the cognitive architecture DUAL that JUDGEMAP-1 is based on and the mechanisms of judgment proposed by the model. Chapter 3 lays out the details of the formation of the so called “comparison set” within the framework of the JUDGEMAP-1 Model. It specifies the role of spreading activation in the process of judgment: namely, how the activation may determine the content of the comparison set. Chapter 4 describes two control experiments that explore the highest possible shift in judgment due to range (*R*) and frequency (*Fr*) of stimulus distribution in judgment of line length. Chapter 5 presents an attempt to increase the impact of the irrelevant information by varying both the *R* and the *Fr* stimulus distribution with respect to the irrelevant to the task dimension on line length judgments. Chapter 6 reports one experiment that tries to increase the influence of the irrelevant to the task dimension by limited stimulus presentation in judgment of line length. Chapter 7 presents two experiments that explore the effect of several correlated irrelevant dimensions. Chapter 8 describes one experiment that investigates the possibility of shift in judgments due to the irrelevant information with abstract stimuli. Finally, Chapter 9 presents the conclusions from the experimental findings and suggests ideas for future work.

CHAPTER 1
Empirical Evidence for Contrast and Assimilation Effects

The literature on the contextual effects in judgment reports two main shifts in the rating depending on the presented context– contrast and assimilation effects. They concern situations in which both the contextual and the target stimulus could be measured on the same scale. ***Contrast*** refers to the displacement of judgments away from the values of the contextual stimuli, while ***assimilation*** refers to the displacement of judgments toward the values of the contextual stimuli.

1.1. Contrast effect

In this section, experiments resulting in contrast effect are conditionally separated by subsections according to the procedure used.

1.1.1. Anchor placed outside stimulus series

Several experiments used the so-called "anchor" procedure to study contextual effects in judgment. Generally, "anchor" means an explicit standard for comparison (a reference point) presented during judgment. For example, Sherif, Taub and Hovland (1958) demonstrated contrast effect in judgment of weights. A within-subject experimental design was used - each participant rated the same weights in two conditions: with or without an anchor, in addition, values of anchor were manipulated in the "anchor" condition. In the first session, participants were asked to lift weights one by one and to judge how heavy they feel them on a 6-point scale. Then, in a second part of the experiment, subjects lifted the same weights but coupled with a standard for comparison. Each weight was presented in pair with an anchor weight which subjects were explicitly told to call 6. The judged weights were between 55 and 141 gm. When the anchor (the contextual stimulus) was heavier in weight (between 168gm and 347gm) than the objects in the regular series, the ratings of the target stimulus were reduced, i.e. *contrast* effect was observed. The judgment of the stimulus paired with an anchor was lower than the judgment of the same stimulus when presented alone.

Sherif et al. (1958) conducted an additional experiment in order to find out whether the same effect would be observed if the anchor is lighter than the stimulus series. The weights in the stimulus series were between 75 and 141gm. Anchoring weights (i.e., 67, 43 and 35 gm) were lighter than the weights in the stimulus set and were called to be equal to 1 on a 6-point scale. The design and instruction used were the same as in the previous experiment. The results were also similar, that is, contrast effect in judgment was observed, i.e. the same stimuli were rated higher when presented with a lighter anchor than without an anchor.

Another experiment, which obtained contrast effect by the means of an anchor placed outside stimulus set, was conducted by Sarris and Parducci (1978). They compared the effect of single and multiple anchoring upon category rating of square sizes. In the single anchoring condition, the same square was presented on every anchoring trial. In the multiple anchoring conditions, squares with different sizes were presented on the anchoring trials. Participants were asked to judge the size of each square on every other trial. The anchoring stimulus, which was presented before each target stimulus, was not judged. The judgments were placed on a 9-point scale. The results show that in both single and multiple anchoring conditions contrast declines with the increasing the distance between the anchor and the target stimulus.

In conclusion, reviewed experiments with anchors point to several important observations about context effects in judgments:

- First, anchors placed outside the regular series redefine the end points of the stimulus set and cause contrast effect.
- Second, distance of the anchor to the regular series can be important for the degree of the contextual shift in judgment;
- Third, the procedure for stimulus presentation was not found to differentiate the results. Although, the first experiment (Sherif et al., 1958) used simultaneous presentation of the anchor and the target stimuli while the second one (Sarris and Parducci, 1978) – a sequential presentation, results were comparable. Both experiments found a significant contrast effect due to an anchor magnitude which was out of the range of stimulus magnitudes.

1.1.4. Primed context

Many experiment in the field of judgment use category priming to investigate contextual effects. Few of them would be presented in order to give an idea about various experimental situations used for studying contrast effect in judgment.

- *Judgment in the presence of an extreme contextual exemplar*

Herr, Sherman, and Fazio (1982) asked participants to rate the size of a target animal on 11-points scale during a “color perception” experiment. The target stimuli were always moderate in size animals - Wolf, Sheep, Pig, Goat, Jabo and Lemphor. The first 4 animals were examples of real animals, while the last 2 – of fictitious ones. This difference between the target stimuli (real/fictitious) was treated as similar to the distinction between the ambiguous and unambiguous stimuli. The unambiguous stimuli were the sizes of the real animals, while the ambiguous ones were the sizes of the fictitious animals. Contextual information was primed in the course of a "color perception" experiment. Context was distinguished into four levels of animal sizes: extremely high, moderately high, moderately low, and extremely low (see Table 1).

Table 1. Examples of the animals used for the priming of a particular size in the experiment of Herr, Sherman, and Fazio.

Primed category	**Mean size** (Pretest on a 10-point scale)	**Examples**
extremely high	mean=9.11	Whale, Elephant, Hippo, Rhinoceros
moderately high	mean=6.81	Antelope, Cow, Lion, Tiger
moderately low	mean=2.93	Porcupine, Gopher, Groundhog, Cat
extremely low	mean=0.36	Snail, Flea, Minnow, Ant

Whenever the *target* was an *unambiguous* stimulus (an exemplar of a real animal) its rating shifted away from the primed size, i.e. a contrast effect appears. When the *target* animals were *ambiguous* ones (fictitious: jabo and lemphor) and the prime animals were extremely high or extremely low exemplars, a contrastive shift in judgments was also observed. This, however, did not happen when participants judged an ambiguous target but the prime was moderately high or moderately low in size animal. The moderate context was found to result in assimilation rather than contrast in judgment of ambiguous target stimuli (for details, see the section about assimilation toward primed context).

The same effect of an extreme prime was obtained in judgment of face attractiveness. For example, Kenrick and Gutierres (1980) found contrast effect in judgment of an average target face when it was presented after a highly attractive face. In the first experiment participants who were watching “Charlie’s Angels" were asked after the end of this TV program to judge the beauty of a moderately attractive female faces on a 7-point scale. The main characters in “Charlie’s Angels" were 3 highly beautiful women. Participants in the control group were also asked to judge the same faces but after watching another TV program. The same face received a lower rating (3.34) from the watchers of "Charlie's Angels" than from the controls (4.00).

The same result was obtained when contextual face appeared during the instruction. Participants were asked to rate a female face of average attractiveness along several bipolar scales. The target scale was beautiful-ugly. Half of the participants saw a picture of a very beautiful female face during the instruction, the others – did not. Those participants who saw the beautiful woman before assessing the beauty of the target face underestimated the test face in comparison with participants who did not see the beautiful face.

Thus, priming of an extreme exemplar could be considered as resulting in contrastive shift in judgment of a moderate target. Similarly to the previous section about contrast due to an anchor, a contextual stimulus, the magnitude of which is out of the stimulus magnitudes seems to cause contrast effect in judgment.

- *Contrast in judgment of a target stimulus belonging to a category associated with an extreme stereotype*

Few experiments focused on judgment depending on an existing stereotype. All of them start with an induction series for establishing a stereotype and a test series for judgment of stereotyped and non-stereotyped stimuli. For example, Manis, Nelson and Shedler (1988) conducted an experiment in order to test how the extremity of the existing stereotypes influences social judgment. During the induction phase participants were asked to judge the pathology of statements coming from Metropolitan or Central Hospital. The statements from one of the hospitals were always normal, while the statements generated from patients of the other hospital were 100% extremely pathological. In this manner an extreme stereotype toward a particular hospital was established - Metropolitan or Central Hospital. During the test session, participants were asked to choose the more pathological case within a pair of moderate in pathology statements. One of the cases was originated from a patient of a Metropolitan Hospital and the other - from the Central Hospital. If the participants had participated in the induction part of the experiment when an extreme stereotype about Metropolitan Hospital had been established, they would choose as more pathological the test statement from the other hospital (Central Hospital). On the contrary, statements from Metropolitan Hospital were chosen when an extreme stereotype toward the Central Hospital had been induced. Because both statements in the test pairs were with moderate pathology, the results were interpreted as showing contrast effect due to an extreme stereotype for a particular Hospital. Contrast, however, was observed only when the stereotype was an extreme one otherwise assimilation appeared (see section for assimilation toward primed context).

This experiment shows once again that the extremity of the context is very important even in judgment of such complex entities like clinical cases. Moreover, context itself was a quite complex stimulus, i.e., a particular stereotype concerning the level of patient's pathology in a hospital.

- *Context excluded from the representation of the target*

The last study in this section concerns also judgment of complex social stimuli like participant's current happiness and life satisfaction. Strack, Schwarz, and Gschneidinger (1985) show contrast effects in judgment of happiness and satisfaction when preceded by the recall of 3 positive or 3 negative past events. When participants were induced to think about 3 negative events from their *remote* past, they judged their current happiness and satisfaction as higher (on 11-point scale) than participants who recalled 3 positive *remote* past events first. In contrast, participants who were first asked to recall 3 positive *recent* events assessed their current happiness and life satisfaction higher (on 11-point scale) then Ss who recalled 3 negative recent events. The authors interpret these results in favor of the inclusion-exclusion theory - contrast appears when the context is excluded from the target's representation.

To summarize, priming of particular contextual information may result in contrast effect if the prime information is an extreme one. Extreme exemplars are able to push away the judgments of unambiguous stimuli and of ambiguous and moderate exemplars of the same category. However, in case of ambiguous stimuli (Herr et al., 1980) assimilation effect could be obtained. It should be noted that it is not clear what is assumed to be an ambiguous stimulus – something having several meanings for the same person or having different meaning for different people (for discussion, see Herr et al., 1980). This vagueness may lead to confounding variables and possibly invalidate the results. Unambiguous stimuli, on the other hand, are more likely to be displaced away from the active contextual information independently of its extremity (Herr et al., 1980).

1.1.3. Contrast depending on a set of stimuli

Many experiments demonstrated that context could be a set of all presented stimuli rather than a single stimulus. This research line was initiated from Parducci in the 60s (Parducci, 1965, 1968,1974). Just few experiments testing the influence of context considered as a set of sequentially presented stimuli would be briefly reviewed here.

- *Range of the stimulus distribution*

Parducci and Perret (1971) demonstrated contrast due to a change in the range of the stimulus set. The results showed that despite the same mean of the stimulus series, judgments of the same square (with the same width, projected on the same screen) differ significantly depending on the range of the stimulus distribution. For example, stimulus distribution that includes squares 1, 3, 5, 9, 13, 17, 20, 22 and 24 has the same mean as the set of squares with sizes 1, 2, 4, 7, 11, 14, 17, 20 and 22. Each number of a square stands for a size, which is 1.16 times bigger then the previous square from in the set of 24 squares. The judgment was made on a 9-point scale. The mean judgments of the common stimuli for the two conditions are presented in Table 2.

Table 2. The mean judgment of the sizes of the common squares depending on the experimental condition.

Stimuli	Range 1 (from 1 to 24)	Range 2 (from 1 to 22)
1	1.09	1.10
17	4.45	5.00
20	5.60	6.47
22	6.80	7.88

The second distribution was restricted with respect to the first one. The ends of the first distribution were 1 and 24, while for the second distribution they were 1 and 22. As if a square number 24 was added in the first set. The effect was reducing the rating of the rest of the squares in the first set, the so-called "contrast effect". Contrast due to variations in the range of the stimulus set was demonstrated with various stimuli with different complexity, i.e. in judgments of physical attractiveness (Wedded, Parducci and Geiselman, 1978), in clinical judgments (Wedell, Parducci and Lane, 1990), judgments of class performance (Mellers and Cooke, 1994).

Thus, the range of the stimulus set may be considered as creating context for judgment of each particular stimulus within this set. The next section will review several studies demonstrating that the frequency of the stimuli within the stimulus set may also matter.

- *Frequency of the stimulus distribution*

Many experiments initiated by Parducci indicate how judgment depends on the frequency of the stimulus presentation within the whole set of stimuli that has to be judged. For example, Parducci and Perret (1971) demonstrated how violations in the frequency of the distribution of several squares with different sizes result in the different ratings of the same square. Consider the situation of having three sets of nine squares with equal ranges, i.e. the smallest and the largest square in these sets are identical. The only difference between these sets is the frequency of square's presentation. In set 1, all nine squares are presented equally often (9 times each) forming a uniform distribution of stimuli. In set 2, small squares were presented more frequently than larger squares, forming a positively skewed set. In set 3, larger squares are presented with higher frequency than smaller ones, forming a negatively skewed set. Participants were asked to rate the size of each square "in comparison with the other squares" on a 6-point scale (p.433, Parducci and Perrett, 1971). Results show a significant contrast effect in judgments of the common stimuli depending on the stimulus distribution. In the positively skewed distribution, ratings of all squares shifted upward, while in the negatively skewed distribution – downward (Table 3).

Table 3. Ratings of the common stimuli in 3 of the experimental conditions in the study of Parducci and Perret.

Stimuli	Normally skewed set	Positively skewed set	Negatively skewed set
1	1.09	1.05	1.00
3	1.62	1.47	
7		2.30	1.60
19		4.45	3.34
22	6.80		4.56
24	8.21	5.83	5.67

Wedell, Parducci and Geiselman (1987) obtained contrast effect due to the frequency of the distribution in judgment of physical attractiveness. Participants were asked to rate the attractiveness of female faces on 5 and 101-point scale. The stimuli were drawn from a set of 118 photographic slides of female faces that were previously rated on a 5-point attractiveness scale. The stimuli were distinguished into seven groups from less attractive to very attractive female faces. Less attractive faces predominate in the positively skewed set, while more faces that were very attractive were included in the negatively skewed set.

Stimuli were presented sequentially and in random order except the ten test stimuli, which appeared always at the same position in both experimental conditions. The effect of the skew of the stimulus distribution was measured trough the ten common stimuli for both skews. Contrast effect of skewing was significant - the same faces were rated higher in the positive condition and lower in the negative one. The effect decreases from the 5-point scale to 101-point scale but was still significant in the latter.

The same context effect due to density distribution of the stimulus set was reported with stimuli with different complexity, e.g. clinical judgments (Wedell et al., 1990), and class performance (Mellers and Cooke, 1994).

- *Contrast due to the frequency in multiattribute judgment*

Cooke and Mellers (1998) demonstrated contrast due to the frequency of the stimuli along one out of 3 target dimensions. Participants were asked to judge an offer for a rent of an apartment described on three dimensions - price, distance from campus and friend's opinion. The distribution over one of the target dimensions (either distance or rent) was skewed while the distribution over the other two were uniform. When the distance was skewed (given as walking time in minutes) and the remote offers predominated, the ranks of the middle stimuli increased, and the opposite, when near apartments predominated in the list, the offers from the middle part of the stimulus distribution were perceived as worse by subjects. Similar results appear in the rent condition (positively and negatively skewed distributions with respect to rent).

Mellers and Birnbaum (1981) tried to differentiate the possible mechanism for contextual effect in judgments of two-dimensional stimuli, which were skewed on both target dimensions: test scores on two exams (Exam1 and Exam2). The scores on both exams were skewed either positively (the low ratings on both exams predominated) or negatively (the high ratings on both exams predominated). Participants were asked to judge the overall performance of each student on a 9-point scale. Each stimulus consisted of two numerical values that depicted the two test scores.

The data for the 40 scores common for both conditions was compared. The authors argue that the distribution of the total scores affects only the transformation from the integrated impression to the overt response. Judges were sensitive to contextual information while combining their own judgments for the two exams into a rating for a student’s overall performance. The effect of visualization of the stimulus skew (i.e., some of the participants were shown a histogram of the exam's distribution) was minimal.

This explanation, however, does not seem appropriate for cross-modality judgments, i.e., judgments of stimuli from two different dimensions. Mellers and Birnbaum (1982) asked participants to rate the "total intensity" of both stimuli (i.e., density of dot-patterns and size of circles) on a scale from 0 (the intensity is very very low) to 90 (the intensity is very very high). The darker the dot-pattern and the larger the circle, the greater the "total intensity" of the stimuli was. Six circles with diameters between 7.6 and 25.4 mm were combined with one of four distributions of dot-patterns (medium range, wide range, positively skewed, negatively skewed). Like in the previous experiment, judgments of 6 common pairs of stimuli were compared across the four experimental conditions. Results were interpreted as being in favor of the hypothesis that context influences judgment on each dimension separately rather than the integrated impression of both dimensions. Participants judged the intensity of each circle (size) and dot pattern (darkness) separately and depending on the distribution of the circle's sizes and dot darkness, respectively.

In conclusion, this review suggests that judgment can be sensitive to properties of the whole set like the range and even the density of the stimulus distribution. The effect of the range seems to be similar to the effect of an anchor and the extreme primed context described in the previous sections, i.e. context changes the range of the stimulus set and causes contrast effect. This effect declines with increasing the distance between the context and the stimulus set. Contrast due to both a range and a frequency violation was repeatedly demonstrated independently of the stimulus complexity and the number of the judged dimensions (e.g., multiattribute judgments). It could be, however, that the size of the observed effect was a result of the explicit instruction to judge each stimulus with respect to the rest, rather than *only* from manipulations of the range and frequency of the stimulus distribution. Unfortunately, this question remains open.

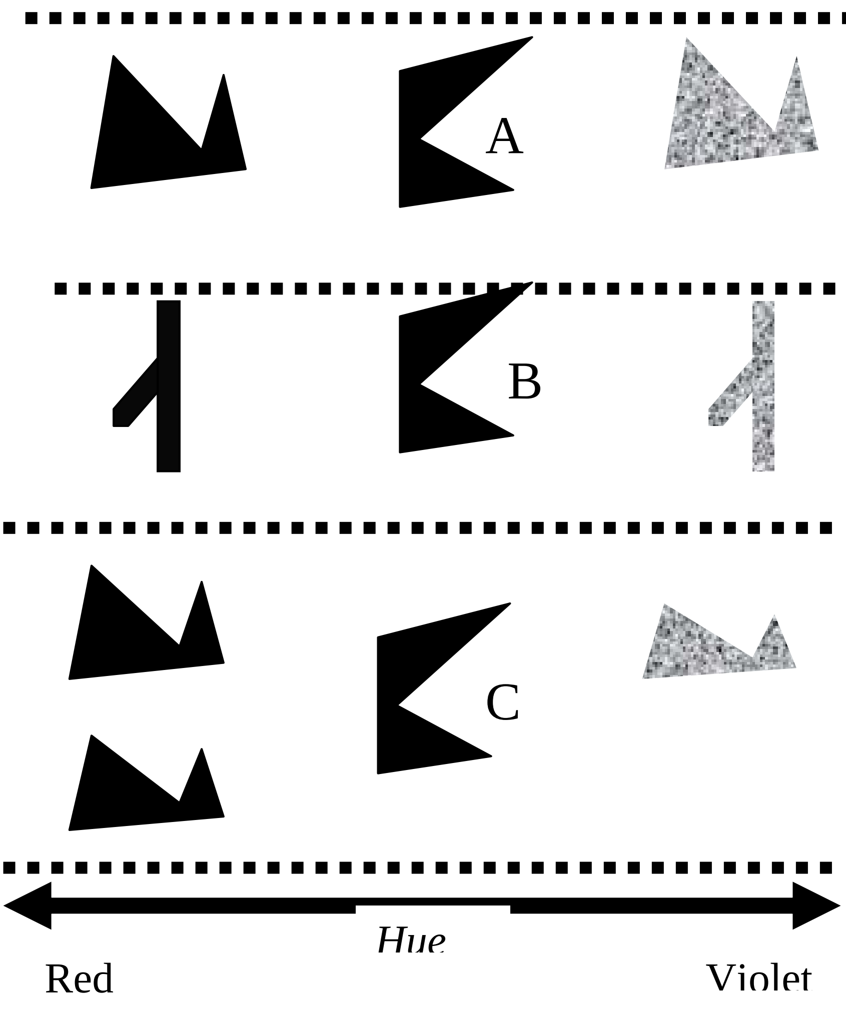
1.1.4. Contrast due to the irrelevant information

Unlike the experiments reviewed in the previous sections, the few studies presented here show contrast in judgment depending on an irrelevant to the task dimension. The irrelevant dimension was considered as a context in judgment of the stimulus along the relevant (target) one.

Goldstone (1995) reported contrast effect on adjustment of the color of a stimulus depending on the range of the stimuli that possess a different irrelevant dimension. The relevant dimension was the color of the stimuli, which ranges from red to violet. The irrelevant dimension was the form and direction that the figures faced. Examples of the stimulus material in two controls and one of the experimental conditions are shown in Figure1.

Participants were asked to adjust the color of a five-sided polygon with right-pointing prongs (polygons A, B and C) in order to match the color of another figure. The standard figures actually depicted the context in this experiment. Both contextual and target figures were five-sided polygons but they differed in the pointing direction of their prongs. The prongs of the standard polygons pointed upwards. This situation is illustrated in the lower part of fig. 1. The adjusted color of polygon C was slightly different from the color of the standard polygons. For example, if the color of the standard polygons was predominantly red, the color of the target polygon was adjusted to be more violet than the standard's color. The adjusted color of C was more violet than the adjusted color of A (the colors of the two controls A, B did not differ). This result indicates contrast in the adjustment of C with the color of the rest of the polygons. Thus, the C polygon was probably excluded from the set of red up-pointing polygons and was perceived as more violet than it actually was.

Figure 1. Examples of stimuli used in 3 conditions of Goldstone's (1995) experiment. The shading of the figures indicates their color- darker figures are red, lighter ones are violet.



Another example of contrast effect obtained through differentiating the range of the stimuli depending on an irrelevant to the task dimension was the experiment by Arieh and Marks (2002). Participants were asked to judge length of lines (vertical and horizontal lines). The irrelevant dimension was the orientation of the lines – vertical or horizontal. Context A comprises relatively long vertical and short horizontal lines, whereas Context B comprised relatively short verticals and long horizontals. The length of the stimulus lines presenting in context A and context B are presented in table 4.

Table 4.The lengths of stimulus lines depending on the context.

Experimental conditions		Lengths of the stimulus lines					
Context A	***Vertical lines***			1.9cm	2.85cm	4.12cm	6.0cm
	Horizontal lines	0.65cm	0.95cm	1.9cm	2.85cm		
Context B	***Vertical lines***	0.65cm	0.95cm	1.9cm	2.85cm		
	Horizontal lines			1.9cm	2.85cm	4.12cm	6.0cm

The lines of a middle length, i.e. the ones with length 1.9 and 2.85cm were common for all contexts and were presented both in vertical and horizontal position. The common lines presented within Context A received lower ratings when they were presented vertically than horizontally on the screen. On the contrary, the same lines shown in Context B were estimated as longer when presented vertically than horizontally. The size of the horizontal-vertical illusion (HVI) was reduced by 7.15% under contextual condition A and was enlarged by 5.75% under Contextual Condition B compared with the base line condition. These effects were called “differential context effects” (DCEs).

