NEW BULGARIAN UNIVERSITY

DEPARTMENT OF COGNITIVE SCIENCE AND PSYCHOLOGY

Perceptual Learning – Perceptual Changes in Learning New Categories

Alexander Enchev Gerganov

An extended abstract (autoreferat) submitted in partial fulfillment of the requirements for the degree of Ph.D. in Psychology

Supervisor: Assoc. Professor Maurice Grinberg, NBU
External Adviser: Professor Robert Goldstone, Indiana University

Sofia • 2015

The dissertation is submitted in English language.

- Volume: 139 pages (without bibliography, appendices and title pages)
- Used literature: 101 references in English
- Tables: 8
- Figures: 39

Summary

Perceptual learning traditionally focuses on studying early plasticity in the sensory pathways. Categorization is a task typically attributed to relatively higher-level cognition. This dissertation explores whether visual categorization of simple objects involves perceptual learning.

A simple, physiologically plausible neural network model is put forward to demonstrate how both supervised and unsupervised learning could take place in visual categorization. The model shows how early perceptual representations could be formed, adjusted and reorganized without feedback. At the same time, the re-weighting of the connections between low-level representations and mid-level perceptual structures allows the selective filtering of early-level information and the formation through supervised learning of perceptual detectors for characteristic parts of objects. The model demonstrates how perceptual learning processes could take place in seemingly higher-level tasks like categorical learning of simple objects.

Several experiments, using a position transfer paradigm in categorical learning of simple objects, are presented in addition to the new model. They confirm empirically the claims of the model: that visual categorical learning involves lower-level perceptual learning processes. Evidence for such processes is found in the incomplete transfer of learning when the stimuli are presented at a new location in the visual field. This effect is observed even when the participants' categorization strategy is very simple and explicit. These results complement a large pool of similar findings for arguably lower-level perceptual tasks like vernier discrimination or orientation judgment.

This dissertation argues that visual categorical learning of simple objects involves low-level perceptual learning processes, similar to the processes in typical sensory perceptual tasks.

Table of Contents

Chapter I. Introduction

A contemporary definition of perceptual learning would usually describe it as a relatively long-lasting performance improvement in perceptual (low-level, sensory) tasks that is the result of training, practice or experience (Fahle & Poggio, 2002; Sagi, 2011). Such "low-level" or "sensory" tasks in the visual domain are vernier acuity (Fahle & Edelman, 1993), orientation judgment (Schoups, Vogels, & Orban, 1995), phase discrimination in compound gratings (Fiorentini & Berardi, 1981), and others. Performance improvement in these tasks is assumed to require changes in low-level perceptual structures because of the nature of the tasks. Indeed, such low-level changes for these tasks are demonstrated through research in psychophysics, neuroscience and computational modelling (Fahle & Poggio, 2002; Sagi, 2011). For example, various studies have shown that learning in these tasks is specific for the retinal location where the training takes place and does not transfer to other locations (Dill, 2002). Studies from neuroscience have found neuronal correlates for these empirical findings in primary structures such as V1 (Schoups, Vogels, Qian, & Orban, 2001). Still, there is no consensus about the exact mechanisms and the level of neuronal plasticity, connected with perceptual learning in these simple tasks (Sagi, 2011).

At the same time, researchers in cognitive science have also adopted the term "perceptual learning" for higher-level tasks like, for example, visual categorization (Goldstone, 1998). Some of the mechanisms, discussed in this context, such as "attentional weighting", "stimulus imprinting", "differentiation", "unitization" (Goldstone, 1998) can be found functionally in various psychological studies. There isn't, however, the clear connection between these phenomena, observed in a variety of psychological experiments and the particular brain structures or basic perceptual mechanisms which are the primary focus of discussion in perceptual learning.

The goal of this dissertation is to study the role of perceptual learning in a typical higher task such as visual categorization of simple stick-figure objects (Pevtzow & Goldstone, 1994). This is achieved through series of simulations with a newly developed model of categorical perceptual learning and through several experiments which explore the position specificity of categorical perceptual learning of simple objects.

This dissertation is organized in the following way:

Chapter II consists of a review of the general findings in perceptual learning and in particular – research connected to position specificity of learning.