A series of further experiments (Arieh and Marks, 2002) tested the spatial specificities of the DCEs with the same stimuli (vertical and horizontal lines). The visual field was divided into 4 quadrants and the contextual set of vertical and horizontal lines was induced in just one of them. The effect was underestimating the long vertical lines in condition A and the long horizontal lines presented in condition B. Thus, it seems that DCE did not transfer from the specific part of the visual field to another. If the contextually induced set of lines, however, was presented to the same part of the fovea, the effect was found in different regions of the external visual space. The possibility for projecting the lines on relatively the same place of the retina was controlled through the instruction to look always to a cross that appears before the lines. The lines that comprise context A were projected at one spatial location and the stimulus lines from context B were presented at another. Participants judged a random sample of the lines from Context A and B. Thus, if the effect from Context A was transferred to the judgments of the lines of Context B, the effect should be cancelled out. Results showed that the magnitudes of the HVI obtained in Context A, Context B and the base line were all similar in size. According to the authors, these results demonstrate that DCA is produced at the retinal level.

The DCE was repeatedly demonstrated with different stimuli: e.g., in judgment of length of vertical and horizontal lines (Potts, 1991), taste (Rankin and Marks, 1991, 1992), haptic touch (Marks and Armstrong, 1996) and olfaction (Rankin and Marks, 2000).

To conclude, context could be presented by the means of an irrelevant to the task stimulus dimension. It was argued that this effect was a result from low level processes like sensory adaptation or perceptual learning. Goldstone (1995) demonstrated, however, that slight differences in the stimulus set could result in assimilation of the same target rather than in contrast (for details, see section for assimilation due to the irrelevant information).

1.1.5. Summary and discussion of the empirical data for contrast effect

Most of the studies that reveal contrast effect in judgment actually manipulate the range and the frequency of the stimulus set. Moreover, changes in range and frequency of the stimuli seem to describe quite well both contrasts with simple and complex stimuli (Parducci and Perret, 1971; Mellers and Birnbaum, 1981; Wedell, et al., 1987; Manis et al., 1988).

The distinction of the levels of contextual influences seems to be important and profitable one. If particular channels are responsible for processing particular contextual information or at least some effects could be observed on this level, it is possible to find some particular contextually sensitive mechanism of processing of the input information. For example, Marks and his collaborators (Marks, 1988, 1992, 1994; Marks and Warner, 1991) argue for contextually sensitive adaptation. Goldstone (1995) discussed the possibility that the contextual influence on the input information is a form of perceptual learning. The effects of lower level information processing, however, are hardly traceable to the higher levels. It is difficult to predict how contextually dependent judgment of the length of lines reported by Arieh and Marks (2002) would affect judgment of woman attractiveness. Moreover, it is quite possible that the effects from the lower level blot out on the upper level. For example, the assimilation toward a perceptually similar category demonstrated by Goldstone (1995) may counter-balance the contrast found at the level of human receptors. It could be, however, that context influences judgment on different levels – both lower and higher levels of information processing.

1.2. Empirical Evidences for Assimilation Effect

Several groups of experiments demonstrating assimilation effect in judgment are reviewed in this section. Like in the previous section for contrast effect, empirical data was divided into groups of studies using similar experimental procedure.

1.2.1. Anchor placed inside the stimulus series

A few of the experiments showing assimilation of simple stimuli to the current context rely on the anchoring procedure. For example, Sherif et al. (1958) obtain assimilation of the judgments of the stimulus series toward the anchor (a standard for comparison), which was equal to the heaviest or the lightest stimulus in the original series. Each subject rated once each of the stimulus weights alone and then coupled with an anchor on a 6-point scale. When the weight of the anchor was equal to the heaviest stimulus in stimulus set (141gm) judgments of the stimuli were higher than judgments of the same stimuli presented alone.

The same effect was observed when the anchor was equal to the lighter weights in the original series. An assimilation of the rating of the stimuli in the stimulus series toward the value of the anchor was observed.

The experiments of Sherif et al. were replicated by Parducci and Marshall (1962) and the same shift in judgment was observed except for the lighter anchor condition, where judgments of the stimuli in the original stimulus series did not differ significantly from judgments of the same stimuli presented without an anchor.

In sum, both experiments briefly described in this section reveal the possibility for assimilation to be a consequence from an anchoring stimulus belonging to the range of stimulus set – in fact, assimilation was unambiguously obtained only when the anchor was equal to the heaviest stimulus in the stimulus series. Unfortunately, the results are not so definite about the assimilation toward the lightest member in the stimulus set. This casts some doubt on the general conclusion that an anchor that belongs to the stimulus set causes an assimilation, while an anchor that does not – a contrast effect.

1.2.2. Primed context

Most of the experiments showing assimilation effect in judgment used priming of particular contextual information. These studies usually require judgment of ambiguous or complex stimuli.

- *Assimilation of the ambiguous target toward moderately low/high context with respect to the judgment dimension*

An experiment of Herr, Sherman, and Fazio (1983) demonstrates assimilation of an ambiguous target to the moderate context. They primed 4 types of animal sizes: Extremely high, Moderately high, Moderately low, Extremely low and asked participants to rate the size of 2 ambiguous target animals (fictitious) on the 11-point scale (for examples, see table 1 in section about contrast due to primed context).

When the primed stimuli were animals of moderately high or moderately low size targets' ratings were assimilated toward them. The opposite contrast effect, however, appeared when the primed context was an extreme one (extremely high or extremely low, for details, see the section about contrast).

- *Assimilation in judgment of an ambiguous target toward the accessible contextual information*

Srull and Wyer (1979) studied how judgment of an ambiguous target may depend on the accessible context. The accessibility of a certain trait was manipulated through a sentence-construction task, i.e. to choose 3 out of 4 words that would make a complete sentence. Some of them primed hostility others did not (fillers). For example, "leg break arm his" conveys hostility, while "her found knew I" does not. Then participants were asked to read a short text describing the stimulus person (Donald) whose behavior was ambiguous with respect to the primed concept. The task was to form an impression of the target person and to rate him on a 10-point scale along a series of different trait dimensions. Six of the dimensions implied either high or low hostility (hostile, unfriendly, dislikable, kind, considerate, thoughtful), and the rest of them - were descriptively unrelated to the primed concept (boring, selfish, narrow-minded, dependable, interesting, and intelligent).

Ratings of the Donald's behavior along the hostility dimensions increased with the number of times hostility-related concepts had previously been activated. The ambiguous with respect to hostility Donald's behavior was rated as more hostile depending on how accessible the concept was. The effect of priming decreases with the time-interval (immediate, after 1 hour, after 24 hours) between the priming task and presentation of Donald's behavior. The same results following the same procedure were obtained with the priming of kindness.

- *Assimilation on the encoding stage of information processing*

Srull and Wyer (1980) studied when the contextual influence takes place on encoding stage or during the integration of information. The procedure and stimulus material were identical to the study described above. First, priming of hostility, second, acquisition of Donald's behavior and third, judgment of the target person along several dimensions (6 trait dimensions that convey hostility and 6 trait dimensions, which were not related to it). When participants experienced a *delay* between the *activation* of the trait category and the *acquisition of Donald's behavior*, their ratings along the hostility dimensions increased with the number of the presentations of the priming category but decreased with the length of the delay. On the contrary, when participants experienced a *delay* between *acquisition* of the stimulus behavior and *judgment* of this behavior, their ratings increased with both the number of prior activations and the length of the delay. However, when the trait category was activated after acquisition of the stimulus behavior, none of these effects occurs. Authors interpret these results as an evidence for the influence of the context on encoding the information upon judgment.

- *Expectation-driven assimilation*

Several experiments focus on assimilation toward a stereotype. These studies start with an induction series of creating a particular stereotype. Then a judgment of stimuli that belongs or not to the stereotyped category is required. For example, Manis, Nelson, and Shedler (1988) study the influence of the stereotypes (extreme/moderate) on the choice of a more pathological statement within a pair of midscale in pathology items. During the induction phase a stereotype with different intensity to either Central or Metropolitan Hospital was created. The extremity of the induced stereotype was controlled through the percentage of the extremely pathological statements presented during the induction phase (52%, 68%, 84%, or 100%). Participants' task was to choose the more pathological case within a pair of midscale statements and to evaluate their confidence about the choice they made on a 7-point scale.

Although the test items from both hospitals were chosen to be moderate, participants who received in the induction part 52%, 68%, or 84% extremely pathological statements consistently chose as more pathological the statements that originated from the stereotyped hospital. Authors call this effect assimilation toward the stereotype. It was observed only when the *stereotype was moderate* while for the extreme one a contrast was observed. Moreover, the assimilation was higher when the stereotype was established through 52% distorted statements than if the induction series included 68% or 84% extreme in distortion items.

Manis and Paskewitz (1984) induced a stereotype by presenting a set of highly pathological statements to part of the participants and low in pathology statements to the others. Participants were asked to judge whether each statement was produced by a schizophrenic person or not. Then both groups received the same moderate in pathology test items - 30 definitions rated on the pretest between 4.1 and 7.8 on an 11-point scale. Part of the participants was asked to “guess” the number of schizophrenic patients in the subsequent test series and whether the next patient is schizophrenic or not.

The respondents’ general expectation for the overall test series was consistent with the type of statements they were presented in the induction phase. Participants who rated highly pathological statements expected the rest of the statements to be highly pathological, while participants who rated low pathology definitions first expected the rest of the definitions to come also from patients with low pathology. Guesses before each trial followed the same fashion - they were congruent with the type of definitions that was presented in the induction series. Moreover, the correlation between guesses (either for the overall degree of pathology in the test sample or trail-by-trial expectations) and subsequent judgments of the amount of thought distortion was consistent and positive but quite low (from 0.07 to 0.18).

Similar expectation driven assimilation was reported also in judgment of the height (Manis, Biernat and Nelson, 1991) and judgment of self and others (Biernat, Manis, and Kobrynovicz, 1997).

- *Context included in the representation of the target*

Strack, Schwarz, and Gschneidinger (1985, experiment1) asked participants to write about 3 recent events or 3 events from the past. Part of the Ss was instructed to think about positive events, while the other part -

about negative events. Then, participants judged their happiness and life satisfaction along an 11-point scale. Ss who recalled 3 positive recent events rated their happiness and satisfaction as being higher then participants who recalled 3 negative recent events first. Participants who recalled 3 negative recent events assessed themselves lower on both dimensions then participants who recalled 3 negative past events. Authors argue that when the activated information belongs to the judged period judgment is assimilated to it. On the contrary, when the activated contextual information is excluded from the judged category a contrast could appear. Participants who were induced to think about 3 positive events from the past then rate themselves as less happy than participants who first recalled 3 negative past events or 3 positive recent events.

To conclude, experiments reviewed in this section demonstrate assimilation toward an available category, which represents the context of judgment. Usually such experiments rely on judgment of moderate target stimulus within a moderate context. The exception, however, was the experiment of Herr et al. (1983), where the judged stimuli were ambiguous (fictitious) animals. Their ratings were assimilated only to the moderate but not to the extreme context. Thus, the reported experimental result may be roughly summarized in the following way: judgments can be assimilated toward the activated contextual information when the target and the contextual stimulus belong to the same category.

1.2.3. Sequential effects

Correlation between the ratings or the values of the current and previous stimulus is known as sequential assimilation. Stimuli are overestimated when judgment of the prior stimulus was high and are underestimated when judgment of the prior stimulus was low. Sequence effects seem to be quite robust (Ward, 1973, 1979; Jesteadt, Luce and Green, 1977; Lockheed, 1992; Petrov and Anderson, 2000, 2005). Such effects are found independently of stimulus complexity. For example, Lockhead (1992) studied the integral nature of bivariate stimuli (auditory tones with loudness and pitch) through the sequence effect. He argues that people cannot attend only to the target attribute neglecting the rest attributes of the stimuli but rather they process the entire stimulus before abstracting the attribute value. Lockhead designed a set of 10 tones with two nonlinearly correlated characteristics: loudness (amplitude of 79 to 88 dB SPL in 1dB steps) and pitch (frequency of 1.000 to 1,045 Hz in 5Hz steps). Each amplitude was paired consecutively with a particular frequency. Participants were asked to identify only the intensity of each tone on a 10-point scale. The data shows that sequence effect reflects the bivariate space. Judgments of each tone tend to be assimilated toward the value of the previous bivariate stimulus. The sequence effect was not detected only for the target dimension (intensity of the tones) but rather for the combination of both stimulus characteristics (amplitude and pitch). The assimilation toward the previous response was greater for no-feedback than for the feedback condition but the assimilation toward the value of the prior stimulus was greater for the feedback than for the no-feedback condition. Response time (RT) also depends on sequence - judgments are faster when successive stimuli were more similar (the Euclidian distance in the frequency-amplitude space).

To sum up, Lockhead's (1992) experiment draws attention to several things:

- First, sequential assimilation can be found both in ratings and in stimuli themselves;
- Second, feedback/no-feedback conditions may determine whether assimilation would be toward the previous rating or toward the value of the

previous stimulus. In no-feedback conditions assimilation toward the previous response was greater than for the feedback condition;

- Third, judgments reflect characteristics of the whole stimulus rather than only the target attributes;
- Fourth, RT in judgment depends on the similarity between the successive stimuli.

1.2.4. Assimilation due to an averaging between simultaneously presented stimuli

A body of research initiated by Anderson (1966) demonstrates assimilation due to an averaging between simultaneously presented stimuli and context. Anderson argued that the reported assimilation seemed to be independent of stimulus ambiguity.

- *Component ratings in the impression formation task*

Anderson (1966, 1971b) found an assimilation effect in judgment of the likeability of personal traits toward the value of the contextual ones. He asked people to rate first, the likeability of each person described by 3 adjectives on a 9-point scale and then to judge how likable/dislikable is one of the aspects (the test adjective) of the person on a 20-point scale. The other two adjectives actually represent the context in these experiments. Both the test and contextual adjectives were divided into 4 groups: H (highly favorable), M+ (moderately favorable), M- (moderately unfavorable) and L (highly unfavorable). Ratings of target adjectives vary with context. For example, the rating of the same H adjective decreases from 16.6 in the context of two H adjectives to 14.1 in the L context. The author interprets these results as a positive context effect, i.e., judgment of the component was displaced toward the value of the other traits of the person presented simultaneously on a card.

In another experiment (Anderson et al., 1973) participants received 3 pairs of adjectives per card, but each pair described a different person. They had to rate (1) attractiveness of each person in a group, (2) attractiveness of the group as a whole, and (3) importance of each member in the group. The judged attractiveness of each member in the group rose or became lower depending on the attractiveness of the other two persons of the group. When the test person was described as a leader, his/hers attractiveness became even more positive in the context of 2 very attractive adjectives and more negative in the context of two very unattractive adjectives. The judgment of the group attractiveness was the average of the attractiveness of all members of the group and could increase when the test person was described as a leader.

- *Assimilation of the ambiguous target toward the available context*

Another set of experiments goes into contextual shift in judgments of ambiguous target (a target with more than 1 meaning). In contrast to Anderson[1] (1965, 1971b, 1973), researchers who work in this direction consider adjectives as having variety of interpretations. The context determines which interpretations of the adjectives are inappropriate. Wyer (1974) demonstrates an increasing contextual influence upon judgment of an ambiguous target with the rise in the ambiguity of the target stimulus. The test adjectives were divided with a pretest into moderately high in likableness adjectives (M+), neutral in likableness adjectives (0) and moderately low in likableness adjectives (M-) and in the same time into High and Low in ambiguity adjectives. For example, *high in ambiguity* and moderately *high* in likableness adjectives were cautious, ingenious, intellectual. *Low in ambiguity* and moderately *high* in likableness adjectives were innocent, courteous and healthy. *High in ambiguity* and moderately *low* in likableness adjectives were loud-mouthed, ill-tempered, sarcastic. *Low in ambiguity* and moderately *low*

[1] Anderson claimed that the meaning of the adjectives is contextual-free.

in likableness adjectives were self-centered, troublesome, showy. These types of ambiguous stimuli were grouped with four groups of unambiguous contextual adjectives: H (high in likableness), M+ (moderately high in likableness), M- (moderately low in likableness) and L (low in likableness).

Participants were told that each set of three adjectives (1 ambiguous+2 unambiguous) describe a particular person and their task was first, to estimate how well they like him/her and then, how well they like a particular trait of that person (represented by the test adjective) on a 10-point scale. The ratings of the ambiguous test adjectives increased with the favorableness of the contextual pair of the personality adjectives. For example, adjectives from the test group M- were rated as more likable when they were presented with 2 high in likableness unambiguous adjectives than two moderate or law in likableness personality adjectives. An ambiguous test adjective that came from group M+ received lower ratings when it was coupled with a contextual pair from the low in likableness adjectives and so on. Moreover, c*ontextual effect increased* with the *ambiguity* of the test adjectives. Thus, Anderson's (1965, 1971b, 1973) averaging seems to depend also on the ambiguity of the target stimuli. The ratings of the high in ambiguity adjectives were assimilated more toward the meaning of the contextual stimuli then the ratings of the low in ambiguity adjectives.

- *Assimilation within a pair of stimuli*

Like in Anderson's (1965, 1971b) experiment, Wedell et al., (1987) found assimilation between simultaneously presented stimuli. They designed pairs of female faces. Six pairs were contextual ones and consisted only of high, moderate or low in attractiveness faces. Another 6 pairs were test one and comprised one moderate in attractiveness face and one contextual that was either high or low in attractiveness female face. Assimilation was found in the ratings of the faces within each of the 6 test pairs. The same face was judged as more beautiful when presented with a highly beautiful face within the test pair and as less beautiful, when the contextual face in the test pair was low in attractiveness. In order to facilitate assimilation participants were told that each pair of faces represents "friends from the high school". Another implicit factor for the assimilation within each test pair was probably the procedure for skewing the stimulus distribution by the firs six contextual pairs. Faces in these pairs were always drawn from the same group of attractiveness (high, moderate or low in attractiveness faces), which may implicitly to predispose participants that the next 6 pairs are also including comparable in attractiveness faces. Authors interpreted the observed assimilation similar to the averaging found in Anderson’s studies.

In conclusion, experiments reviewed in this section rely on the specific instruction that the simultaneously presented stimuli have something in common, e.g. “traits of the same person”, “traits of group members” or “friends in the high school”. Thus a question about what causes assimilation remains, namely, whether the simultaneous presentation or the specific instruction is responsible for the obtained assimilation.

1.2.5. Assimilation due to an irrelevant to the task dimension

Goldstone (1995) found a significant assimilation effect in color perception. This effect appeared both toward some existing conceptual categories which were irrelevant to the task and to contextually sensitive on-line groupings of the stimuli depending on their irrelevant characteristics.

In the first experiment, participants were asked to level the color of a black object to the color of a particular sample. Samples were letters and numerals, which color ranged from red to violet, however, the color of the letters was more reddish than the color of the numerals.

Judgment of the color of the sample was made on the objective continuum of the hue between red and violet. Participants report their judgments by means of the color chosen for the second identically shaped object. The context in this experiment was the irrelevant to the task category of the sample and the modified object (letters and numerals). When letters and numerals with the same physical color were samples in the experiment, the modified color of the target stimuli was more reddish for letters and more violet for numerals. The color of the target was assimilated toward the categories of “red letters” and “violet numerals”.

The same assimilation was obtained when the stimuli were five-sided polygons and two-line branches. Judgments on their color were found to depend on an on-line categorization of the stimuli depending on the irrelevant to the task shape. For example, if the polygons were judged within a set of two-lined branches and five-sided polygons its color was adjusted with respect to the color of the rest of the polygons presented in a particular experiment. Moreover, participants were inclined to judge the color of the target stimulus depending on the perceptual similarity between the rest of the stimuli. When among the stimuli was only one polygon except the target one, the color of the target stimulus was adjusted toward the color of the other polygon. However, if the color of the same target polygon was judged within a sample of other polygons, its color was adjusted in the opposite direction – a contrast effect due to the dissimilarity between the target and the other polygons with prongs pointing in different direction (for more details, see section about contrast effect).

These experiments demonstrate assimilation effect at the level of perception of the object’s color. Participants judge the color of the objects depending on contextually sensitive categories. Moreover, it seems that participants spontaneously create such on-line categories depending on some irrelevant to the current task dimension like the shape or the conceptual identity of the objects.

1.2.6. Summary and discussion of empirical data that reveals assimilation effect

Empirical data that reveals assimilation effect does not fall easily into any grouping. Different researchers imply different meaning of context. Some of them perceive context as the most accessible and applicable information (Srull and Wyer, 1979, 1980), others specify that such information could be elicited from memory (Strack et al., 1985). Third part, comprehend context as external stimulus, e.g. stimulus presented simultaneously with the target one (e.g., Anderson, 1966, 1971; Wedell et al., 1987; Goldstone, 1995), stimulus presented successively to the target one (e.g., Sherif et al., 1958; Parducci and Marshall, 1962; Mellers and Birnbaum, 1982), or prime stimulus (e.g., Barg and Pietromonaco, 1982; Herr et al., 1982). Moreover, some data reveals assimilation of ambiguous stimuli (e.g., Herr et al., 1982; Srull and Wyer, 1980), other data - of unambiguous ones (e.g., Sherif et al., 1958; Parducci and Perrett, 1971, Goldstone, 1995). These substantial distinctions in the existing experimental studies are a challenging obstacle for any classification. Thus, the only criterion that seems to generalize the findings described so far sounds quite abstract –if the current context is affiliated with the target category the judgment of the target could be assimilated towards it.

It is important to mention that affiliation to a particular category in the judgment could be considered as contextually dependent (Goldstone, 1995). Category membership is assigned or not to a certain stimulus depending on the current context. The same stimulus could be categorized differently in different situations. Even very simple and meaningless stimuli like polygons and branches are sensitive to the current context. This *contextually dependent on-line categorization* seems also to be *perceptually grounded*. People distinguish objects in the environment depending on their perceptual similarity even without an explicit task to do so.

1.3. Discussion of the empirical data and conclusion

This section attempted to reveal the controversies and weak points in the field of relative judgment. The major problem, in my opinion, is the absence of unambiguous answer 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? It is also possible that these factors compete with each other and the outcome becomes not easily predictable.

Stimulus presentation mode is an appropriate example of the first idea (variations of the same factor results in both contextual effects): successive presentation of the contextual and target stimulus result in contrast, while simultaneous one – in assimilation (Geiselman, Haight, and Kimata, 1984). This idea, however, is not evident from the empirical review. For example, Sherif, Taub and Holland (1958) demonstrated both effects through the same procedure and within a single experiment. The only variable that they varied was the place of the anchor bur not the procedure for anchor presentation. Other experiments that contradict the same hypothesis, namely that the presentation of the stimuli may cause a particular contextual effect, are the experiments on expectation driven assimilation (Manis and Paskewitz, 1984; Manis at all, 1988) and multiattribute judgments (Melers and Birnbaum, 1882).

Another example of the idea that variations of the same factor may result in both contrast and assimilation is the hypothesis that category membership of contextual and target stimulus is crucial for the subsequent context effect (Sherif et al., 1958; Strack et al., 1985). C*ontrast* may result from the presentation of a context that does not belong to the category to which the target stimuli are assigned. *Assimilation*, on the other hand, may result from the presentation of a context that belongs to the target category if the context *does not change the frequency* of the stimulus distribution. Moreover, the categorization can be considered an on-line context-dependent process (Goldstone, 1995). Experimental data that seems to contradict to category membership hypothesis come from the experiments on contrast effect that rely on frequency of the stimulus distribution (e.g., Parducci and Perrett, 1971, Wedel et al, 1987). Although, stimuli that provoke contrastive shift in judgment belong to the same category as the target one they push the ratings of the target away from them if they are densely packed. Other contradicting experiments are the ones that require judgment of ambiguous stimuli. It seems that people are more susceptible to assimilate their ratings of an ambiguous target towards the available context (Srull and Wyer, 1979; Herr et al., 1983; Wyer, 1974; Anderson, 1966, 1971b).

The most "popular" and implicitly shared idea in the field seems to be that the two contextual effects are caused by different factors. Some researchers explore the causes for contrast (e.g., Parducci and Perret, 1971; Arieh and Marks, 2002; Mellers and Birnbaum, 1981; Wedell et al., 1987), others – for assimilation (e.g., Anderson, 1966, 1971b; Wyer, 1974; Srull and Wyer, 1979; Manis and Paskewitz, 1984). Moreover, research on contrast has been typically performed with simple rather than complex stimuli and vise versa, most research on assimilation effect has been performed with complex rather than simple stimuli.

The situation, of course, may be even more complex if we consider the evidence for low-level contrast (Arieh and Marks, 2002; Goldstone, 1995) and low-level assimilation (Goldstone, 1995). Context, however, may influence information processing at different levels (for a review, Wedell, 1994). Unfortunately, the question about the level of appearance of contextual effects is a relatively new one. Most of the studies that reviewed empirical evidence for contrast and assimilation do not differentiate where the effect appeared.

Finally, review of the experimental data on judgment evidently showed that there is no terminological agreement on the stimuli used in the experiments for studying contrast and assimilation. On one hand, context was not consistently specified across studies and on the other hand, stimuli that differ in ambiguity and complexity were judged. Moreover, there is no methodological clarity in the obtained contextual effects. Contrast and assimilation were demonstrated with various methodologies, e.g. different types of primes such as category priming, anchoring, variation of the instructions. Thus, the stability of the obtained contextual effects seems to be questionable.

CHAPTER 2
Theories of Judgment

Current theories rarely refer to mechanisms. The theories can roughly be divided into two camps: psychophysical and social. The first one focuses on the product of judgment rather than on the process itself. The strength, however, is in its precise methodology for testing models, and especially in the use of simple stimuli for studying judgment. Simple stimuli have an advantage over complex ones in giving researchers the chance of knowing what actually people judge. Complex stimuli (e.g., faces, personalities), on the contrary, may posses an innumerable number of features. Moreover, some features may form a new and meaningful entity within the stimulus itself. Theories that focus on explaining judgment of such complex and often ambiguous stimuli are the social ones. They have the advantage of proposing mechanisms for explaining contextual effects in judgment. These two distinct points of view are brought together because of several reasons:

- First, it is more efficient to find and formulate mechanisms that are simple enough, but explain many phenomena of interest. Moreover, it is quite possible that the mechanisms underlying judgment of simple and complex stimuli to be the same.

- Second, there is *no* evidence that the contextual effects in judgment of simple and complex stimuli are governed by essentially different processes.

- Third, several experiments demonstrate the existence of contrast effect due to change in range and/or frequency of the stimulus set which holds both in psychophysical and social judgments. For example, contrast was observed in judgment of the size of squares (Parducci and Perrett, 1971; Sarris and Parducci, 1978), darkness of dot-patterns (Mellers and Birnbaum, 1982), physical attractiveness (Wedell et al., 1987), clinical judgments (Wedell et al., 1990) and judgments of the self and others (Biernat et al., 1997).

Thus, since various approaches and methodologies have obtained the same contrast effect, it would be possible to assume that the same mechanisms may govern the judgment process. Moreover, to the best of my knowledge, there are no experimental results that differentiate judgment mechanisms based on level of complexity of the stimuli.

The next section reviews the tree most relevant to the aims of the present research psychophysical theories for judgment.

2.1. Psychophysical point of view

Tree theories for judgment of simple stimuli are briefly reviewed in this section. The first two of them are focused more on describing the product of judgment than the process itself, while the third theory proposes particular mechanisms that may underlie judgment.

- Adaptation-Level (AL) Theory
- Range-Frequency Theory
- ANCHOR Model

- **Adaptation-Level Theory**

The Adaptation-Level (AL) theory elaborated on the analysis of size-contrast illusions (Helson, 1964). It states that judgment (J) of a test stimulus (X) depends on the ratio of X to the adaptation level (A).

$$J(X)=X/A \qquad (1)$$

Consider a situation of judgment of the length of line *X*. If the longer line *Y* is presented along with the target *X*, the *A* will shift toward the value of the contextual line *Y* (i.e., *A* will become higher at this particular example). Thus, if you consider equation (1), judgment of line *X* will decrease. This is a typical explanation from the AL theory. It predicts always contrast from the context presented during judgment.

Contradictory to the AL Theory explanation were the results of Parducci and Perrett (1971). In their experiment judgment of stimuli from sets with the same mean shifted depending on the range of the stimulus set.

In sum, AL Theory describes judgment in the presence of external contextual stimulus, i.e. stimulus from the surrounding environment. It predicts always contrast. Restle (1978), however, argued that after modifications of theory's equations AL Theory could account also for assimilation. Unfortunately, AL Theory did not focus on the mechanisms underlying contextual effects. The equations, which the theory propose describe the product rather than the process of judgment.

- **Range-Frequency Theory**

Range-Frequency Theory proposed by Parducci (1965, 1968, and 1974) assumes that judgments are determined by range and frequency principles. Each rating results from a compromise between the two principles. According to the *range principle (R),* the judge assigns the two extreme categories to the two extreme stimuli and then divides his psychological range into sub-ranges whose relative sizes are independent of the stimulus condition. The *frequency principle* (*Fr*) asserts that the judge employs the alternative categories at equal frequency. *R* and *Fr* are characteristics of the so-called comparison set. Whenever the comparison set is known the judgment of the target stimulus could be predicted.

Suppose that the task is to judge the length of lines on a 7-point scale. The range value of the stimulus depends on: (*a*) the subjective value of the target stimulus; (*b*) the value of the smallest line presented to the subject before the target line; (*c*) the longest line presented to the subject before the target line.

Parducci proposed that the standardized range value could be calculated on the base of the following equation:

$R_i = (E_i - E_{min})/(E_{max} - E_{min})$,

where E_i is the subjective value of the stimulus, E_{min} is the minimum value of the observed set of stimuli and E_{max} is the observed maximum among the values of the stimulus set. R_i is a number between 0 and 1. If the judgment has to be situated on a scale from *a* to *b*, R_i has to be rescaled with respect to the two extremes of the required scale:

$R_i' = a + (b-1) R_i$,

where *a* is the lower end of the scale on which the stimulus has to be judged (i.e., 1 on a 7-point scale) and b represents the upper end of the scale (i.e., 7 on a 7-point scale), and R_i' – a rescaled value of the stimulus range (R_i).

The frequency value of a certain stimulus, on the other hand, represents the mean rating that a given stimulus would elicit if each of the rating categories were used with equal frequency. It could be calculated through the following equation:

$F_i = (r_i - 1)(N-1)$,

where r_i is the rank of the stimulus within the stimulus distribution and N is the total number of stimulus presentations. F_i is a number between 0 and 1. It has to be rescaled with respect to the scale on which the judgment of the stimuli has to be done:

$F_i' = a + (b-1) \times F_i$,

where *a* and *b* are the two extremes of the judgment scale. For example, if the stimuli have to be judged on a 7-point scale, *a* and *b* would be respectively, 1 and 7.

Approximating to a whole number, the frequency value could be calculated if the total number of stimuli presentations is divided into the number of available categories. For example if the subjects in our hypothetical experiment were presented 112 times with some target lines, then we could find the frequency value if first divide 112 to the 7 scale categories, gets 16 and then count by rank. The frequency value of the first 16 stimuli would be 1, for the next 16 - 2, etc.

Finally, judgment is accomplished according to the following formula:

$$J_i = (R_i' + F_i')/2$$

Where J_i is the subjective judgment of the i-th stimulus, R_i' is the judged stimulus location in the range of stimulus values, and F_i' is the frequency value of the stimulus in the distribution of stimulus values.

Later Range-Frequency Theory was further elaborated to account for transfer effects of changing the distribution of the comparison set (Parducci and Wedell, 1986). It was assumed that the comparison set includes the most recent (e.g., last 15) trials.

In general, the result of the compromise between these two principles is in the direction of contrast. In other words, when contextual stimulus is added beyond the lower or the upper end of the stimulus distribution, the ends of the stimulus set change and the ratings of the target stimuli will increase or respectively decrease. Alternatively, when there is a change in the frequency of the stimuli in any part of the range, the result is again a contrast effect.

Range-Frequency Theory has provided a good quantitative fits to judgments of simple stimuli (e.g., Parducci and Perrett, 1971; Sarris and Parducci, 1978), multiattribute judgments (e.g., Mellers and Birnbaum, 1981, 1982; Cooke and Mellers, 1998; Mellers and Cooke, 1994) and judgments of complex stimuli (e.g., Weddel et al., 1987, 1990).

Unfortunately, Range-Frequency Theory did not focus on the mechanisms that may underlie judgments. It is not specified neither how the comparison set is formed nor the mechanisms for extracting the *R* and the *Fr* values from the comparison set.

- **The ANCHOR Model**

The ANCHOR Model (Petrov and Anderson, 2000, 2005) is a computational model of judgment based on the ACT-R cognitive architecture (Anderson and Lebiere, 1993). Unlike, Range-Frequency Theory, the ANCHOR Model proposes specific mechanisms for judgment of simple stimuli. ANCHOR models judgment and absolute identification (i.e., judgment with a feed-back) of line segments. The length of the lines was transformed by the model into internal magnitude, using calculations consistent with Weber's low. Thus, the whole perceptual subsystem was reduced to a single stochastic equation that abstracts away factors like habituation, perceptual contrast, Gestalt etc. The obtained magnitude of the stimulus was compared to a set of anchors (prototypes for each scale category) kept in the Long Term Memory (LTM). The best fitting anchor (the most active one) became a reference point for further comparisons that result in the final judgment produced by the model. Anchor selection is a probabilistic process, sensitive to similarity between the internalized stimulus magnitude, base-level activation and recency.

Although the ANCHOR model provides a detailed description of the mechanisms underlying the judgment process, it has several limitations. Unlike, the more general Range-Frequency Theory it can process only simple stimuli. Moreover, the ANCHOR model was specifically tuned to generate answers only on a 9-point scale. Therefore, the issue for transferring judgment from one scale to another and how a 100-point scale could be represented in LTM via a set of anchors remains open.

- **Conclusion**

To conclude, the theories presented in this section seem to share the assumption that target stimuli are judged against a particular context. The context, however, is considered as external by the AL Theory (e.g., contextual stimulus that changes *A*) and as internal by the Range-Frequency Theory (i.e., *R* and *Fr* of the comparison set) and the ANCHOR Model (i.e., anchors kept in LTM). Moreover, context may dynamically vary with time as the ANCHOR Model describes it.

Another open issue in the field concerns the complexity of the stimuli: namely, whether the theories can explain the judgment of both simple and complex stimuli. The Range-Frequency Theory is the only theory that takes an effort at this direction. It could, however, be the case, that Range-Frequency Theory describes contextually sensitive *mapping* between the internal magnitude and the scale (Parducci, 1965), while others, e.g. AL Theory, are more interested in contextually sensitive *perception*.

Unlike the three theories reviewed in this section, the social theories of judgment briefly presented in the next section focus more on the mechanisms than on the product of judgment. In addition, social theories unlike psychophysical ones more often predict assimilation than contrast.

2.2. Social Cognition Theories

The following section discusses contrast and assimilation from the point of view of:

1. Two-Path Theory
2. Integration Theory
3. “Change in Meaning” Theory
4. Inclusion-Exclusion Theory
5. Norm Theory

- **Two-Path Theory**

Two-Path Theory is a judgment model proposed by Manis and Paskewitz (1984). They assume that context exerts both assimilative and contrastive forces upon human judgment. Context violates judgment through these two independent paths. On one hand, prior exposure to some category members increases the expectation to encounter another one of the same kind. This bias is called *assimilation*. For example, participants who were initially presented with relatively tall women (in the induction series) then rated the tall women in the test sample as being taller than the respondents who initially assessed short women (Manis, Biernat and Nelson, 1991). Similar results were found when participants judge the degree of psychopathology of patients (Manis and Paskewitz, 1984; Manis et al., 1988) and judgments of self and others (Biernat, Manis, Kobrynovicz, 1997). Manis and Paskewitz (1984) link expectation with judgment through a Bayesian-like information processing. Assimilation was explained through some sort of readiness of the cognitive system to encounter exemplars that concur with the base-rate expectations.

On the other hand, *contrast* is provoked by the comparison between the stimulus that has to be judged and the stimuli that have been recently judged. Let us consider a case in which respondents were asked to judge the height of a short female and a medium-height male and then to rate the height of target women and men in the test phase. These subjects perceived the female target as taller in comparison with judgments that followed tall female in the induction phase (Manis, Biernat and Nelson, 1991). The authors assume the Range-Frequency Theory explanation for contrast effects in judgment.

The final judgment according to the Two-Path Theory is a result from contrast due to the *R* and *Fr* principles and the expectation-driven assimilation.

An advantage of the Two-Path Theory is that it tries to explain both contrast and assimilation. The theory, however, did not reach detailed explanation of the mechanisms underlying both expectation-driven assimilation and contrast effects.

- **Integration Theory**

The Integration theory of Anderson (1971a) states that whenever the stimulus is presented within a context it's rating would shifts toward the contextual value.

s`= ws + (1-w) I,

where ***s***` is the value of a stimulus in the given context, ***s*** is its value without context, ***I*** is the overall impression (the subjective impression of the stimulus and its context together), and ***w*** is a weighting factor. The context may change either the weighted average (because of the "*I*" parameter) or the importance of some stimulus (*w*). As a result, however, always, assimilation may appear.

The assimilation due to an averaging of the response is demonstrated mainly in the domain of impression-formation (Anderson, 1965; Anderson, 1966; Anderson, 1971a; Anderson, 1971b; Anderson, Lindner and Lopes, 1973; Wyer, 1974).

Change in the w-parameter (the importance of a particular component) was demonstrated in situations in which respondents judge the leader of a group. The assessment of the group as a whole is directly connected with the evaluation of the leader (Anderson et al., 1973). When the leader is a very attractive person, attractiveness of the whole group is judged higher than the attractiveness of the same group but presented with a less attractive leader.

The Integration Theory has one problem on the experimental level. Ss were always told that the stimuli belong to the same entity - one person, one group etc. It is quite probable that not the context but rather the instruction influences judgment. It would be much more convincing the same results (assimilation effect) to be found without such instruction manipulations.

Unlike Integration Theory, the theory ("Change in Meaning" Theory) reviewed in the next section locates contextual effects in judgment on the encoding rather than integration of the information. However, the two theories possess different points of view about the nature of the stimuli. The

Integration Theory accepts that the meaning of the stimuli is context free while the "Change in Meaning" approach - that it has a variety of interpretations. In addition, Wyer (1974) found that when judges are asked to evaluate an adjective that describes a particular person, presented in a group of another two adjectives (adjectives that represents the other two persons), the “halo effect” explanation fits better to the results. In comparison, in a situation in which subjects have to assess one person through three adjectives, the Change in Meaning approach is more adequate (Wyer, 1974).

- **“Change of Meaning” Model**

Change in Meaning model (Wyer, 1974; Srull and Wyer, 1979, Srull and Wyer, 1980) accounts for "how" the ambiguity of the target stimuli contributes to the appearance of the assimilation effect. According to the Change in Meaning theory, the meaning of the stimuli tends to be consistent with the meaning of the most accessible and applicable category in a particular instance of time. In the Change in Meaning approach the focus is mainly on the questions “when” and “how” context affects human judgment. Several experiments demonstrate that ambiguity of the stimulus has an important role for the assimilation to occur. What is more, the strength of the contextual effects increases with the ambiguity of the test stimuli (Wyer, 1974; Srull and Wyer, 1979; Srull and Wyer, 1980). Herr et al., (1982) however, demonstrated that the extremity of the context also may matter. Ambiguous stimuli primed with moderately extreme exemplars assimilated their meaning to the primes, while the same ambiguous stimuli could be judged in contrast of the extreme contextual prime.

With respect to the question concerning the time of contextual influence in judgment, Srull and Wyer (1980) demonstrated that the delay between category activation and acquisition of the target stimulus may decrease assimilation. As the delay between category activation (the context-usually a certain trait category) and the acquisition of the stimulus information (an ambiguous behavior with respect to the primed category) gets longer (none, 24 hr., 1 week), the assimilation decreases but never occurs if the trait category is presented after the target information (Srull and Wyer, 1980). Those results suggest that context exerts its effect on the way in which information is encoded.

Stapel, Koomen and Plight (1996) also demonstrated that characteristics of the context could be important for the subsequent contrast or assimilation in judgment of ambiguous stimuli. Participants were exposed to the photos of the actors with real names of some trait implying sentences like: "Peter peddled even harder as he fell further behind in the race" (persistent); John knew he could handle most problems that would come up" (confident) and so on. The subsequent evaluations of the Donald's ambiguous behavior (with respect to the trait dimensions implied by the first part of the experiment) along the two applicable dimensions[2] on 7-point scales shifted away from the primed examples. Participants rated Donald as more conceited and stubborn rather then persistent and respectively confident. On the contrary, when the primed concept was general behavior labels (abstract concept), the judgment of the target behavior assimilates to the primed one. Furthermore, it turns out that descriptive applicability of the primed category is not so important precondition for occurrence of contextual effect. The activation of a sufficiently broad and inapplicable construct (e.g. good) may prime a descriptively applicable trait (e.g. persistent) and hence, produce assimilation of the ambiguous target (e.g. persistent - stubborn Donald) (Stapel and Koomen, 2000). In other words, general (abstract) labels are more likely to propose an interpretative framework for an ambiguous and new target than narrow and specific categories. Priming with narrow and specific constructs, on the other hand, will result in contrast effects, because they represent sufficiently distinct and relevant information, which is more likely to serve as a subjective standard for comparison in the subsequent judgment.

In conclusion, "Change in Meaning" Theory focuses on the mechanisms underlying judgment of ambiguous stimuli. The theory suggests that judgment is performed within the context of the most active information. Depending on whether the accessible information is general or specific, assimilation or contrast, respectively, may appear. The mechanisms suggested by the "Change in Meaning Theory" are a kind of categorization as far as judged stimuli within this paradigm are mostly ambiguous ones.

The next theory also treats categorization as a source of contextual effects in judgment. Unlike, "Change in Meaning" Theory, however, Inclusion-Exclusion Theory describes both contrast and assimilation in judgment of complex stimuli.

- **Inclusion-Exclusion Theory**

According to the inclusion-exclusion theory (Schwarz and Bless, 1991), the default operation is to include information that comes to mind in the representation of the target. If this happens, assimilation would occur. On the contrary, if the accessible information is excluded from the representation of the target, a contrast is more likely to appear. The factors that determine which information would be included or excluded concern the source (Does the information comes to mind for a wrong reason?), categorization of the most accessible information at a given time (Does the information belong to the target category?), and conversational norms for its use (Am I supposed to use it?).

If the judge knows that some irrelevant factors brought to his/her mind certain information, he/she would exclude it from the representation of the target and hence a contrast would appear. When, however, individuals'

---

[2] Confident-conceited and persistent-stubborn

cognitive capacity is restricted (participants work on another task or are unmotivated), the opposite assimilation effect occurs (Martin et al., 1990). On the contrary, when a person is not aware why a certain category is highly active, he/she would include it in the target's relevant information by default.

Salience of the category boundaries is also among the possible factors that may produce contrast or assimilation effects. Schwarz and Bless (1991) altered the temporally salient boundaries of the target. They asked first-year students to assess their current life-satisfaction after recalling a positive or a negative event from the last two years. Subjects reported higher life-satisfaction when the recall of the positive event preceded the judgment than when a negative recall was first. However, if the instruction pointed that they were not at college two years ago, the result was reversed. Subjects, who recalled a negative event excluded it from the college period and rated their current life-satisfaction higher than those that recalled a positive one.

A category width is another variable that determines the use of the accessible information. Wider target categories allow much more information to be included than the narrow ones. For example, if the task is to assess the trustworthiness of politicians in the United States, the target category includes all the politicians in the United States and one can include any politician from the same country in this category. However, if the question concerns the trustworthiness of the politicians of the Democratic Party, you could include only the Democrats etc.

Conversational norms may also affect judgment. People tend to make their conversations informative and for that reason, they usually avoid providing information that is already "given" (given-new contract). That is why, when unhappily married respondents judged their life-satisfaction, they rated it lower if they were asked to rate their general life-satisfaction before that (Schwarz and Strack, 1991).

To conclude, Inclusion-exclusion Theory explains both contrast and assimilation depending on some constructive process of on-line building of the representation of the target stimulus. These mechanisms, however, were specified on a quite abstract level of description. Much more detailed, in this respect, seem to be the attempt of the Norm Theory reviewed in the next section to describe the process of judgment.

- **Norm Theory**

The Norm Theory proposed by Kahneman and Miller (1986) can hardly be classified as belonging to the social cognition approach or as a theory for judgment. It is a theory that deals with constructing norms for categories. The authors applied the theory to some complex social phenomena as emotional responses, social judgment, and conversations about causes. Although, the Norm Theory does not explain the whole process of judgment - from the stimulus to the rating assigned to it, the theory proposes some plausible mechanisms that underlie the process of judgment of both simple and complex stimuli.

The Norm Theory states that a stimulus including its context and its category name evokes a set of memory representations of objects, events or fragments of them. A probe activates simultaneously and to a different degree a number of elements that could be both retrieved and constructed. Each element could be characterized in terms of features (values of attributes). The evoked set could be described in terms of norms for each of the attributes. For example, if the set could be defined as having attributes X and Y, each feature of every element is described along them through a particular distribution of activation (profile). The sum of the profiles of the elements for X and Y constructs the attributes' norms for the whole set. As Kahneman and Miller (1986) pointed out, these aggregate profiles provide a measure of availability, while the measure of normality could be received through rescaling the profiles according to the ratio of their availability to the maximum of the mode. Evaluation of the stimulus normality is based on the comparison with the evoked norms.

The Norm Theory proposes a general frame for explanation of judgment as well as some references to possible mechanisms and structures (e.g., "spreading associative activation", "exemplar models", "norm construction") but the theory does not specify the connections among them. Still, it predicts that each particular stimulus would be compared with a particular norm, which was constructed depending on the elements that the stimulus retrieved from the memory.

The Norm theory does not specify neither the mechanisms, which underlie the comparison between the probe (stimulus and its context) and the norm nor the process of assigning a particular rating to a particular target stimulus. It also does not explain in details the contextual effects in judgment that we are interested in. However, the theory could explain contextual effects in judgment as result from stimulus comparison with a contextually sensitive norm.