Chapter III reviews briefly several examples of the types of models used in perceptual learning and some of the more influential theories in the field.

Chapter IV provides some examples of higher-level categorization tasks which could involve perceptual learning processes and focuses on two particular studies which are later used as a basis for simulations and experiments.

Chapter V presents the research goals and the methodological approach for the dissertation.

In Chapter VI, there is a brief description of the used modeling mechanisms and a neural network model of categorization, developed by Goldstone (2000), is replicated, studied and discussed in detail.

Chapter VII introduces a new, improved model of categorical perceptual learning and describes two simulations of supervised and unsupervised categorical learning.

Chapter VIII presents several experiments (2 pilot and 3 main experiments) which study the position specificity effect in categorical learning of stick-figure objects – a classical effect from psychophysics which is taken as evidence for lower-level changes during perceptual learning in low-level tasks. This effect is seldom studied for higher-level tasks like categorization of simple objects based on characteristic elements.

Chapter IX discusses the experimental results and simulations, their theoretical implications as well as their shortcomings.

Chapter X describes how the model could be developed further and other perspectives for future work based on the research, presented in the dissertation.

Chapter XI concludes the dissertation with a short discussion of the contributions of this research.

Finally, there is a bibliography of the literature, used in the dissertation and a list of the author's publications based on this research.

Chapter II. Perceptual Learning

Psychophysicists have studied perceptual learning using a variety of different "sensory" or "low level" tasks (Fine & Jacobs, 2002). A task which is often used in perceptual learning is vernier discrimination (Fahle, Edelman, & Poggio, 1995; Fahle & Edelman, 1993; McKee & Westheimer, 1978; Saarinen & Levi, 1995). In this task participants are presented with two parallel horizontal or vertical lines (or dots) that are slightly displaced (Figure 1). The task is simply to judge if the target line (or dot) is to the right or to the left (respectively up or down) of the other one. Participants improve their performance really fast and even perform above chance at resolution which is finer than the spacing between individual photoreceptors in the retina. The last phenomenon is known as visual hyperacuity (Fahle et al., 1995). The vernier task is assumed to be mediated by low-level structures and mechanisms (Herzog & Fahle, 1997). Another set of tasks involve orientation discrimination where observers are trained to identify orientations of lines, Gabor patches, noise fields, etc. (Matthews & Welch, 1997; Schoups et al., 1995; Vogels & Orban, 1985). Pattern, object and face discrimination tasks are also quite common – familiar object discrimination (Furmanski & Engel, 2000), novel face discrimination (Gold, Bennett, & Sekuler, 1999).

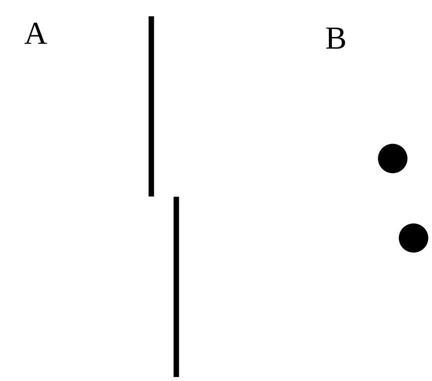


*Figure 1.*Two examples of vernier discrimination task: with lines (A) and dots (B). Participants have to make a judgment about the direction of displacement of the two lines (A) or the dots (B)

Performance improvement to the level of hyperacuity in vernier tasks or learning in tasks, relying heavily on orientation judgment can be easily traced to structures as early as V1 (Schoups et al., 2001), however object and face discrimination involve more complex processes and arguably higher brain structures than the primary visual cortex. Therefore, an important question for the whole perceptual learning field is what constitutes a "low-level" task and is perceptual learning limited to these tasks only. Since the exact mechanisms which make perceptual learning possible are still a matter of debate, it is very difficult to say where perception end higher-level cognition starts. In order to study these phenomena, a deliberate blurring of the boundary between conception and perception has been suggested (Goldstone & Barsalou, 1998).

Several mechanisms are considered to play a role in perceptual learning. Among them are the change of low-level representation in the visual cortex and selective reweighting of the connections between low-level representations and higher-level judgment structures (Fahle & Poggio, 2002; Sagi, 2011). Selective attention could also explain some experimental findings (Z.-L. Lu & Dosher, 2004; Zhong-Lin Lu, Lesmes, & Dosher, 2002; O'Toole & Kersten, 1992; Xiao et al., 2008). There is no common agreement about the exact role of each of these possible mechanisms in perceptual learning during different tasks (Sagi, 2011).