- **Conclusion**

Social Theories outline two important starting points for further discussions. One of them concerns the ambiguity of the stimulus and contextual category, the other – the level of appearance of contextual effects. The type of the stimulus or context seems to be crucial for contextual effects. In social judgment, assimilation is usually demonstrated with ambiguous stimuli or wide contextual categories, while contrast – with unambiguous stimuli. This distinction, however, is not very precise because assimilation effects were also demonstrated with simple and unambiguous stimuli in the field of psychophysical judgments (see section for sequential assimilation in Chapter 1).

The time of contextual influence, i.e. during the encoding or during the integration of information, was also in the focus of social theories. The task (judgment of a group member / trait of a particular person), however, was shown to be able to switch the level of contextual influence on judgment (Wyer, 1974).

In sum, Social Theories explain contextual effects by referring to memory – both memory processes (e.g. encoding, integration, retrieval, accessibility and applicability of the active categories) and its content (e.g., broad or narrow categories).

Although social theories were interested in the mechanisms underlying contextually sensitive judgment, they did not specify them in details. Only the Norm Theory suggests some possible mechanisms for explaining the way stimulus elicits its set for comparison and represents it with a single norm, like - spreading associative activation, norm construction etc. Since the Norm Theory is not a theory explicitly designed for judgment it does not specify the direction of contextual shift depending on the norm.

2.3. Theories of Judgment - Conclusion

To sum up, it seems that psychophysical theories focused more on contrast, while social cognition ones – on assimilation effect in judgment. Although the main difference between these two approaches concerns the complexity of the judged stimuli, there are examples of theories that successfully account for judgment of both simple and complex stimuli, e.g. the Range-Frequency Model. Thus, stimulus complexity may not be so crucial for distinguishing two different approaches to judgment.

Psychophysical and Social Theories are, however, interested in different aspects of judgment. Psychophysical theories describe the outcome of the judgment usually by the means of equations. Social theories focus more on the process of judgment and on the possible mechanisms underlying contextual influence in judgment. The proposed mechanisms, however, were specified at a quite general level, e.g. expectation-driven assimilation (Manis and Paskewitz, 1984), activation of particular information (e.g. Srull and Wyer, 1979, 1980; Stapel et al., 1996), inclusion or exclusion of the target stimulus from the contextual information (Schwarz and Bless, 1991). In contrast to most theories in the field, the ANCHOR Model (Petrov and Anderson, 2000, 2005) and the Norm Theory (Kahneman and Miller, 1986) describe in more details the mechanisms of contextually sensitive judgment. Unfortunately, the ANCHOR Model is very restricted in several aspects, i.e. it focuses only on judgment of simple stimuli on a 9-point scale. Norm Theory, on the other, hand is more general but describes only contextually sensitive construction of the norm. The way this norm may shift judgment was not defined, neither as direction nor as mechanisms.

In general, the mechanisms indicated in the theoretical review (this chapter) seem to rely strongly on memory, e.g.. anchors were kept in memory, norms were constructed from recalled exemplars, category information was activated from particular context. Therefore, the mechanism specified in more details in the field of judgment seems to be connected in one way or another to memory. It could be, however, that context also influences judgment by changing perceptual input as Arieh and Marks (2002) argued. Goldstone (1995) also discusses the possibility context effects to result from perceptually grounded on-line categorization of judged stimuli. Unfortunately, theories explaining the contextual shifts of perceptual input in terms of detailed mechanisms have not yet been proposed.

In addition, both psychophysical and social theories seem to have the same terminological disagreement about the character of context, i.e., whether context is external or internal one. One part of the theories define context, as external one, i.e. stimulus presented at the environment during judgment (e.g., Helson, 1964; Anderson, 1966, 1971b). Second part, seems to share the idea that context is internal, i.e. information retrieved from memory (e.g., Petrov and Anderson, 2000, 2005; Manis and Paskewitz, 1984; Wyer, 1974; Srull and Wyer, 1980; Schwarz and Bless, 1991). A third part, defines context as possessing both external and internal aspects, i.e. context from the environment (e.g. other stimuli and all characteristics of the target stimulus) directs memory retrieval, which elicits the attributes that comprise the constructed norm (Kahneman and Miller, 1986). In the next chapter, the JUDGEMAP-1 (Judgment as mapping) Model will be reviewed. It defines context in a similar way to the Norm Theory, i.e., context as comprising both stimuli from the environment and exemplars retrieved from memory. Since JADGEMAP1 is a computational model, it proposes detailed mechanisms for contrast and assimilation effects in judgment.

CHAPTER 3

JUDGEMAP-1

The empirical and theoretical review outlined several important weaknesses and open issues for further development in the domain of judgments. First, any theory of contextual effects in judgments needs a starting definition of context. This would certainly allow a generalization of the empirical and theoretical findings, and new tests for the proposed theory.

Second, both the theories and the experimental results do not define the factors responsible for contextual effects in judgment. It is difficult to predict when contrast or assimilation would appear in a particular judgment situation. This seems to be the greatest problem in the field, because it poses doubts on the systematic nature of contextual effects. 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. Moreover, if the mechanisms underlying judgment are shared with other cognitive processes and can still produce assimilation and contrast, they can possibly explain also other contextual effects related to other cognitive processes. This will also allow contextually sensitive judgment to be integrated with other cognitive abilities. JUDGEMAP-1 (Kokinov, Hristova, Petkov, 2004), for example, integrates judgment with analogy-making and memory within the cognitive architecture DUAL (Kokinov, 1994a, 1994b).

Unlike, most theories in the field of judgment, JUDGEMAP-1 starts from an explicit definition of context and suggests detailed mechanisms for judgment of both simple and complex stimuli. JUDGEMAP-1 treats context as equivalent to the content of the WM. Thus, context may comprise both perceived elements from the surrounding environment and old instances recalled from LTM. Thus, it brings together the points of view of different researchers viewing context as external or internal. In this respect, WM seems to be a good substitute for the notion of context proposed by the Dynamic Theory of Context (Kokinov, 1995), namely that context could be considered as a “state of the mind” with no clear cut boundaries that dynamically vary in the course of time and comprise all associative relevant elements in memory.

Since, JUDGEMAP-1 is a computational model it proposes concrete mechanisms describing the process of judgment. It is naturally integrated with AMBR (Associative Memory-Based Reasoning) Model (Kokinov, 1998; Kokinov and Petrov, 2001) within the cognitive architecture DUAL (Kokinov, 1994b, 1994c). Both Models rely on memory (for constructing a dynamic set of entities in WM) and both use the same mechanisms of mapping.

This chapter consists of several sections. First the DUAL architecture will be briefly presented, than the JUDGEMAP-1 Model and finally a specific prediction of the JUDGEMAP-1 Model.

3.1. DUAL Cognitive Architecture

The JUDGEMAP-1 Model is based on the mechanisms proposed by the cognitive architecture DUAL (Kokinov, 1994a; 1994b). Knowledge representation and information processing in DUAL are carried out by small entities called DUAL agents. Each Dual agent is a hybrid entity that has symbolic and connectionist aspects. On the symbolic side, each agent "stands for" something and is able to perform certain simple manipulations on symbols. On the connectionist side, it sends activation to and receives activation from its immediate neighbors. Each DUAL-based system consists of a large number of such agents. There is no central executor in the architecture that controls its global operation. Instead, each individual agent is relatively simple and has access only to local information, interacting with a few neighboring agents. The overall behavior of the system emerges out of the collective activity of the whole population. This so-called "society of mind" (Minsky, 1985) provides a substrate for concurrent processing, interaction and emergent computation. Only a small fraction of this large network (LTM), however, may be active at any particular moment. The active subset of long-term memory together with some temporally constructed agents constitute the WM of the architecture. The mechanism of spreading activation plays a key role for controling the size and the content of the WM. There is a treshold that sets the minimum level of activation an agent must have in order to enter the WM. There is also a spontaneous decay factor that pushes the activation level back to zero. As the pattern of activation changes over time, some agents from WM fall back to dormancy, others are activated, etc. Only active agents may perform symbolic computation. Moreover, the speed of these computations depends on the level of activation of the respective agent.

The activation in DUAL is supplied from two constant sorces: Input and Goal nodes. The Input node simulates perceptual input, while the Goal node represents the presure to the system to perform a particulat task. Local computations and spreading activation, which determines the relevance of the various pieces of knowledge, make DUAL a dynamic and context-sensitive model.

3.2. JUDGEMAP-1

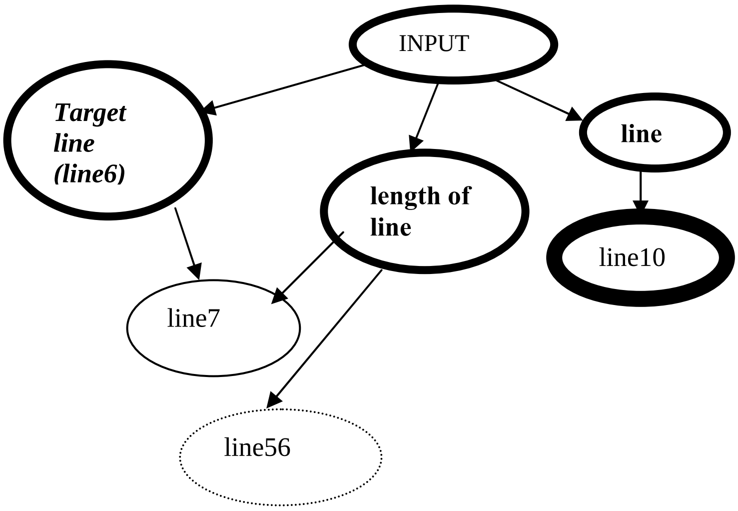
JUDGEMAP-1 views the process of judgment as an *analogical mapping* between a set of stimuli and a set of ratings (i.e. a mapping that will keep the ordering relations of the sets – better stimuli should receive higher ratings). The set of stimuli (called comparison set) includes the target stimulus and possibly some recently judged stimuli, some familiar exemplars of the judged category and some similar to the target, previously encountered stimuli. The set of ratings represents the scale defined by the particular task. As a rule, the stimuli in the first set can vary with time (i.e. the comparison set is dynamic) – some receive additional activation and enter the set, others, on the contrary, lose activation and fade away. Usually all (or most) of the stimuli in the comparison receive ratings but, as far as the task is to judge the target one, only its rating is reported. The content of the set of ratings is constant, because the task requires that (judgment on a particular scale). Nevertheless, some of the ratings can be more active than the others (e.g. favourite numbers or the ratings recently used in previous judgments). The correspondence between the set of stimuli and the scale values could be a partial one but the pressure, which is important for the analogical mapping between the set of objects and the set of ratings, is that objects and relations between them have to be similar[3]. In the framework of judgment, the objects are the stimuli and the ratings, while the relations represent the order of the stimuli with respect to the judged characteristic and respectively, the order of the numbers in the scale. The memorized objects from the comparison set and the rating from the set of the ratings would send markers to a common node, called "element of scale". When markers emitted from the object and the rating cross in this common node, a hypothesis for correspondence is constructed. The activations of the rating, and of the object determine the strength of their connections with the hypothesis and hence its level of activation. Hypotheses that are more active have a bigger chance to win.

[3] JUDGMAP-1 is highly restricted in the application of analogical mapping. It is based on calculations of the relative position of both the object and the rating on their min-to-max scale and generation of hypotheses about correspondence between objects and ratings with similar relative position. The next version of the model – JUDGEMAP-2 has already more sophisticated mechanisms for computing and representing the ordering relations in each of the sets and mapping these ordering relations rather than the objects and ratings directly.

Such hypotheses of correspondence are constructed between most of the elements of the comparison set and the ratings. In this way, a constrain-satisfaction network is established. It is possible, a single stimulus to be linked to several hypothesis representing possible correspondences with several different ratings. Then the hypotheses that connect the same object with several different ratings will be competing and between them inhibitory links will be established. Similarly, if a single rating corresponds to two or more objects these hypotheses will also be competing and therefore connected with inhibitory links. On the other hand, the links between supporting each other hypotheses are exhibitory. Each agent (a rating or a object) checks the activation of the hypotheses that are connected to it. If one of them holds as a leader for a long time, it is promoted to become a winner. The result of the constraint-satisfaction network determines the correspondence between the objects (including the target stimulus) and the ratings.

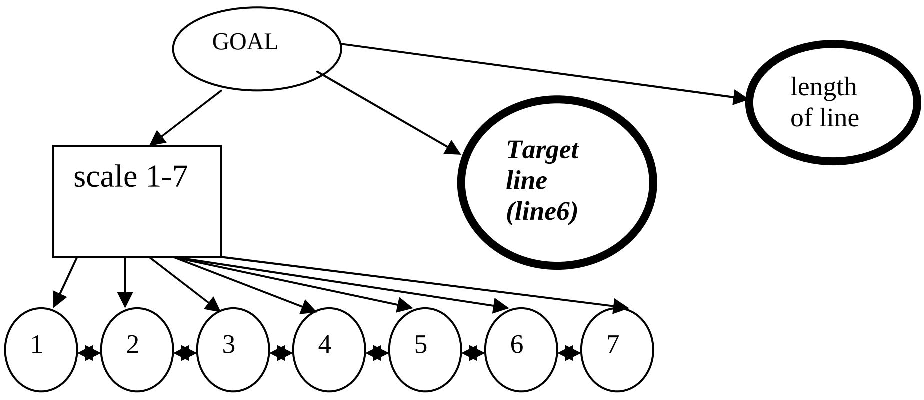
Since JUDGEMAP-1 is a DUAL–based model of judgment it adopts mechanisms from the DUAL architecture. The activation from the GOAL and the INPUT nodes spreads through the neighbouring nodes and combined with the residual activation of ratings and stimuli (if activated from a previous task) determine the content of the WM (Diagram1). If the task is to judge the length of lines, the WM will consists of a number of agents: the representation of the target (activated by the INPUT node), the concepts of line and length of lines (activated by the INPUT node as well), and as a consequence some typical examples of lines seen before (for example, line10), examples of lines similar to the target one (activated by the target stimuli via the mechanism of spreading activation, for example, line7) and recently seen examples of lines, which still have some residual activation (for example, line56).

Diagram1. Activation spreading from the INPUT node in the Dual-based model of judgment. The content of the comparison set is formed.



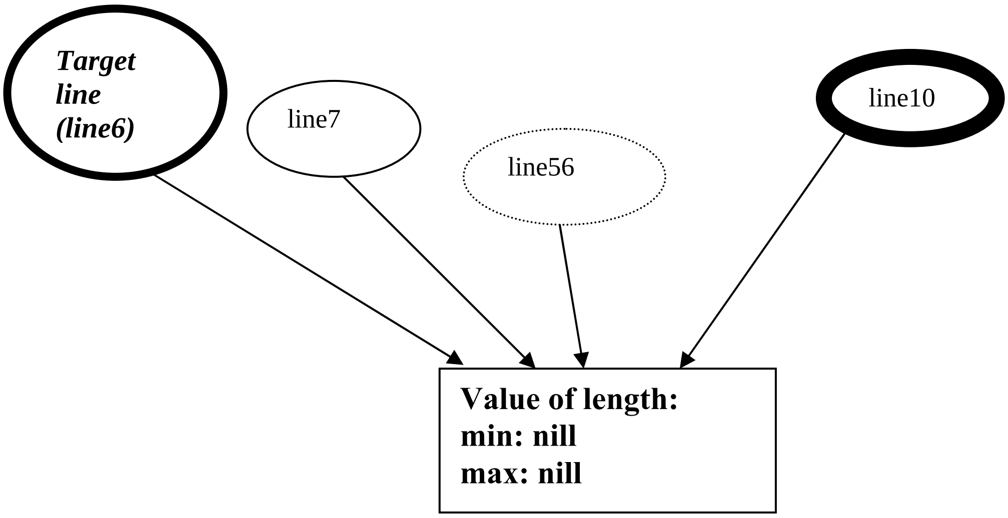
The GOAL node, on the other hand, will activate the target's representation, the concept for length of lines and the scale representation that is required (Diagram 2).

Diagram 2. Activation from the GOAL node in the DUAL-based model of judgment.



Examples of lines that were activated somehow from the input send markers to the node value of length and receive the temporal values of the range (minimum and maximum). Each node calculates its own relative value (rv=(value-min)/(min-max)) and sends a marker to the agent called "a value of length" (Diagram3).

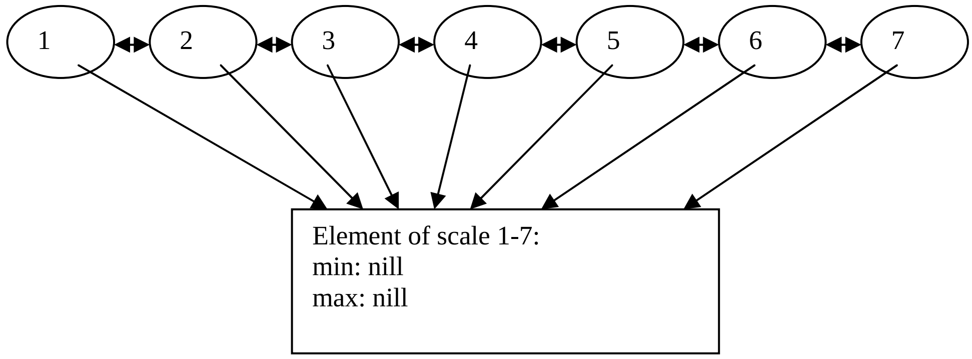
Diagram3. The set of stimuli (the target stimulus and the retrieved stimuli) before each agent to send a marker with its rv to the node "value of length".



The nodes that belong to the set of ratings (the scale activated from the GOAL node) also send markers with their current relative value to the node called "element of scale 1-7" (suppose that the required scale is a 7-point scale) and receive the current min and max. Each node from the coalition of the ratings calculates it's current rv depending on the min and max value in a given moment (rv=(value-min)/(min-max) (Diagram4).

Diagram4. The set of the ratings before each agent to send a marker with its rv to the node "element of scale 1-7".

The agents "value of length" and “elements of scale 1-7” transmit the



markers of the agents in both coalitions (the coalition of the objects and the coalition of the ratings) toward the node “element of scale”. When markers from “elements of scale 1-7” and “value of length” intersect in the node “element of scale”, a hypothesis for correspondence is generated. It connects a particular agent representing a object and a particular agent representing the rating.

This newborn hypothesis would receive activation from three places: the two nodes that correspond to each other and the node “element of scale” (Diagram5). The weight of the first two connections would be calculated based on the difference between the “rv” of the corresponding nodes (the smaller the difference, i.e. the relative positions of the object and of the rating are close to each other, the greater the weight is).

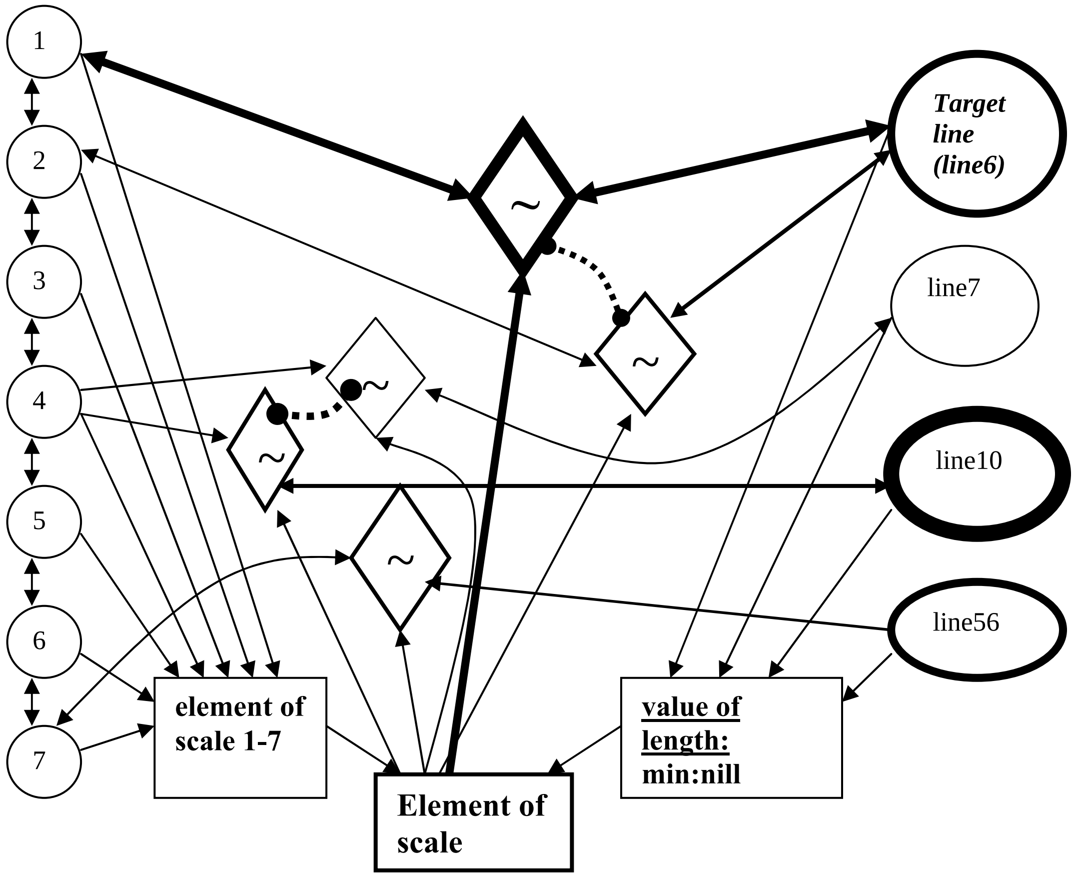


Diagram5. Hypotheses of correspondence. A constraint-satisfaction network.

The hypothesis that receives the highest activation for a sufficiently long period of time will become the winner. If it happens, the particular line instance (the target, for example) will receive the rating to which it is connected by this winner hypothesis.

Assimilation effect arises in the JUDGEMAP-1 model from simultaneous or residual activation of a certain rating on the scale. Recently or simultaneously judged stimuli (context stimuli) activate additionally the particular ratings that they have received and indirectly their neighboring ratings. Suppose that line 10 was the line that was judged in the previous trial and its rating was 4. Suppose then that on the next trial the target line is very small and should be evaluated by 1. Line10 and the rating-4 are still active. The node of the rating-4 would spread activation to its neighboring ratings of 3 and 5. They would be less active than 4 but more active than 2 and 6. Suppose now that the target line is connected with the hypotheses of correspondence to the ratings 1 and 2. Since rating 2 would be more active than the rating 1 (being a closer neighbor of 4) this could help the hypothesis, which connects the target line with the rating 2 to win. Thus, an assimilation effect will arise – the target will receive a rating 2 instead of 1.

Contrast effect is explained in the model by the competition between the hypotheses connecting one and the same rating to two or more lines. This competition may seem unnecessary since there shouldn't be a problem to rate two different lines with the same rating, however, this competition is inherited in the JUDGEMAP-1 model from the AMBR model of analogy-making that requires a one-to-one mapping. Thus, JUDGEMAP-1 makes the pressure two different lines to be rated with two different ratings and therefore makes the pressure each rating to be used not more often than the others. As a direct consequence, we obtain the frequency principle of the Range-Frequency Theory. Now suppose that we have obtained a rating 2 for line 10 and we have now to rate a target whose main hypotheses are those of rating 1 and 2. Since its hypothesis that it can be rated by 2 will have a competitor (the hypothesis that the previous line was rated also 2) this hypothesis will be inhibited by the competitor and is not very likely to win. Therefore, it is more probable to obtain a rating of 1 for this target, i.e. to obtain a contrast effect.

3.3. Predictions of the Model.

3.3.1. How the comparison set is formed?

The JUDGEMAP-1 model postulates that the target stimulus would be included and judged within a dynamically formed set of memories for other objects. This set is formed by several factors such as similarity, recency and familiarity. All these factors contribute to the level of activation of the corresponding memory element in the model. Thus, recently judged objects would have some residual activation and therefore participate in the comparison set. Objects that are more familiar would have stronger links from various other nodes and therefore would be easier to be retrieved (activated) in the comparison set. Finally, objects, which are highly similar to the target, would have a high chance to be activated by the many shared features and therefore will also have greater chances to participate in the comparison set. The process of formation of the comparison set and the process of judgment run in parallel and interact with each other. Moreover, the level of activation of each element of the comparison set would be proportional to the level of its contribution to the result. Thus, if the representation of one particular object is highly activated (e.g., because it is very similar to the target object) then this object will play a more important role in the comparison set than another one. In other words, we may consider the comparison set as a fuzzy set where the degree of membership is computed as the level of activation of the corresponding memory element.

3.3.2. Idea for research on JUDGEMAP-1's predictions

Suppose we have multidimensional objects to be rated on a one-dimensional scale, e.g., "Rate on a 7-point-scale how appropriate this dress is for an official dinner". The model would predict that dresses similar to the target will be retrieved from memory and form the comparison set. These dresses may be similar on various dimensions that are not necessarily relevant for the current judgment, but the very fact of similarity may bring them into WM and make them to participate in the comparison set. Let now take an extreme example. Suppose that we have to rate "how tall this person is". Again, other persons that are similar to that one will tend to form the comparison set. This means that if the person is a lady, predominantly images of other ladies will be retrieved and thus the "tallness" of ladies will be computed based on a different set than for men. This sounds very intuitive. It has, however, further implications: if the target lady is blond, predominantly blond ladies will be retrieved, if the target person is a teacher, predominantly teachers will be retrieved, etc. This is a quite unusual prediction. It says that even irrelevant to the task features may take part and influence the final result based on their contribution to the formation of the comparison set.

The simplest case would be to study what happens if we have two-dimensional objects that are rated along one of their dimensions and we manipulate the other irrelevant dimension. Let us suppose that we have to rate line segments that vary in length and color and we have to rate their length. In this case it is quite clear that the length is the relevant dimension, while the color is irrelevant.

Let the target stimulus be a red line of certain length. In this case, according to the JUDGEMAP-1 Model, we may expect that there will be more red lines in the comparison set (Figure 2) – they will be activated through the RED concept, which is shared with the target. On the other hand, if the target stimulus is a green line of the same length, more green lines will become part of the comparison set (Figure 3). Now, if it happens that the known red lines are longer than the known green lines, then the two target stimuli (differing only in color) will be included in different comparison sets and thus judged differently, i.e., green target would be judged to be longer than red target with the same length.

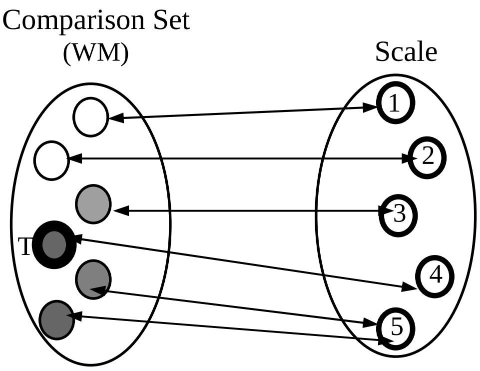


Figure 2. The target stimulus is red and therefore we expect more red exemplars in the comparison set. They happened to be larger in size and thus they compete for the upper part of the scale. In this case the target stimulus (of the same size as in Figure 7) will compete with them and will be mapped onto 4.

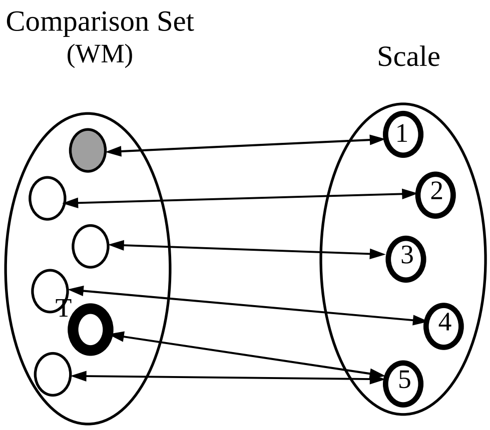


Figure 3. The target stimulus is green and therefore we expect more green exemplars in the comparison set. They happened to be smaller in size and thus they compete for the lower part of the scale. In this case the target stimulus (of the same size as in Figure 6) will compete with them and eventually will be mapped onto 5. In this way we receive an upward shift in the judgment.

This prediction was verified with a simulation experiment with JUDGEMAP-1 that is briefly described in the next section.

- ***Simulation Experiment***

In this simulation, experiment a stimulus set of 56 lines was used. They are all in the long-term memory of the model. The lines differ in length and color. There are 7 different sizes (from 10 units of length to 34 units with increment of 4 units) and two different colors (red and green). Thus in each size group there are 8 lines. The frequency of the red (respectively green) lines varies across the size groups. In size group one (the shortest lines - length 10 units) there are 7 green and 1 red line, in the second shortest group (length 14 units) there are 6 green and 2 red lines, etc. In the largest group size (length of 34 units) there are 7 red lines and one green line. Thus, the green lines were positively skewed while the red lines were negatively skewed.

Each line is represented by a coalition of 5 agents standing for the line itself, for its color, for its length, and for the two relations (color_of and length_of). In addition there are agents representing the numbers from 0 to 8, but only the agents standing for 1 to 7 are instances of “scale element” (i.e., elements of the 7-point scale).

On each run of the program one of these lines was connected to the input list simulating the perception of the target stimulus and simultaneously, the agent standing for “scale_from_1_to_7” was connected to the goal node simulating the instruction “rate the elements on a 7 point scale”.

42 variations of the knowledge base of the system were produced simulating 42 different participants in the experiment. The knowledge bases differ mainly in the associative and instance links among the agents, thus although all our “artificial participants” will know the same lines and the same concepts, they will activate different instances in the comparison set (i.e., in WM).

For each of these knowledge bases we have run two judgment trials: one for a red line of size 22 (middle size) and one for a green line of the same size. Since, the most active exemplars in WM receive their ratings on the scale along with the target stimulus, the ratings of all judged lines within each run of the model was considered in the data analysis.

Simulation Results

7 line lengths averaged the data for each "participant". The results from the simulations are presented in Figure 4. The mean rating of the green lines are in most cases slightly higher than the mean rating of the red lines with the same length.

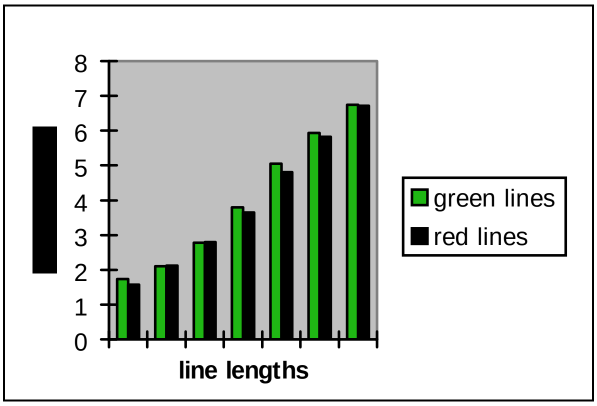


Figure 4. Simulation data. The mean rating of each line with a certain length (1-7) and color (green and red) obtained from all "participants".

Thus the mean ratings of all red lines is 4.012, while the mean rating of all green lines is 4.065, which makes a difference of 0.053 which turns out to be almost significant tested with repeated measurements analysis ($F(1,41)=3.917$, $p=0.055$). The data shows that the possible size of effect of color is very small, but may turn out to be significant. This prediction makes sense: on one hand it is small enough, so that we can ignore it in everyday life and this explains why our intuition says that irrelevant information does not play a role in judgment. On the other hand, the simulation predicts that the irrelevant information does play a role and shifts a bit the evaluation. This means that under specific circumstances this shift could be larger and become significant.

The experimental question that follows from the model prediction is whether the color of the line would or would not matter in this judgment. By analogy with the simulation, the model's prediction was tested experimentally with a set of green and red lines that form respectively, positively and negatively skewed distribution (Kokinov, Hristova, Petkov, 2004).

- *Psychological Experiment (Kokinov, Hristova, Petkov, 2004)*

Participants were asked to rate the length of 14 lines each presented 8 times forming a set of 112 trials. 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. The green and red lines were mixed and presented randomly for judgment on a 7-point scale. Each line was shown horizontally at the center of the computer screen against a gray background. 18 participants took part in this experiment. Since, the green and red lines were mixed together and forming, overall, a uniform distribution of lines (14 line lengths x 8 times), different ratings for the same line depending on its color would be considered as evidence in favor of the JUDGEMAP-1 Model.

Results and discussion

Like in the simulation, the green lines received higher ratings than the red lines with the same absolute length. The repeated measurements analysis showed that the difference (0.046) between the mean judgment of the green lines (4.239) and the mean judgment of the red lines (4.193) is significant ($F(1, 17)=5.966$, $p=0.026$).

Surprisingly enough we obtained a difference (0.046) that is almost the same as the difference we obtained in the simulation (0.053). No tuning of the model was possible since we did not have the experimental data in advance.

Thus, the prediction of the JUDGEMAP-1 model was considered to be experimentally confirmed.

3.3.3. Aims of the present research

Although, the prediction of JUDGEMAP-1 Model for the influence of the irrelevant to the task stimulus dimension was experimentally confirmed several questions remain:

- Whether the effect of the irrelevant dimension was an *accidental* one, since we received a very small thought significant effect of color upon judgment of the line's lengths? Thus replications are needed. Several experiments with similar design and the same experimental question were conducted. If all experiments demonstrate the same effect of the irrelevant dimension we can be positive about its existence.
- Whether the influence of the irrelevant dimension can be increased? What could be the specific circumstances in which this negligible shift in judgment due to the irrelevant to the task dimension(s) may become larger? Each of the experiments reported in the thesis carefully manipulates a particular variable that may potentially increase the effect of the irrelevant stimulus dimension(s). This was considered as the main goal of the presented experimental work, namely, to explore the upper boundary of the effect produced by the irrelevant context.
- Whether the effect of the irrelevant dimension is due to contextually sensitive retrieval of similar to the target stimulus exemplars (i.e. is due to the judgment process itself) or it is rather due to some early low-level perceptual changes in the stimulus representation with respect to the stimulus irrelevant characteristics? This question concerns the mechanisms that may underlie the reported effect of the irrelevant dimension, namely whether spreading activation mechanism (Kokinov, et al., 2004) may account for the presented results or rather the results could be described by means of early "recalibration" of perceptual system's sensitivity (Marks, 1988, 1992, 1994; Marks and Warner, 1991; Arieh and Marks, 2002)

- Whether the effect of the irrelevant dimension increases, decreases or rather is independent of the number of trials? If the effect of the irrelevant dimension is considered like a noise it seems highly probable the effect to disappear with increasing the number of trials. Moreover, the effect of the on-line contextually sensitive categorization that Goldstone (1995) reported may also decrease the impact of the irrelevant dimension in the course of time since assimilation toward perceptually similar and contrast from perceptually dissimilar objects work against the expected contrast effect due to the irrelevant dimension in all presented experiments. It could be, however, that the influence of the irrelevant information increases with time if participants enrich the range and frequency information for the stimuli with the number of trials. Then, it is more likely judges to become more sensitive to the irrelevant stimulus dimension with the number of trials. In order to test this contradictory issue in all experiments presented in the thesis, participants had to judge 3 times the same skewed stimuli set.

The next several chapters present 7 experiments. Chapter 4 describes 2 control experiments, which focus on the difference in length judgment depending on the *Range* and the *Frequency* principles. Although this difference was obtained through manipulation of the relevant to the task dimension it was considered as an empirical maximum effect for the rest 5 experiments. Chapter 5 presents an experiment, which vary both the *R* and the *Fr* distribution of stimuli with respect to their irrelevant to the task dimension to see whether the obtained effect will be increased. Chapter 5 includes 2 experiments that study the effect of several correlated irrelevant dimensions both with simple and complex stimuli again aiming the possible increase of the size of the effect. Chapter 6, presents an experiment on judgment of abstract stimuli, where the effect of the irrelevant dimension (if any) could hardly be described by “recalibration” of perceptual system sensitivity.

All experiments have 2 common expectations. First, the positively skewed with respect to their irrelevant dimension stimuli to receive higher rating than the negatively skewed ones. Second, the impact of the irrelevant dimension was expected to be higher in the middle range (i.e., stimuli with magnitudes close to the mean the whole stimulus set) than in the extreme stimuli, since according to the range principle the smallest and the biggest stimuli should always be evaluated with 1 and 7, respectively.

CHAPTER 4

Exploring Empirically the Maximal Possible Size of the Effect in Ideal Conditions

This chapter is devoted to two experiments that study the maximum size of the shift in segments *length* judgment depending on the *R* and *Fr* principles (Parducci, 1965, 1971, 1974). Both experiments manipulate the two principles with respect to the relevant dimension, i.e. the length of the stimulus lines. The first experiment (control 1) explores difference between judgments of length of lines that differ in their length range and frequency of distribution. The second experiment (control 2) is focused on the difference between length judgments of positively and negatively skewed line sets.

These experiments rely on manipulation of the *R* and *Fr* with respect to the relevant to the task length of the lines. Therefore, the observed effects will be considered the maximum possible shift in judgment of line length due to the irrelevant dimension.

4.1. *Control 1 Experiment:* Judgment of lines that differ in length range and frequency distribution

Control 1 explores the situation where both ***R*** and ***Fr*** "work" together, i.e., they push judgments in the same direction. This has to result in the highest displacement in judgment due to the impact of the whole stimulus distribution (Parducci, 1965, 1971, 1974). Control 1 aims at finding empirically the size of this effect in judgment of lines.

Two stimulus sets were designed. One of them included relatively short and positively skewed lines. The other included relatively long and negatively skewed lines. These two sets of stimuli were given to two different groups of participants for judgment on a 7-point scale. Because the stimuli in both sets differed in ***R*,** the impact of the context generated from the whole stimulus set was measured only on the 8 overlapping by length lines.

4.1.1. Method

- *Design*

A between subject design was used. Independent variables was the group that differentiated two target sets of lines, i.e., relatively short positively skewed stimuli and relatively long negatively skewed stimuli. The dependent variable was ratings on a 7-point scale for each of the 8 common lines received from each group.

- *Stimuli*

14 black lines with different lengths were designed. The smallest one was 12 pixels long, the longest one was 727 pixels long and the increment was 55 pixels. The first 11 lengths of lines were included into positively skewed set of stimuli, i.e., shorter lines were presented more frequently than longer lines. Longer lines (the last 11 lengths), on the contrary, were presented more frequently in the negatively skewed set. The presentation frequency of the lines is presented in Table 5.

Table 5. Presentation frequency of the lines in Control 1 depending on the group.

lines	***Length in pixels***	presentation frequency in group 1	presentation frequency in group 2
1	12	16	-
2	67	16	-
3	122	16	-
4	177	14	2
5	232	12	4
6	287	10	6
7	342	8	8
8	397	8	8
9	452	6	10
10	507	4	12
11	562	2	14
12	617	-	16
13	672	-	16
14	727	-	16

- *Procedure*

Lines were presented randomly and horizontally against a gray background at the center of the 17-inch PC monitor. The lines appeared one by one on the screen and stay until participants judge them. The experimenter registered participant's rating and changed the slide manually. The experiment took about 10 minutes.

Participants were instructed to judge the length of each line presented on the screen on a 7-point scale: where 1-"it is not long at all" and 7-"it is very long".

- *Participants*

13 students (10 females and 3 males) from the introductory classes in psychology at New Bulgarian University participated in the experiment. They received credits for participation. Six students participated in Group 1 and seven in Group 2.

4.1.2. Results and discussion

Data from each group was averaged by item (8 common lines) over participants. Overall, the 8 common for both sets lines were judged higher (mean=5.00; SD[4]=0.94) when participated in the set of relatively short positively skewed lines than within the set of relatively long and negatively skewed lines (mean=4.16, SD=0.96) (Figure 5).

[4] SD-Standard Deviation

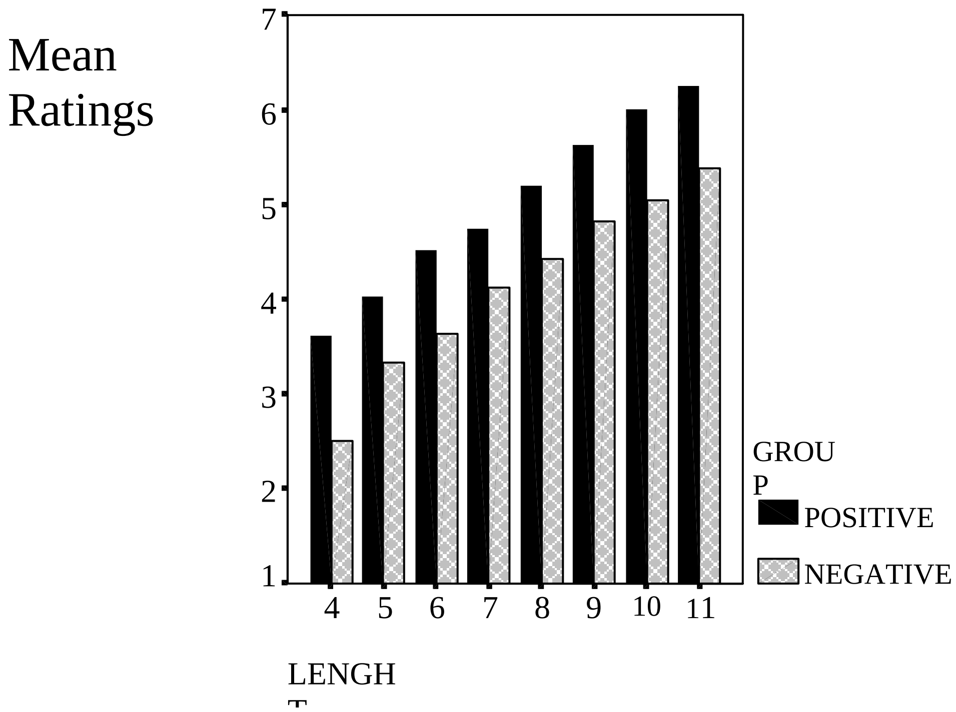


Figure 5. Mean ratings of the common lines for each group. Black bars represents the mean ratings in the positively skewed set of short lines, while the gray textured bars, stand for the mean ratings in the negatively skewed set of long lines.

The difference between the 8 mean ratings received from both experimental groups was 0.84 (Table 6). It turns to be a significant difference tested with One-Sample T-test: t (7)=15.43; p=0.000. These results confirm the theoretical predictions of the Range-Frequency Theory, i.e., ratings of the same lines were shifted away from regions with high stimulus frequency and depending on the stimulus range.

Table 6. Mean, Standard Error, Standard Deviation and 95% confidence interval for the difference between judgments of short positively skewed and long negatively skewed lines.

	Mean	Std. Error	Stand. Deviation	95% Confidence Interval of the difference	
Diff	0.84	0.054	0.154	0.711	0.968

The correlation between ratings and line length was 0.980 (p=0.01). Thus, contrast due to the R and Fr variations seems to exist on the background of high judgment precision (i.e. correlation between the judgment and the length of the lines).

4.2. *Control 2 Experiment:* Judgment of positively and negatively skewed lines

The aim of control 2 was to find the empirical maximum of the difference in ratings of positively and negatively skewed distributions of lines keeping the range equal in both distributions. In the positively skewed set short lines were presented more frequent than long lines. On the contrary, in the negatively skewed set the long lines were presented more frequent than the short ones. These two unevenly distributed sets of stimuli were presented to two different groups of participants.

4.2.1. Method

- *Design*

The experiment had a between subject design, where the independent variable was the group. It was varied on 2 levels depending on the stimulus distribution presented for judgment to the participants: positively and negatively skewed lines. The dependent variable was the rating of each line received from each group.

- *Stimuli*

A set of 14 black lines with different length was constructed. A shortest line was 180 pixels, the longest one was 505 pixels, and the increment was 25 pixels. The frequency of the line’s presentation is described in Table 7.

Table7. Presentation frequency of lines in Control 2 for each group.

lines	***Length in pixels***	Positively skewed distribution (group 1)	Negatively skewed distribution (group 2)
1	180	14	2
2	205	14	2
3	230	12	4
4	255	12	4
5	280	10	6
6	305	10	6
7	330	8	8
8	355	8	8
9	380	6	10
10	405	6	10
11	430	4	12
12	455	4	12
13	480	2	14
14	505	2	14

- **Procedure**

Each line was presented horizontally against a light gray background at a random position on the screen. The lines were randomly selected from the set of positively or negatively skewed stimuli.

Participants were instructed to judge the length of each line on a scale from 1 to 7, giving their rating through the keyboard. The slide was changed when participant pushed the respective button. The experiment lasts around 10 minutes.

- *Participants*

32 (24 females and 8 males) students from different departments of New Bulgarian University took part in the experiment. 15 students participated in group 1 and 17 students in group 2. Participant's age varied from 19-46 years. Part of them was enrolled for course credits, others- were paid (1 lv.) for their participation.

4.2.2. Results and Discussion

Data from each group was averaged by item (14 line lengths), i.e., 14 mean ratings for the positively skewed set and 14 mean ratings for the negatively skewed set. The mean of these 14 differences was 0, 73 (Table 8). T-test showed that the two groups were significantly different from each other: *t* (13) =11.816, *p*=0.000. Results were congruent with the predictions of Range-Frequency Theory that judgments of the same stimuli depend on the skew of the whole stimulus set. Although equal in length, positively skewed lines were rated higher (Mean=5.18, SD=1.34) than negatively skewed one (Mean=4.45, SD=1.43). Rating profile for each line in each skew condition is presented in Figure 6.

Table 8. Mean, Standard Error, Standard Deviation and 95% confidence interval for the difference between judgments of short positively skewed and long negatively skewed lines.

	Mean	Std. Error	Stand. Deviation	95% Confidence Interval of the difference	
Diff	0.73	0.062	0.231	0.597	0.864

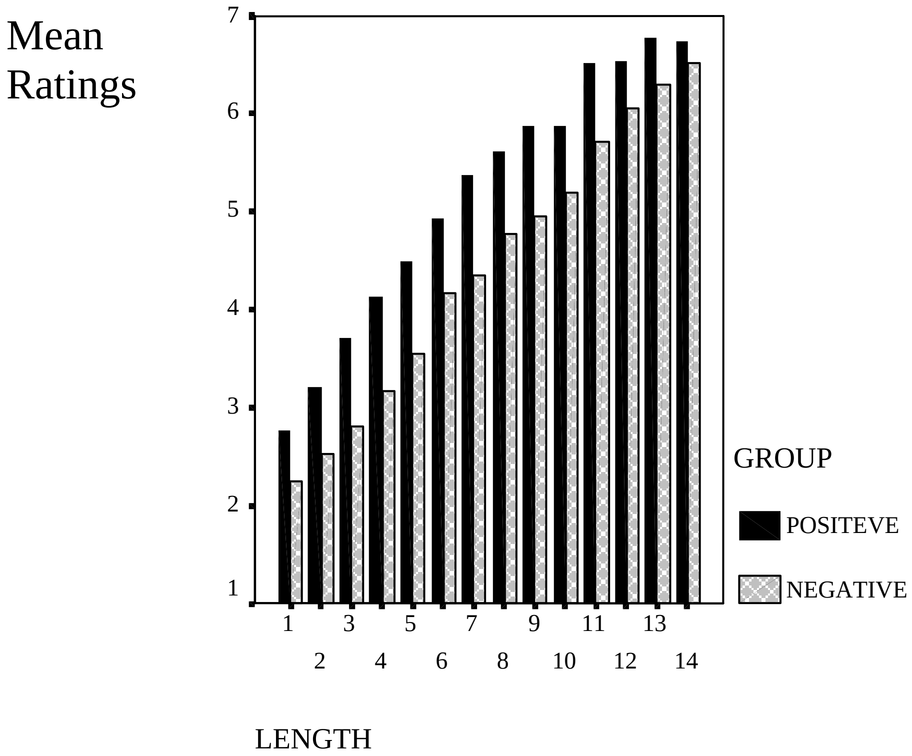


Figure 6. Mean ratings of the lines for each group. Black bars represents the mean ratings in the positively skewed set of short lines, while the grey textured bars, stand for the mean ratings in the negatively skewed set of long lines.

This experiment, like the previous one (control 1) also indicates significant correlation between ratings and length of the lines (r=0.954, p=0.01).

Results from both control experiments outline the maximum difference due to the skew of the stimuli in judgment of line length. The difference, however, in judgment of positively and negatively skewed stimuli is calculated based on the responses of different experimental groups, i.e. one group judge positively skewed stimuli another negatively skewed stimuli. Experiments from the next 3 chapters rely on within subject design, where the same participants judge the lines that comprise both positively and negatively skewed set depending on their irrelevant to the task color. Thus, the difference in judgment of line length depending on the *R* and *Fr* could not be expected to be as high as in the control experiments. The obtained empirical differences outline the upper boundary of the effect rather than the expected size of the effect. Moreover, JUDGEMAP-1 predicts that the comparison set will be formed dynamically for each stimulus separately and the color of the lines will take part in this process of comparison set formation as well as many other factors (e.g., recently judged stimuli, familiar exemplars of the target category). Therefore, JUDGEMAP-1 states that the target lines will retrieve in memory mixed sets of lines with respect to their color (i.e., the comparison set will contain both green and red lines but the green lines will be more when a green target is evaluated and the red lines will be more when a red target is evaluated). Thus, the mixed comparison set will be less positively or negatively skewed than the given stimulus set. Therefore, the expected difference would be much less that 0.73.

CHAPTER 5
Experiment 1: Judgment of lines that differ both in Range and in Frequency with respect to their color

Experiment 1 studies the impact of an irrelevant dimension on judgment of lines length. The lines differ both in their R and in Fr of presentation. Lines could be differentiated into two sets depending on their color, which was the irrelevant dimension in this experiment. Lines with color *P* were positively skewed, while lines with color *N* were negatively skewed one. Moreover, both sets differ not only in skew but also in range, i.e. positively skewed lines were relatively short one, while negatively skewed - relatively long ones. Thus, both contextual forces (the ***R*** and the ***Fr*** principles) were taken into account in this experiment.

R and the ***Fr***, however, were coded through *the irrelevant to the task dimension*, i.e. the color of the lines. Hence, the influence of both principles would be possible *only if participants consider stimulus color* in judgment of the length of the lines. Otherwise, judgment of the same line would not depend on its color but rather on the characteristics of the whole set which includes 14 uniformly distributed lines. Hence, the same line should receive the same rating independently of its color.

The color was either green or red forming a set of positively or negatively skewed lines that differ in range. The experimental procedure was counterbalanced. One part of participants judged a set of short green lines and long red lines. Another part judged a set of short red lines and long green lines. The impact of the irrelevant dimension was measured on the eight common lines, which belong to the range of both positively and negatively skewed lines. Two middle length lines were expected to be more sensitive to the experimental manipulation than the rest 6 common lines, because of the ***Fr*** principle. It states that shift in judgments due to the skew of the stimulus distribution is greater for the middle stimuli than for the end ones. This effect of context is usually called a nonlinear one.

The number of trials was also considered. Participants rated twice the same mixed set of randomly presented green and red lines.

5.1. Method

- *Design*

The group was a between subjects factor - the color of the relatively short positively skewed lines and relatively long negatively skewed lines depend on the group.

The within-subjects factors color of the lines and number of stimulus presentation were varied on two levels: color - red and green; number of presentation - first 112 trials and second 112 trials.

The dependent variable was rating of the eight identical in length but different in color lines.

- *Stimuli*

The same 14 lines used in Control 1 (line length varied from 12 to 727 pixels, with the increment of 55 pixels) were presented 16 times each but with different color. The lines with *color N* formed a negatively skewed distribution, which range was between the first 11 lengths of the stimulus lines. The lines with *color P*, on the opposite, formed a positively skewed distribution, which range falls between the lengths 4 and 14. Stimulus distribution is presented in Table 9. The skew and color were counterbalanced between groups, i.e., positively skewed red and negatively skewed green lines were judged in the first group, while in the second group, on the opposite, the color of the positively skewed lines was green, while the color of the negatively skewed lines - red.

Table 9. Frequency and Color of the stimulus lines within a block of 112 trials.

lines	***Length in pixels***	***Presentation of lines with color P***	***Presentation of lines with color N***
1	12	8	-
2	67	8	-
3	122	8	-
4	177	7	1
5	232	6	2
6	287	5	3
7	342	4	4
8	397	4	4
9	452	3	5
10	507	2	6
11	562	1	7
12	617	-	8
13	672	-	8
14	727	-	8

- *Procedure*

Each line was horizontally presented 16 times in the center of the 17-inch PC monitor against a gray background. Each line was presented 8 times within each block of 112 trials. The frequency and color of the lines distribution is shown in Table 9. Participants rated twice the same mixed set of randomly presented green and red lines, forming a set of 224 trials.

Participants were instructed to judge the length of each line presented on the screen on a seven point scale, where 1-is not long at all and 7-it is very long.

The experimenter registered participants' ratings for each line and then changed the slides manually. Experiment lasted about 20 minutes.

- *Participants*

83 (49 female and 34 male) students from the introductory classes in psychology at New Bulgarian University participated in order to satisfy a course requirement. In Group1 participated 42 students in group 2 - 41. All participants denied knowing to be color-blind.

5.2. Results and Discussion

All the data was averaged by item (8 line lengths) for each participant. Repeated Measurement ANOVA show a non-significant influence of the between-subject factor group: F (1, 81) = 0.534, p= 0.467. The original (red or green) color of the lines was recoded in *color P* and *color N* depending on the stimulus skew. The *color of the positively skewed stimuli* (the red color of the stimuli from the first group and the green ones from the second group) *was recoded into color P*. The color of the *negatively skewed set of lines* (the green color of the stimuli from the first group and the red ones from the second group) *was recoded as color N.*

The impact of the irrelevant dimension was examined first for 2 middle-length lines (lines 7 and 8) common for both distributions. The mean ratings and Standard Errors for lines presented with *color P* and *color N* are shown in Table 10 and are visualized in Figure 7.

Table 10. Mean ratings and Standard Errors of middle lines (lines 7 and 8) for color *P*, color *N*.

color	**Mean**	Std. Error
P (positive skew)	4.45	0.047
N (negative skew)	4.41	0.050

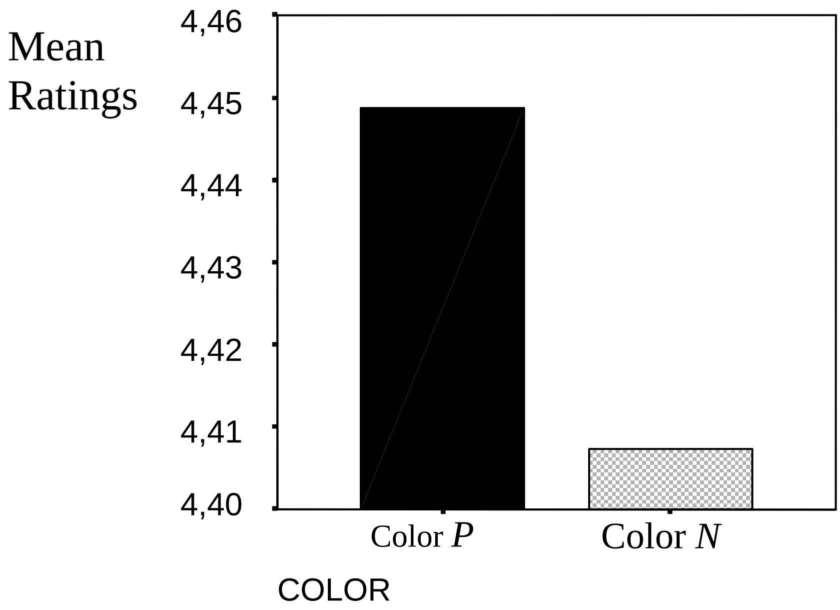


Figure 7. Mean ratings of the middle length lines for each color.

The difference in the mean ratings of the middle-length lines was 0.04. This difference turns to be significant tested with the Repeated Measurement statistics: F (1, 82) = 5.542, p=0.021. The effect size (ES) is 0.063. As was expected, positively skewed middle-length lines were rated higher than negatively skewed middle-length lines despite their equal length. Thus, the results could be considered as confirming the influence of the irrelevant dimension. The effect of the within-subject factor “number of trials” was non-significant: F (1,82)=0.193, p=0.662. The interaction between the factors for the middle-length lines was also non-significant.

Contrary to the expectations, however, the difference between judgments of the eight common lines depending on their color was not significant tested with the repeated measurement analysis (F (1,82) = 1.098, p=0.298). The effect of within-subject factor number of trials was also non-significant (F (1, 82)=1.001, p=0.320) but the interaction between number of trials (first and second 112 trials) and color reached significance (F (1, 82) =3.090, p=0.083). The mean ratings and Standard Errors for lines with *color P* (i.e., short positively skewed lines) and lines with *color N* (i.e., long negatively skewed lines) for each block of trials is presented at Table 11.

Table 11. Mean and Standard Error of ratings of length of all lines for line's color and number of stimulus presentation.

Color	**Number of trials**	**Mean judgment**	Std. Error
P **(**relatively short positively skewed lines)	First 112 trials	4.40	0.048
	Second 112 trials	4.39	0,043
N **(**relatively long negatively skewed lines)	First 112 trials	4.41	0.046
	Second 112 trials	4.36	0.044

Repeated measurement analysis on ratings from the second 112 trials reveals a significant main effect of color (F (1, 82) =6.021, p=0.016, ES=0.069). The appearance of the expected effect of color during the second block of presentation contradicts the idea that contextual effects may disappear with time. Therefore, they cannot be considered noise during the initial calibration process.

Like in Control 1 the correlation between ratings and line lengths was significant: r= 0.93, p=0.01. It seems that, the observed effect of the irrelevant to the task dimension did not affect judgment precision. Thus, it can be considered that either participants were very precise in judgment of the line length or the impact of the irrelevant dimension was not strong enough to reduce this judgment precision.

Experiment 1 demonstrated the effect of irrelevant dimension on judgments of the middle lines and also on judgment of all lines during the second 112 trials. The judgments of the lines were displaced by the context originated from the irrelevant to the task stimulus dimension. The observed contrast effect was congruent with the experimental manipulations and the data obtained with similar methodology (Kokinov et al., 2004). The same line was judged to be longer when its color was the same as that of short positively skewed lines set than when its color was like the one of the long negatively skewed line set.

The reported shift in judgments due to the irrelevant dimension of the stimuli was further enough from the empirical maximum of 0,84 outlined by Control 1, but JUDGEMAP-1 did not predict that the effect of irrelevant dimension would be comparable to this empirical maximum. Moreover, the shift in judgments was not substantially higher compared with the difference of 0,046 measured in the previous experiment (Kokinov et al., 2004). Therefore, addition of the R principle was not enough for enlarging the effect of the irrelevant to the task dimension. The effect increased with the number of trials which may indicate that participants became more sensitive to the irrelevant information at the end of the experiment rather than at its beginning.

CHAPTER 6
Experiment 2: Judgment under "perceptual limitations"

Stimulus lines were projected for a *short time* (100ms) that allows participants only to detect stimulus but not to compare it unrestrictedly to the rest of the surrounding sources of useful information. This peculiarity of the experimental procedure seemed enough for causing troubles on judgment process. Since contextual effects in judgment could be considered as mistakes, the contextual effect due to the irrelevant dimension may increase under these circumstances. It was assumed that presentation limitations would make judgment harder and would cause more mistakes and smaller precision in judgment. The restricted external sources of relational information and the possibility for more mistakes seem to enable the possible influence of irrelevant information upon retrieval of similar elements in the WM.

Participants were asked to judge the length of lines that appear always horizontally, but in random positions on the screen. Each line was projected for a very short time on the computer screen- for only 100ms. The subsequent answer did not require a prompt, rather the computer "wait" for the participants' answers. The Ss were instructed to press the button when they are sure what rating the target line deserves.

6.1. Method

- *Design*

The within subject independent variables were color (varying at 2 levels) of the lines and number of trials (first, second, and third 112 presentations). The experimental design was counterbalanced in order the positively and the negatively skewed stimuli to be presented either in green or in red. In the first experimental group, the green lines were positively skewed, while red lines form negatively skewed distribution. In the second experimental group, red lines were positively skewed, while green lines were negatively skewed. The dependent variables were the mean rating of line lengths on a 7-point-scale.

- *Stimuli*

The same 14 lines as in Control 2 (the length varied from 180 pixels to 505 pixels with 25 pixels increment) were presented 8 times each forming a basic set of 112 trials. Participants rated 3 times the same set of randomly presented 112 trials. Thus, experiment comprises 336 trials (112 trials for a block $_x$ 3). Each line was presented either in red or in green. The frequency distribution of lines in the first experimental group, for one block of 112 trials is presented in Table 12. In second experimental group the presentation of lines was just on the opposite, i.e., red lines formed positively skewed distribution (include relatively short lines) and green lines formed negatively skewed one (include relatively long lines).

Table 12. Frequency and color of the lines for a block of 112 trials presented to group 1, where green lines were positively skewed and red lines were negatively skewed one.

Lines	***Length in pixels***	Number of the green lines (Positively skewed distribution)	Number of the red lines (Negatively skewed distribution)
1	180	7	1
2	205	7	1
3	230	6	2
4	255	6	2
5	280	5	3
6	305	5	3
7	330	4	4
8	355	4	4
9	380	3	5
10	405	3	5
11	430	2	6
12	455	2	6
13	480	1	7
14	505	1	7

- *Procedure*

Each line was presented horizontally against a gray background in a *random position* on the screen for *100 ms.* Participants were instructed to press a button from 1 to 7 on the keyboard whenever they are sure on their rating. They were also asked to leave their hand on the desk after each pressing of the key in order to be sure that the time for reaching each button on a keyboard is relatively equal. When participant pressed the button corresponding to his/hers answer the next line appeared on the screen. The lines were presented within three blocks with the same range and frequency distribution like the one presented at Table 12.

The experiment was conducted in sound-attenuated room and lasts around 15 minutes.

Participants: 31 students (17 female and 14 male) from New Bulgarian University participated in the experiment. Participants' age varied between 19 and 31 years. Part of them participates in order to satisfy a course requirement, others were paid 1 lv. In group 1 participated 16 students, in the second group – 15 students.

6.2. Results and Discussion

Data was averaged by item (14 lengths). Each participant has 28 mean judgments (14 lines*2 colors). Color and number of trials were analyzed as a within-subject factor, while group was a between subject factor. Repeated Measurement Analyses shown a non-significant main effect of group: F (1, 30) = 0.215, p=0.646. Thus for plainness, I will use *color P* as indicating positively skewed lines and *color N* – negatively skewed lines rather than the particular red or green color of the lines used during the experiment.

The main effect of the irrelevant dimension (color *P* vs. color *N*) on rating of the middle lines was significant estimated with the Repeated Measurement Analysis: F (1, 30) =4.400, p=0.045, ES=0.132. The difference between mean judgment of positively skewed lines (5.01) and mean judgment of negatively skewed lines (4.92) was 0.09 (Table 13). Positively skewed middle-length lines were rated higher than negatively skewed middle-length lines despite they were equal in length (Figure 8).

Table 13. Mean, Standard Error of middle lengths ratings (lines 7 and 8). *Color P* stands for positively skewed stimuli, *color N* – for negatively skewed ones.

Color	**Mean ratings**	Std. Error
color P (positively skewed lines)	5.01	0.088
color N (negatively skewed lines)	4.92	0.088

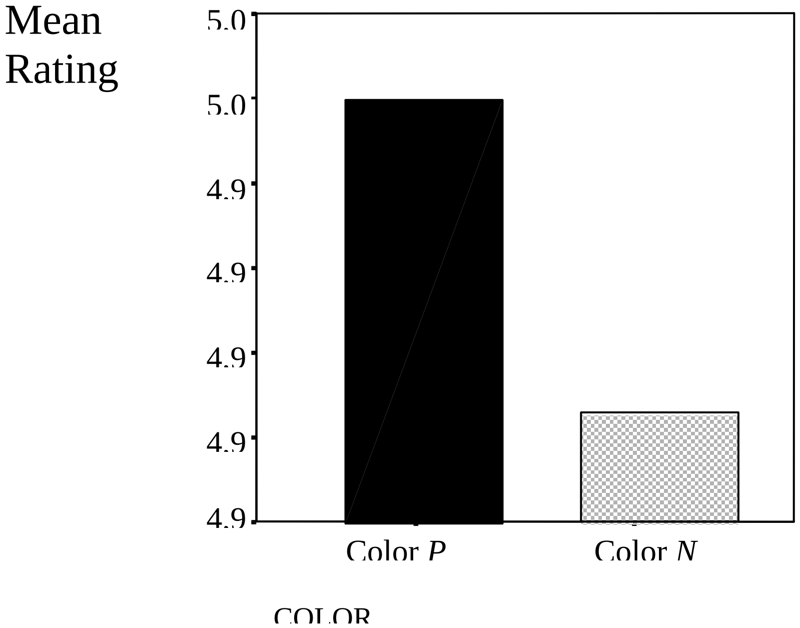


Figure 8. Mean ratings of the middle line lengths for each color. Black bar stands for ratings of the positively skewed lines with respect to their color, while gray textured bar – for negatively skewed with respect to color lines.

Repeated Measurement Analysis, however, did not show a significant main effect of color and of number of trials on the mean ratings of all 14 lines (color: $F(1, 30) = 0.070$, $p = 0.793$; number of trials: $F(2,30) = 1.354$, $p = 0.266$). The effect of the skewed with respect to lines' color context, appeared not strong enough to transfer from the middle toward all lengths of the target lines.

Like in the previous experiments, the correlation between ratings and line length was significant (*r*= 0.8, p=0.01).

The effect of the irrelevant dimension differs in decimals from contextual effects reported in the previous experiments (i.e., control 2 and experiment 1). The effect of color did not increase under the presentation limitations imposed on participants in this experiment. It could be, however, that participants got the regularity in stimulus presentation. Each line is presented immediately after the answer was given by the push of a corresponding key. If participants got this rhythm, they may adjust their behavior in order to have enough time for the stimulus perception and judgment.

CHAPTER 7
Two or More Correlated Irrelevant Dimensions

Experiments reported in this section explore the possibility of enlarging the impact of the irrelevant information by adding several correlated irrelevant dimensions to the target stimuli. This idea closely represents a prediction of JUDGEMAP-1 model, namely that the activation spreading along several dimensions that feed the same information raise the possibility this particular information to be elicited in the WM compared with the situation where only one dimension activate the same information. Thus, the most active element in WM is the element that shares nearly all features of the target stimulus. Hence, if several irrelevant dimensions correlate the irrelevant information would have higher chances to direct memory retrieval. As a result, the context understood like the content of the WM would include more elements that share the same irrelevant to the task information.

Experiment 3 used judgment of line length accompanied always by a particular sound associated with the lines' color. Experiment 4 studies the impact of several correlated dimensions on judgment of more complex stimuli than lines, namely, of the size of rabbits.

7.1. *Experiment 3:* Color and Sound

The goal of this experiment was to explore the situation in which stimuli could be characterized as sharing one relevant to the task characteristic and several irrelevant. In DUAL, the activation of particular agents in the LTM should increase if the agents receive activation simultaneously from several nodes. To test this possibility, positively skewed lines were joined with a particular sound ("laser") and negatively skewed lines - with another sound ("whoosh"). Thus positively skewed exemplars should receive greater activation when a line from the same set is judged in comparison with a situation in which the lines differ only by their color. This manipulation may help to elicit with higher probability the exemplars that belong to the target category. Then if the content of WM depends on irrelevant information and retrieved exemplars are used for the consequent judgment, the effect of irrelevant dimensions should be greater.

7.1.1 Method

- *Design:*

The irrelevant dimension (color-sound variable) was a within-subject independent variable that was varied in two levels: red-whoosh and green-laser. Another within-subject independent variable was the number of trials, which was varied on 3 levels. The color-sound variable was counterbalanced across groups. Color of the positively skewed lines was green in the first group and red in the second group. The color of the negatively skewed lines was red in the first group and green in the second group.

The dependent variable was the 7-point rating of each line length.

- **Stimuli**

The same 14 lines used in the experiment 2 were presented with the same frequency depending on their irrelevant characteristics within each block of 112 trials: lines with lengths 1 and 2 were presented 7 times in green to the accompaniment of the “laser” sound, lines with lengths 3 and 4 - 6 times and so on. Participants rated 3 times the same set of 112 randomly presented trials.

- **Procedure**

Procedure was similar to the experiment 2 described in the previous section, but without any time limitations for presentation of lines. Each line appears simultaneously with a particular sound - green line with “laser” sound, red lines with “whoosh” sound. The lines were randomly presented within 3 sets that share the same stimulus range and frequency forming a sequence of 336 trials (112 trials $_x$ 3 times).

- **Participants**

29 students (15 female and 14 male) from New Bulgarian University distributed in two groups (group 1 -13 students, group 2- 16 students) take part in this experiment. The age of the participants ranged between 18 and 35 years. Part of them participates in order to satisfy a course requirement, others were paid 1 lv.

7.1.2. Results and Discussion

Data was averaged by item (14 lines). Repeated measurement analysis show a non-significant effect of group: F (1, 28) =1.363, p=0.253. Thus, the original red or green color of the lines was recoded into *color P* and *color N* with respect to the skew indicated by the irrelevant to the task color of the lines. *Color P* stands for positively skewed green lines associated with the “laser” sound at group 1 and for positively skewed red lines associated with the “whoosh” sound at group 2. *Color N*, represents negatively skewed red lines associated with the “whoosh” sound at group 1 and negatively skewed green lines associated with the “laser” sound at group 2.

Repeated Measurement Analysis showed a non-significant main effect of color (*F* (1, 28) = 1.929, *p*=0.176) but significant interaction between the number of trials and the color of the lines (F (2, 28) = 4.944, p=0.011) for the middle-length lines. The expected contrast due to the irrelevant dimension appeared during the third 112 presentations (Figure 9).

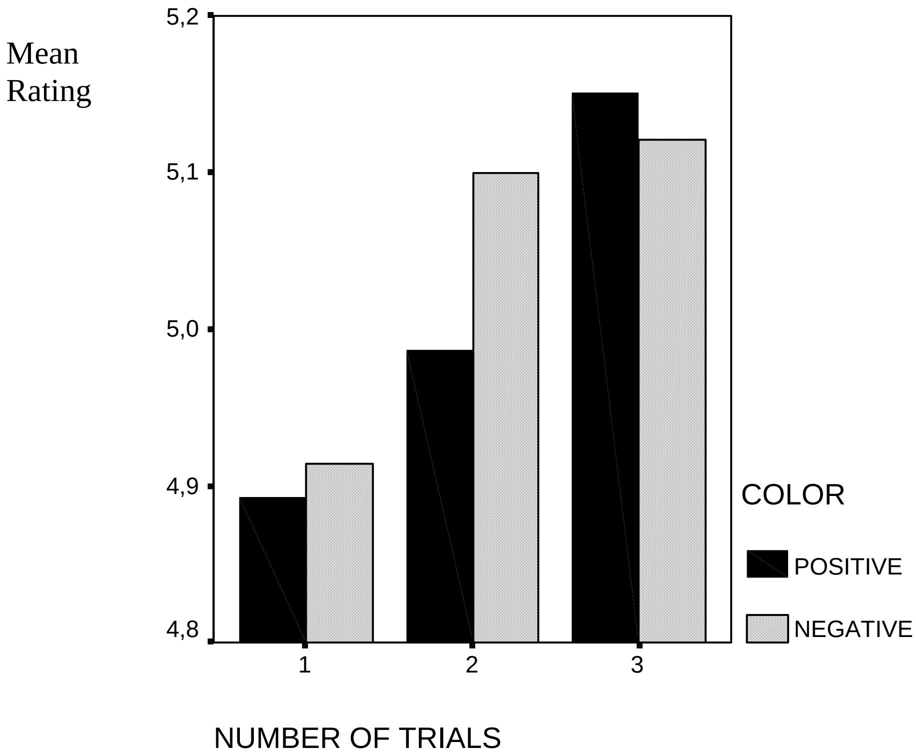


Figure 9.Mean ratings of the middle length lines for each color and number of trials, i.e., 1 stands for the firs 112 trials, 2-for the second 112 trials, 3 – for the third 112 trials.

During the first and second 112 presentations negatively skewed lines were judged higher than positively skewed lines, while this tendency was turned around during the last 112 presentation and the expected contrast appeared (i.e., positively skewed lines were started to be judged higher than negatively skewed ones with the same length) (Table 14). This tendency, however, observed at the end of the experiment was not strong enough to be estimated as significant from the Repeated Measurement Analysis: F (1,28)=0.418, p=0.523.

	Color *P*		Color *N*		Difference
	mean	Std. Error	mean	Std. Error	Color *P*-color *N*
First 112 trials	4.862	0.105	4.914	0.105	-0.052
Second 112 trials	4.987	0.116	5.099	0.108	-0.112
Third 112 trials	5.151	0.105	5.121	0.108	0.030

Table 14. Mean, Standard Error of middle lengths ratings (lines 7 and 8) for each color and number of trials. *Color P* stands for positively skewed stimuli, *color N* – for negatively skewed ones.

Like for the middle lines, the effect of color on judgment of the length of all 14 lines was also non-significant tested with the Repeated Measurement Analysis (*F* (1, 28) =2.080, *p*=0.160). The within-subject factor number of trials caused a significant main effect on judgments: *F* (2,28)=10.830, *p*=0.000, ES=0.279. The insignificant tendency observed in judgment of all 14 lines to be rated higher when was negatively skewed than positively skewed decreases significantly with the number of trials.

Like in the rest of the reported experiments, the correlation between rating and line length was significant 0.82 (p=0.01).

The lack of statistically significant effect may be due to the second irrelevant dimension added in this experiment in comparison to the first one. The sounds could be perceived as quite different from participants and thus giving the opportunity for opposing shifts in their judgments. As Goldstone (1995) showed in his experiments, people are inclined to assimilate their judgments toward a category if they perceive the target as similar to it. Moreover, judgments may be assimilated toward categories that are constructed on-line based only on some sort of contextually sensitive similarities. It is possible that, in experiment2 two contextual forces are acting against each other – contrast that is due to the skew of lines against assimilation toward the categories of lines, which appear along with “laser” or “whoosh” sound. “Whoosh” and “laser” sound were different enough to provoke such effect.

On the other hand, the experimental manipulation could be obvious enough and participants may realize the idea of the experiment. Most of the participants shared the idea that the experiment was probably about the sound. Although none of them was sure what exactly was expected to happen, this according to Martin et al. (1990) may motivate participants to overcome the effect of context, e.g. to neglect the irrelevant information and to focus only on the relevant to the task line length.

7.2. *Experiment 4:* Judgment of the size of the Rabbits

Experiment 4 also tests the same hypothesis about the effect of several correlated dimensions on judgment. The stimuli, however, were more complex in comparison with the lines used in experiment 4, namely, rabbits. People were asked to judge the size of dappled white rabbits with red eyes and yellowish ear, and of homogeneously black rabbits with blue eyes and pink ears. The similarity between stimuli with respect to their irrelevant to the task characteristics was varied on five dimensions: color (black or white), pattern (homogenous or dappled), color of the eyes (red or blue) and color of the ears (yellowish or pink). All these irrelevant dimensions were correlated in the experiment. The expectation was that more differences between stimuli may increase the effect of the irrelevant to the task stimulus information.

7.2.1.Method

- Design

The independent variables were the color of the rabbits (dappled rabbits with red eyes and yellowish ear and homogeneously black rabbits with blue eyes and pink ears) and the number of trials (first and second 112 presentations). The group counterbalanced the irrelevant to the task color of the rabbits and the stimulus skew. Group 1 judged a mixture between positively skewed white dappled rabbits and negatively skewed black rabbits. Group 2 judged a mixture between positively skewed homogeneously black rabbits and negatively skewed white dappled rabbits. The dependent variable was the participant's 7-point rating depending on the stimulus color.

- *Stimuli*

A set of 14 rabbits with different sizes was designed. The smallest image of the rabbit was 150 pixels in width and 119 pixels in height, the largest image was 837 pixels in width and 663 pixels in height. Each stimulus image was 12% larger than the previous in the series. The frequency of the stimuli was calculated in order to receive a positively and negatively skewed sets depending on the stimulus color. The color distribution within the set of all 112 stimuli (14 sizes $_x$ 8 times) is presented in Table 15.

Table 15. Frequency distribution and color of the stimuli (where 1 represents image size of 150 pixels in width and 119 in height, stimulus 2 - 204 pixels in width and 162 in height and so on).

size	**Presentation frequency for rabbits with *color P***	**Presentation frequency for rabbits with *color N***
1 and 2	14	2
3 and 4	12	4
5 and 6	10	6
7 and 8	8	8
9 and 10	6	10
11 and 12	4	12
13 and 14	2	14

- *Procedure*

The experiment was conducted individually in sound-attenuated room. Each stimulus was presented 8 times within set of 112 trials. All 112 stimuli were randomly presented for judgment 2 times forming a sequence of 224 trials.

Participants were instructed to judge the size of each rabbit presented on the screen on a seven-point scale: where 1-“it is not big at all” and 7-“it is very big”. The experimenter registered participant’s rating and changed the slides manually. No time restrictions have been imposed on the participants. The experiment lasted about 20 minutes.

- *Participants*

85 students (56 female and 29 male) from the introductory classes in psychology at New Bulgarian University participated in order to satisfy a course requirement. 46 of the participants took part in group 1 and the rest 39 participated in group 2.

7.2.2. Results and discussion

Like in previous experiments, the data was averaged by 14 sizes of the rabbits. The Repeated Measurement ANOVA revealed a non-significant effect of the group (F (1, 84) =0.156, p=0.694) but a significant interaction between color and group: F (1, 84)=17.928, p=0.000). Figure 10 illustrates the mean

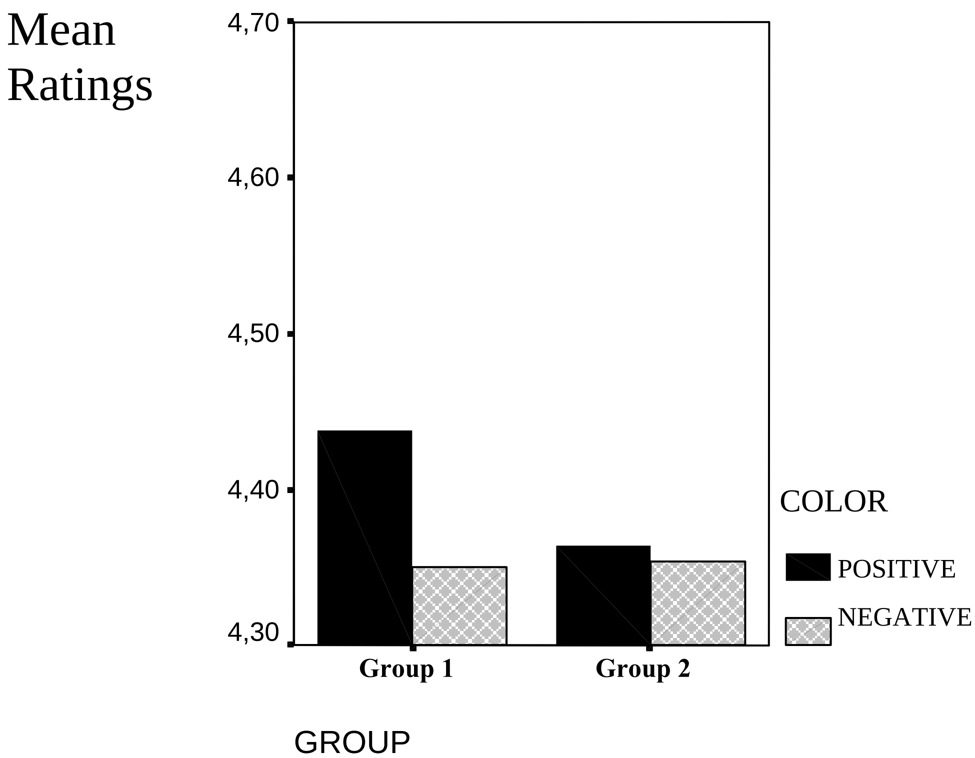


Figure 10. Mean differences in judgment of rabbit’s size depending on the stimulus color and experimental group.

differences in judgments of the 14 rabbit sizes depending on their color and for each group. The effect of color was significant at the first experimental group (F (1, 45) = 38.727, p=0.000, ES=0.463) but insignificant at the second group (F (1,38)=0.937, p=0.339, ES=0.24).

It seems that black rabbits were underestimated in comparison with the white rabbits. Probably, participants saw the black rabbits as smaller than the white spotted ones. Since the data from both groups was combined, the possible perceptual size illusion should be counterbalanced. Thus, the recoded data (from the original white and black color to *color P* and *color N*) will reveal only the effect of the irrelevant dimension rather than the effect of the perceptual size illusion that possibly were registered in the data.

The Repeated Measurement Analysis performed on this collapsed data revealed a significant effect of color (*color P* and *color N*) on judgment of the middle rabbits (F (1,84)=7.318, p= 0.008, ES=0.087). Middle-sized positively skewed rabbits were judged higher (mean=4.45) than negatively skewed ones (mean=4.43). The difference was 0.02. The effect of number of trials was estimated as non-significant: F (1,84)=2.595, p=0.111.

Repeated Measurement Analysis on all 14 rabbits sizes also showed a significant main effect of the within-subject factor color: F (1,84)=27.478, p=0.000, ES=0.246 (Figure 11).

The mean ratings and Standard Errors for each color (*color P* and *color N*) are presented in Table 16. The difference (0.06) between all positively and all negatively skewed rabbits was in the expected direction, i.e., rabbits that share *color P* were rated higher than rabbits that share *color N* despite their equal size.

The mean ratings and Standard Errors for each color (*color P* and *color N*) are presented in Table 16. The difference (0.06) between all positively and all negatively skewed rabbits was in the expected direction, i.e.,

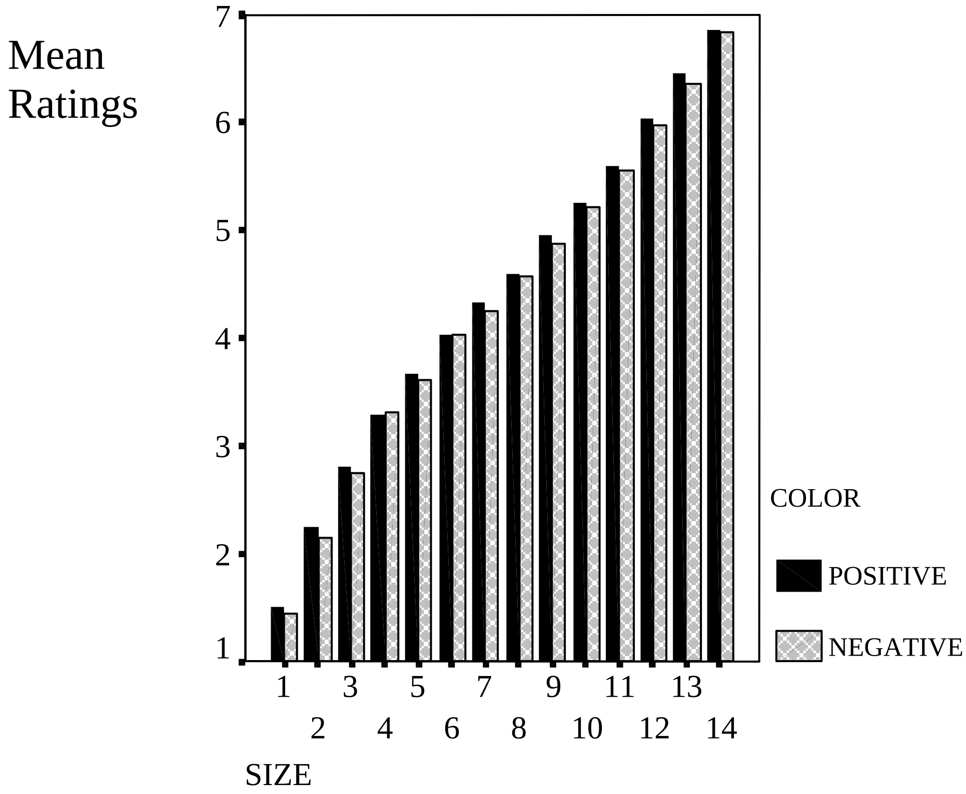


Figure11. Mean judgments of each stimulus size depending on its color.

rabbits that share *color P* were rated higher than rabbits that share *color N* despite their equal size.

The mean ratings and Standard Errors for each color (*color P* and *color N*) are presented in Table 16. The difference (0.06) between all positively and all negatively skewed rabbits was in the expected direction, i.e., rabbits that share *color P* were rated higher than rabbits that share *color N* despite their equal size.

Table 16. Mean ratings and Standard Errors of all rabbit sizes for each color.

Color	**Mean ratings**	Std. Error
*Color P (*positively skewed stimuli)	4.41	0.045
*Color N (*negatively skewed stimuli)	4.35	0.045

The effect of the within-subject factor number of trials was also estimated as significant for all 14 stimuli: F (1,84)= 7.274, p=0.008, ES=0.079. The effect of color decreases with time. The mean ratings and

Standard Errors for each color and block of 112 trials are presented in table 17.

Table 17. Mean ratings and Standard Errors depending on time and color.

Color	**Time**	**Mean ratings**	Std. Error
Color P (positively skewed stimuli)	First 112 trials	4.50	0.062
	Second 112 trials	4.36	0.058
Color N (negatively skewed stimuli)	First 112 trials	4.43	0.056
	Second 112 trials	4.39	0.056

Although stimuli were slightly more complex than the lines usually judged in the rest of the experiments, the correlation between judgment and size of the rabbits were significantly high: $r = 0.90$, $p=0.01$.

Both experiments reveal contradictory results. On one side, experiment 3 did not manifest any influence of irrelevant to the task color and sound. On the other side, experiment 4 reveals the expected effect of color of the rabbits on judgment of their size. The effect of the irrelevant dimension was confirmed at judgment of the size of both the middle size rabbits and the all-14 rabbit sizes. The difference in judgment of the same stimuli depending on their color was comparable in size to the effects reported in previous two sections. Thus, the impact of the irrelevant dimension was not enhanced essentially from correlation between several irrelevant dimensions.

The effect of the number of trials, however, observed in experiment 4 may result from the same forces that probably blocked out the influence of the irrelevant dimension in experiment 3. Since the stimuli were designed to be very dissimilar with respect to several irrelevant dimensions the on-line assimilation toward perceptually similar and on-line contrast from perceptually dissimilar stimuli could be brought into play (Goldstone, 1995). The two experiments, however, differ on the number of the presented for judgment stimuli. During experiment 3, were presented 3 sets of 112 stimuli, while in experiment 4 – 2 sets. Moreover, the effect of color was insignificant at experiment 3. Thus, the comparison between both experiments seems quite speculative. That is why, the time behavior of the effect of the effect of several correlated irrelevant dimensions could be considered as indefinite.

CHAPTER 8

Reduced Perceptual Support

The experiment described in this chapter is trying to differentiate the source of the effect, i.e. whether the effect of the irrelevant to the task dimension is comparable to the effect of sensory adaptation (for discussion see Arieh and Marks, 2002). The stimuli used in this experiment are quite abstract – they are numbers – and thus their color cannot effect their perception like in the previous experiments with line lengths. Thus if there is an effect of their color it cannot be attributed to the calibration process of the perceptual system as in other experiments on judgment of lines' length, loudness of tones, taste or touching (for detailed review, Arieh and Marks, 2002). It is difficult to argue that judgment of men' age based on digits, which stand for their absolute age in years may "induce perceptual system to recalibrate their relative supra-threshold responsiveness"(Arieh and Marks, 2002, p.478).

8.1. Judgment of age: *Experiment 5*

The main idea of this experiment was to investigate whether the effect of irrelevant dimension can be demonstrated with stimuli with higher complexity. Participants were asked to rate men' age based on green and red digits representing their age in years. Several strategies may be used for the fulfillment of this task, e.g. participants may visualize a man on a particular age in order to judge it or participants may base their judgment only on digits' height rather than representation. No matter what the strategy of participants is, it is difficult to suppose that judgments may be influenced from the digits' color. Thus, experiment 5 tests the effect of irrelevant dimension on judgment of stimuli that could not profit by perceptual input, i.e. the age do not depend on the size of the screen nor the age judgments can be affected from any recalibrations of the “relative supra-threshold responsiveness” of the perceptual system (Arieh and Marks, 2002, p.478).

8.1.1. Method

- *Design*

Like all experiments in this research, color and number of trials were within–subject variables. The group counterbalanced age and color: group 1 judged the age of small positively skewed red numbers and high negatively skewed green numbers. The dependent variable was the participant’s mean rating for each age depending on the digit’s color.

- *Stimuli*

A set of 14 ages was designed. The lower age was 10 years, the higher age – 75 years and the increment was 5 years. Like previous experiments, the ages were presented with uneven frequency depending on the digit’s color. The frequency of the stimulus distribution depending on stimulus color is presented in Table 18.

Table 18. Frequency of stimulus distribution depending on the irrelevant to the task color of the digits.

Age category	***The digit representing the target age***	***Frequency distribution of the ages with color P***	***Frequency distribution of the ages with color N***
1	10	7	1
2	15	7	1
3	20	6	2
4	25	6	2
5	30	5	3
6	35	5	3
7	40	4	4
8	45	4	4
9	50	3	5
10	55	3	5
11	60	2	6
12	65	2	6
13	70	1	7
14	75	1	7

- *Procedure*

Stimuli were presented for judgment one by one at the center of the computer screen. The background was gray. Similarly to the rest of the described experiments, stimuli were randomly presented within 2 sets with the same stimulus range and frequency. Participants rated twice the set of randomly presented 112 trials described in Table 18.

Judgments were required on a 7-point scale. Each age stay on a screen until the participant did not judge it. Then the experimenter registers respondents' rating and changes the slide manually. Experiment's duration was 15 minutes.

- *Participants*

31 (25 female and 6 male) on age between 19 and 29 years take part in the experiment for credits. 18 students participated in group 1 and 13 in group 2.

8.1.2. Results and Discussion

Data was averaged for each age. The between-subject factor group showed no effect: F (1,30)=0.083, p=0.775. Thus, like in the rest experiments, the original color was recoded into *color P* for the positively skewed stimuli and *color N* for the negatively skewed ones. Positively skewed middle ages received higher judgments than negatively skewed ones with respect to the recoded color (*color P* and *color N*). The mean ratings and Standard Errors for *color P* and *colors N* are presented in Table 19. The profile of the ratings with respect to the lines' color is presented at Figure 11.The effect of color on judgments of ages 40 and 45 years reach marginal significance tested with the Repeated Measurement Statistics: F (1,30)=3.497, p=0.072, ES=0.108. The main effect of the within-subject factor number of trials was non-significant for the middle-range of the stimulus ages (*F* (1,30)=0.642, *p*=0.429).

Table 19. Mean ratings and Standard Error for the middle ages color.

Color	Mean ratings	Std. Error
Color P (positively skewed)	4.09	0.090
Color N (negatively skewed)	4.04	0.098

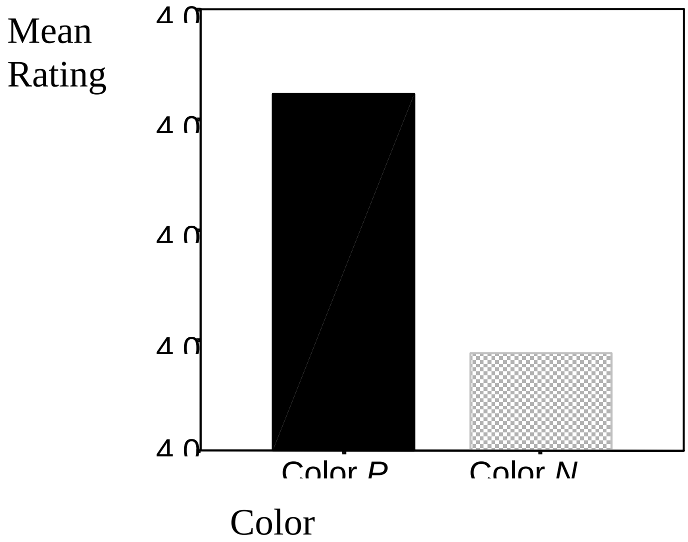


Figure 11. Mean ratings of the middle ages for each color. Black bar stands for the positively skewed lines. Gray textured bar represents the negatively skewed lines.

The main effect of color and number of trials was non-significant for the whole range of ages (10years-75 years) (color: F (1,30)=1.567, p=0.221; number of trials: F (1,30)=0.577, p=0.453). The correlation between judgment and age was significantly high: 0.94, p=0.01.

Results from experiment 5 are congruent with the results obtained from the rest of the experiments reported in this thesis. It demonstrates similar in direction, scope (middle lengths) and size, effect of the irrelevant to the task color on judgment of middle stimuli, i.e. ages 40 and 45.

Surprisingly, the effect of irrelevant dimension was observed on the background of quite high correlation between the judgment and the magnitude (0.94). This correlation was probably a result from the stimulus selection. Participants used quite simple strategy of judging stimuli according to the first digit: e.g., ages 10 and 15 with 1, ages 20 and 25 with 2, ages 30 and 35 with 3. The effect of color, however, appeared despite this stair-like profile of judgment.

Although the effect of the color reached only marginal significance, the results seem encouraging in the sense that they replicate the previous ones obtained with simple stimuli. Moreover, the behavior of the effect seems quite predictable, i.e. the same contextual shift mainly within the middle stimuli at comparable intensity. The most important observation of this experiment is that the effect of irrelevant dimension was demonstrated with complex stimuli. Therefore, it seems difficult, if possible at all, to account for the influence of irrelevant to the task dimension by referring to the low-level mechanisms like "recalibration" of perceptual system sensitivity (Marks, 1988, 1992, 1994, Marks and Warner, 1991, Arieh and Marks, 2002). It is possible, however, that context influences judgment at several different levels of information processing. It could also be the case that contextually sensitive processes are running in parallel and give raise to different shifts in human judgments (some of which can be congruent and others can be canceling each other). This additionally impedes the development of detailed and elaborate description of the judgment process.

Overall, it seems that, the spreading activation mechanism may better describe the effect of the irrelevant information with such complex stimuli than the "early local changes in receptive sensitivity" proposed by Marks and colleges (Marks, 1988, 1992, 1994, Marks and Warner, 1991, Arieh and

Marks, 2002). Moreover, unlike perceptual learning mechanisms (Goldstone, 1995, 1998) that are also argued to be able to account for contextual effects in judgment of complex stimuli, JUDGEMAP-1 proposes mechanisms that are detailed enough to be tested experimentally.

In conclusion, the results of this experiment provide supporting evidence, that judgment of complex stimuli could be influenced by stimulus irrelevant dimension just like judgment of more simple stimuli like lines. The spreading activation mechanism is able to account for these results because it does not pose any limits on stimulus complexity. Thus, it seems reasonable to assume that irrelevant stimulus dimension affect judgment on different levels of the information processing, e.g. relatively early by means of perceptual “recalibration” (Marks, 1988, 1992, 1994, Marks and Warner, 1991, Arieh and Marks, 2002) and perceptual learning mechanisms (Goldstone, 1995) and also relatively late in information processing through contextually sensitive retrieval.

CHAPTER 9

Conclusion

9.1. Conclusion and future work

The main goal of the presented research was to test a specific prediction of JUDGEMAP-1 Model, namely that the irrelevant characteristics of the stimuli may influence judgment. According to JUDGEMAP-1 the target stimulus is not be judged in isolation but rather within a contextually sensitive set that comprises both some perceived environmental stimuli and retrieved exemplars from the LTM (similar, familiar exemplars, generalized prototypes if such exist in LTM, and recently judged stimuli). The mechanism responsible for the retrieval of the relevant information from the LTM is spreading activation. Thus, it turns out, that according to JUDGEMAP-1 model the information needed for judgment embraces the whole information that the stimulus posses. Therefore, judgment process may be affected by both the relevant dimension(s) and the irrelevant stimulus dimensions (i.e., the additional information that is not required to be judged in a particular judgment task). It is quite possible, of course, that the irrelevant information has a smaller weight in the judgment process than the relevant one, but nevertheless, it takes part in judgment.

Other researchers proposed the same idea but they situated the mechanism responsible for this effect on the level of perception (Arieh and Marks, 2002) and peceptual categorization (Goldstone, 1995). For example, Arieh and Marks (2002) argued that the effect of irrelevant dimension was comparable to the effect of sensory adaptation. Receptors recalibrate the input depending on irrelevant to the task dimension. Many experiments that require judgment of simple stimuli from various modalities reported a contrastive shift in judgment depending on the context created by the irrelevant dimension (e.g., Rankin and Marks, 1991, 1992, 2000; Marks and Armstrong, 1996;

Armstrong and Marks, 1997; Potts, 1991; Arieh and Marks, 2002). Some experiments (Goldstone, 1995), however, demonstrated assimilation effect due to the irrelevant stimulus dimension and argue that this may be a perceptually driven process.

Unlike these experiments (Rankin and Marks, 1991, 1992, 2000; Marks and Armstrong, 1996; Armstrong and Marks, 1997; Potts, 1991; Arieh and Marks, 2002; Goldstone (1995)), however, the reported research varied the frequency rather than the range of the stimuli depending on stimulus irrelevant dimension(s). The stimuli were skewed with respect to the irrelevant to the task color rather than differentiated as having different end points. Moreover, the effect reported in this paper was demonstrated across stimuli with different complexity and level of concreteness/abstraction, e.g. judgment of line length, judgment of rabbit size and judgment of the age. The effect seems independent of stimulus complexity and abstraction level and hence sensory adaptation (Arieh and Marks, 2002) can not account for it. It seems difficult to argue for an effect of sensory adaptation in judgment of age described with numbers. It is probable, however, the effect of the irrelevant stimulus dimension reported in the presented experiments to be a result from perceptual learning. As Goldstone (1995, 1998) argued, perceptual learning could account for contextual effects in judgment despite stimulus complexity. In general, the interplay between perceptual learning and high level process like judgment needs further and more detailed discussion.

Overall, the effect of irrelevant dimension was demonstrated in 5 out of 6 experiments (including the initial experiment by Kokinov et. al., 2004) that certainly contradict traditional theories of judgment. The reported series of experiments shows that judgment is sensitive to the irrelevant to the task attributes of the stimuli. On the contrary, traditional theories of judgment consider only the relevant dimension(s) of the target stimuli. For example, *R* and *Fr* principles were demonstrated with stimuli that varied in range and skew with respect to the judged dimension. Neither the Range-Frequency Theory (Parducci, 1965, 1968, 1974) nor the Two-Path Theory (Manis and Paskewitz, 1984) that shares the R and Fr origin of contextual contrast can account for the effect reported in this research without assuming that judgment is sensitive to the context of irrelevant stimulus dimension. JUDGEMAP-1, in addition, relates the effect of irrelevant dimension with particular mechanisms, namely the spreading activation mechanism that can be responsible for the effect's appearance. Although, these mechanisms are in their essence analogous to the once proposed by the Norm Theory (Kahneman and Miller, 1986), the latter could not predict the direction of the judgment's shift since it does not specifies how the target stimulus can be compared to the constructed norm.

In general, the reported experiments demonstrated small but significant effect of irrelevant dimension on judgment. The effect was relatively stable across different experiments even though the experimental manipulations of several variables (R and Fr; correlated dimensions; stress) were expected to enlarge the size of the effect. The effect was smaller than the empirical maximum obtained by the two control experiments but comparable in to the effect of JUDGEMAP-1 simulation (Kokinov et al., 2004).

Several possibilities seem relevant for explaining the small impact of the irrelevant information. First, it is possible that irrelevant information is considered but with lower weight than the relevant (the judged) one. People may pay less attention to the irrelevant to the task stimulus features and thus retrieve exemplars that are similar to the target stimulus along these irrelevant dimensions with smaller weight (probability). This explanation may shift the debate about the mechanisms responsible for the effect of the irrelevant dimension toward attention mechanisms, which, in principle, could be considered as related to the perceptual learning mechanisms (Goldstone, 1998).

Second, it could be that the size of the effect depends on participants' precision in judgment along the relevant dimension rather than the associative weights between judged stimuli and exemplars (or prototypes) stored in the LTM. Overall, participants in the reported experiments were quite precise in their judgments. Thus, if contextual effects in judgment are considered as "mistakes", it seems reasonable to suppose that lower precision in judgment may ensure greater contextual effects. Unfortunately, judgments' precision was not varied in neither of the reported experiments. Even, experiment 5 was not designed to decrease judgment precision because participants judged ages with the same initial digit as belonging to the same scale category, e.g. 30 and 35 were judged as 3, 40 and 45 – as 4. Thus, in order to obtain more information the issue concerning the dependence between judgement precision and the size of contextual effects should be pushed further.

Third, it could be, however, that the effect of irrelevant dimension presented in this paper pushes the judgment opposite to contextually sensitive on-line categorization (i.e., assimilation toward perceptually similar objects and contrast from perceptually dissimilar objects) reported by Goldstone (1995). Suppose that participants judge positively skewed green lines and negatively skewed red lines. Then according to Goldstone's observations, green lines should be judged lower than the red lines with the same length. This is exactly the opposite of the reported results. It could however be, that Goldstone's effect "fights" against the effect of stimulus skew manipulated through the irrelevant dimension reported in this research. This may be one of the reasons for the small size of the contextual shift in judgments due to the irrelevant dimension. Moreover, the effect disappears in Experiment 3, where the stimuli were designed to be very dissimilar with respect to the irrelevant to the task information, i.e. the lines were presented always in the context of particular meaningful sounds. Unfortunately, it seems extremely difficult to study the effect of the irrelevant dimension without activation of the perceptually grounded on-line categorization observed by Goldstone.

Finally, the reported in this thesis series of experiments did not reveal a stable tendency in the behaviour of the effect of the irrelevant information depending on the number of trials. The interest in the behavior of the effect over time was provoked by the contra-hypothesis that context effects could be considered as a noise in the system. If this was the case, the effect of irrelevant dimension as well as any other contextual effect should decrease with time. Overall, the number of trials seems not to influence the effect of the irrelevant information in a systematic manner. Time enhanced the effect of the irrelevant dimension in Experiment 1, decreased the effect in Experiment 4, and did not affect substantially the contrast due to the irrelevant information in the rest of the experiments.

In conclusion, experiments demonstrate systematic shift in judgment depending on an irrelevant to the task dimension of the stumuli. The impact of the irrelevant information was small and relatively stable across different experimental manipulations. Results were in line with JUDGEMAP-1's prediction and hence support it.

9.2.Contributions

The experiments presented in this thesis demonstrate a stable in direction and size contrast effect in judgment due to the irrelevant dimension. In addition, the same effect was shown with both simple and complex/abstract stimuli, which may further challenge the traditional differentiation between judgment of simple (i.e., psychophysical) and judgment of complex (i.e., social) stimuli.

The experiments study the possibility for increasing the impact of the irrelevant dimension trough several experimental and theoretically motivated (with respect to JUDGEMAP-1) manipulations but failed to find any increase of the size of the effect.

This research presents an attempt to explain the effect of the irrelevant dimension by particular mechanisms, namely the spreading activation mechanism used in the formation of the comparison set in WM. This explanation has been supported by the confirmation of JUDGEMAP's predictions.

REFERENCES

Anderson, J., Lebiere, Ch. (1993). The Atomic Components of Thought, *Lawrence Erlbaum Associates,Hillsdale, NJ*

Anderson, N.H. (1965), Averaging versus adding as a stimulus-combination rule in impression formation, *Journal of Experimental Psych*ology, 70, 394-400.

Anderson, N.H. (1966), Component ratings in impression formation, *Psychon. Sci*., Vol.6 (6).

Anderson, N. H. (1971a), Integration theory and attitude change, *Psychological Review*, May, Vol.78, No 3.

Anderson, N.H. (1971b), Two more tests against change of meaning in adjective combination, *Journal of verbal learning and verbal behavior*, 10, 75-85.

Anderson, N.H. ,Lindner, Rh. and Lopes,L.L. (1973), Integration theory applied to judgments of group attractiveness, *Journal of Experimental and Social Psychology*, Vol26, No3, 400-408 .

Arieh and Marks (2002) Context effects in visual length perception. Role of ocular, retinal, and spatial location. *Perception & Psychophysics*, vol. 64 (3), pp. 478-492

Barg, Joan A. and P.Pietromonaco (1982), Automatic information processing and Social Perception: The influence of trait information presented outside of conscious awareness on impression formation, *Journal of Personality and Social Psychology*, No.3, 437-449.

Biernat, Monica, Manis, Melvin and **Kobrynowicz, Diane** (1997), Simultaneous Assimilation and contrast effects in judgments of self and others, *Journal of personality and Social Psychology*, Vol. 73, N 2, 254-269.

Cooke, Allen D.J. and **Barbara A.Mellers** (1998), Multiattribute Judgment: Attribute Spacing Influences Single Attributes, *Journal of Experimental Psychology: Human Perception and Performance*, 24 (April), 496-504

Geiselman,R.E., Haight,N.A., & Kimata,L.G. (1984). Context Effect on the Perceived Physical Atractiveness of Faces, *Journal of Experimental Social Psychology*, 20, 409-424

Goldstone, R. (1995) Effects of Categorization on Color Perception. *Psychological Science*, vol. 6 (5), pp. 298-304

Goldstone,R.,L. (1998). Perceptual Learning. *Annual Review of Psychology*, 49, 585-612.

Jesteadt, W., Luce, R., Green, D. (1977). Sequential Effects in Judgment of Loudness. *Journal of Experimental Psychology: Perception and Performance, vol. 3 (1), pp. 92-104.*

Helson,H. (1964), Adaptation Level Theory: An experimental and systematic approach to behavior, New York: Harper&Row.

Herr, Sherman and Fazio (1982) On the Consequence of Priming: Assimilation and Contrast Effects. *Journal of Experimental Social Psychology,* vol. 19, pp 323-340.

Kahneman, D., and Miller, D.(1986), Norm Theory: Comparing Reality to Its Alternatives, Psychological Review, Vol.93, No 2, 136-153.

Kenrick, D. T., and Gutierres S.E. (1980), Contrast Effects and Judgments of Physical Attractiveness: When beauty becomes a social problem, *Journal of Personality and Social Psychology*, 38, 131-140.

Kokinov, B. (1994a). A hybrid model of reasoning by analogy. In K. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 247-318). Norwood, NJ: Ablex

Kokinov, B. (1994b). The DUAL cognitive architecture: A hybrid multi-agent approach. *Proceedings of the Eleventh European Conference of Artificial Intelligence*. London: John Wiley & Sons, Ltd.

Kokinov, B. (1994c). The context-sensitive cognitive architecture DUAL. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. (1995). A Dynamic Approach to Context Modeling. In: P. Brezillon & S. Abu-Hakima (eds.) *Proceedings of the IJCAI-95 Workshop on Modeling Context in Knowledge Representation and Reasoning.* LAFORIA 95/11.

Kokinov, B. (1998). Analogy is like cognition: Dynamic, emergent, and context-sensitive. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research.* Sofia, Bulgaria: NBU Press.

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*, 2004. Erlbaum, Hillsdale, NJ.

Lockhead, G. (1992) Psychophysical scaling: Judgments of attributes or objects? *Behavioral & Brain Sciences*, 15, 543-559.

Manis, M., Paskewitz, J. (1984) Judging Psychopatology: Expectation and Contrast. *Journal of Experimental Psychology*, vol. 20, pp 217-230.

Manis, M., Biernat, M., and Th. Nelson (1991) Comparison and expectancy processes in human judgment, *Journal of Personality and Social Psychology,* Vol.61, No.2, 203-211

Manis, M., Nelson, Th., Sheder, J. (1988) Stereotypes and Social Judgment: Extremity, Assimilation and Contrast.*, Journal of Personality and Social Psychology,* vol. 55 (1), pp 28-36.

Marks, L. (1988) Magnitude estimation and sensory matching. *Perception and Psychophysics*, vol. 43, pp. 511-525

Marks, L., E.(1992), The slippery context effect in psychophysics: Intensive, extensive, and qualitative continua. *Perception and Psychophysics*, 51, 187-198.

Marks, L., **E**. (1994), "Recalibrating" the auditory system: The perception of loudness, *Journal of Experimental Psychology: Human Perception and Performance,* 20, 382-396.

Marks and Armstrong (1996), Visual and haptic representations of space. In T.Inui and McClelland (Eds.), *Attention and Performance XVI: Information Integration in perception and communication* (p. 263-287). Cambridge, MA: MIT Press.

Marks and Warner (1991), Slippery context effect and critical bands*, Journal of Experimental Psychology: Human Perception and Performance,* 17, 986-996

Martin, L.L., J.J. Seta, and R.A. Crelia (1990), Assimilation and contrast as a function of people's willingness to expend effort in forming an impression, *Journal of Personality and Social Psychology*, 59, 27-37.

Mellers, Barbara A. and **Alan D.J.Cooke** (1994), Trade-offs Depend on Attribute Range*, Journal of Experimental Psychology: Human Perception and Performance,* 20 (October), 1055-1067.

Mellers, Barbara and **Michael H. Birnbaum** (1981), Contextual Effects in Social Judgment, *Journal of Experimental Social Psychology*, 14(July), 157-171.

Mellers, Barbara and **Michael H. Birnbaum**(1982), Loci of Contextual Effects in Judgment, *Journal of Experimental Psychology: Human Perception and Performance,* Vol.8, 582-601.

Minsky, M. (1985). *The Society of Mind*. Simon and Schuster Inc., New York, NY.

Pardicci, Allen (1965), Category Judgment: A Range-Frequency model, *Psychological Review*, 72(6), 407-418.

Pardicci, Allen (1968), The Relativism of Absolute Judgments, *Scientific American*, vol.219 (6), 84-90.

Parducci, A.and L. Marshall (1962), Assimilation vs. contrast in the anchoring of perceptual judgments of weight, *Journal of Experimental Psychology*, Vol.63, No. 5, 426-437.

Pardicci, Allen and **Linda F.Perrett** (1971), Category Rating Scales: Effects of relative spacing and Frequency of Stimulus Values, *Journal of Experimental Psychology*, 89(August), 427-453.

Parducci, Allen (1974), Contextual Effects: A Range-Frequency Analysis, *Handbook of Perception*, vol.2, NY:Academic Press, 127-141.

Parducci and Wedell (1986), The category effect with rating scales: number of categories, number of stimuli, and method of presentation, *Journal of Experimental Psychology:Human Perception and Performance*, 12, 496-516

Petrov, A., Anderson, J. (2000), ANCHOR*:* A Memory-Based Model of Category Rating. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference the Cognitive Science Society* (pp. 369-374). Hillsdale, NJ: LEA.

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

Petrov, A., Dosher, B., & Lu, Z.-L. (in press). The Dynamics of Perceptual Learning: An Incremental Reweighting Model. *Psychological Review*.

Potts, B., **C**., (1991), The horizontal-vertical illusion: A confluence of configural, contextual, and framing factors. *Unpublished doctoral dissertation, Yale University*.

Rankin and Marks (1991), Differential context effects in taste perception. *Chemical Senses*, 16, 617-629.

Rankin and Marks (1992), Effects of context on sweet and bitter tastes: Unrelated to sensitivity to PROP (6-n-propylthiouracil), *Perception and Psychophysics*, 52, 479-486.

Rankin and Marks (2000), Differential context effects in chemosensation: Role of perceptual similarity and neural communality. *Chemical Senses*, 25, 747-759.

Restle, F. (1978), Assimilation Predicted by Adaptation-Level Theory with Variable Weihgts, *Cognitive Theory*, vol.3, Hillsdde, NY:Erlbaum, 75-81.

Sarris, Victor and **Parducci, Allen** (1978), Multiple Anchoring of Category Rating Scales, *Perception and Psychophysics*, vol.24 (1), 35-39.

Schwarz, N., and H. Bless, (1991), Constructing reality and its alternatives: assimilation and contrast effects in social judgment, In L.L. Martin and A. Tesser, eds, *The construction of social judgment*, 217-245, Hillsdale, NJ: Erlbaum.

Schwarz, N., and F. Strack (1991), Context effects in attribute surveys: Applying cognitive theory to social research, In W. Stroebe and M. Hewstone, eds., *European review of social psychology*, vol.2, 31-50, Chichester: Wiley.

Sherif, M., Taub, D., and Hovland, C. (1958), Assimilation and contrast effects of anchoring stimuli on judgments, *Journal of Experimental Psychology*, Vol.55, No.2, 150-155.

Srull, Th. and Wyer, J.(1979), The role of category accessibility in the interpretation of information about persons: Some determinants and implications, *Journal of Personality and Social Psychology*, 37,1660-1672.

Srull, Th., Wyer, J. (1980) Category Accessibility and Social Perception: Some Implications for the Study of Person Memory and Interpersonal Judgment. *Journal of Personality and Social Psychology,* vol. 38 (6), pp 841-856.

Stapel, D. W. Koomen, and Plight (1996), The refference of trait inferences: the impact of trait concepts versus actor-trait links on subsequent judgements, *Journal of Personality and Social Psychology*, vol70, No.3, 437-450.

Stapel and Koomen (2000), How far do we go beyond the information given? The impact of knowledge activation on interpretation and inference, *Journal of Personality and Social Psychology*, V0l.78, No.1, 19-37.

Strack, F., N. Schwarz, and E. Gschneidinger (1985), Happiness and reminiscing: The role of time perspective, mood, and mode of thinking. Journal of Personality and Social Psychology 49, 1460-1469.

Ward, L. (1973) Repeated magnitude estimations with a variable standard: Sequential Effects and Other Properties. *Perception and Psychophysics, vol.13, pp. 193-200.*

Ward, L. (1979) Stimulus Information and Sequential Dependencies in Magnitude Estimation and Cross-Modality Matching. *Journal of Experimental Psychology: Perception and Performance, vol. 5 (3), pp. 444-459.*

Wedell, Douglas H., Allen Parducci, and **R. Edward Geiselman** (1987), A Formal Analysis of Ratings of Physical Attractiveness: Successive Contrast and Simultaneous Assimilation, *Journal of Experimental Social Psychology*, 23(May), 230-249.

Wedell D.H., Parducci A., and **Lane M**.(1990), Reducing Contextual Effects for Ratings of Clinical Case Histories, *Journal of personality and Social Psychology* ,Vol.58, N 2, 319-329.

Wyer, R.S. (1974), Changes in meaning and halo effects in personality impression formation, *Journal of personality and Social Psychology,* Vol.29, No6, 829-835.

List of publications, related to the dissertation:

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, 2004. Erlbaum, Hillsdale, NJ.*

Hristova, P. (2005a) Fragile Contextual Effects in Judgment of Length. *In: Technical Report LIP 2005/007 of the Laboratoire d'Informatique de Paris 6, URL http://www.lip6.fr/fr/production/publications-rapport-fiche.php?RECORD_KEY%28rapports%29=id&id(rapports)=243*

Hristova, P. (2005b) The Mechanisms of Contextual Change in Judgment. *In: Proceedings of the Balkan Conference of Young Scientists, 2005, Plovdiv. In press.*

Petkov, G., Hristova, P., Kokinov, B. (2005) How Irrelevant Information Influences Judgment *In: Technical Report LIP 2005/007 of the Laboratoire d'Informatique de Paris 6, URL http://www.lip6.fr/fr/production/publications-rapport-fiche.php?RECORD_KEY%28rapports%29=id&id(rapports)=244*

Hristova, P., Petkov, G. Kokinov, B. (2005). Influence of Irrelevant Information on Price Judgment. *In: Proceedings of the International Conference on Cognitive Economics*. NBU Press.