Position Transfer in Perceptual Learning

Visual perception is to a large degree invariant for various sizes, positions in the visual field and even rotations. Usually, models of object recognition are focused on the problem how to simulate computationally translation and rotation invariant object recognition (Riesenhuber & Poggio, 1999).

At the same time, there are many experiments which show that perceptual learning is not necessarily translation invariant (Dill, 2002). When learning occurs at a particular position in the visual field, its lack of translation invariance (or in other words the position specificity of learning) would mean that when the stimuli are presented at a different location in the visual field (and at the same distance from the fovea) the performance improvement found at the first location doesn't transfer to the new one and performance starts from its initial levels, as if no learning has occurred.

Receptive fields of the cells increase in size, the higher they are in the visual pathways (Bruce, Desimone, & Gross, 1981; Hubel & Wiesel, 1962). Based on this fact, the position transfer paradigm attempts to link the presence or lack of transfer of learning with the exact level of the cortical structures involved in the learning process. The rationale here is that if a stimulus is presented outside of the receptive fields of the cells that were involved in the learning process, then the performance will decline, sometimes to the starting levels.

There are a few studies which demonstrate position specificity of the perceptual learning taking place in vernier discrimination tasks. For example, Beard et al. (1995) found only partial transfer (along meridians) at a very small distance of 2 degrees of visual angle.

Fahle (1994) and Fahle et al. (1995) tested whether there is transfer of learning at distance of 10 degrees visual angle at several locations. They found that learning in vernier task was position specific. However, there was some evidence that learning might be unspecific for naïve observers (participants who hadn't participated in perceptual learning experiments before).

Line orientation discrimination is another low-level task which is used very often in perceptual learning paradigms. Shiu and Pashler (1992) demonstrated there wasn't transfer of learning at distance of 11 degrees. Schoups et al. (1995) found no position transfer at distances as small as 2.5 degrees of visual angle in a grating orientation discrimination task.

Pattern and object recognition and discrimination tasks involve arguably more complex visual processing and it is very interesting that there are various studies which again show position specificity of learning. Pattern discrimination turned out to be position specific even at distances as close as 1 degree of visual angle (Dill & Fahle, 1998; Nazir & O'Regan, 1990).

Visual perceptual learning is usually associated with tasks like vernier discrimination, line orientation discrimination, and simple pattern recognition. One reason for these tasks to be considered low-level is the inability of the participants to explain explicitly the changes in their strategy/performance over the course of improvement. The assumption for implicitness of perceptual learning, however, is more or less intuitive and presumed on the basis of the specific nature of the learning tasks that are used.

Probably because of the assumption of implicitness of perceptual learning, categorical learning is seldom studied in the classical perceptual learning paradigms (focused on different forms of specificity of learning). One example of a study which addressed the question of position specificity or translation invariance of categorical learning used 15 compound Gabor gratings as stimuli. The 15 gratings were organized in 3 classes – open circle, open square and open triangle (Figure 2). The authors argue that position invariance is an emergent property of category learning, because categories became gradually accessible at new locations where position specific feedback was not received.

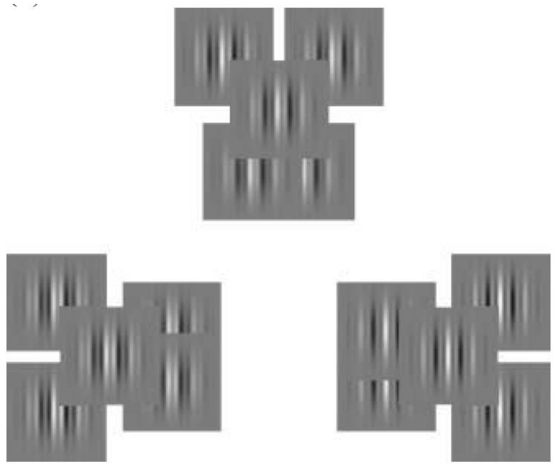


Figure 2. Stimuli used by Jüttner and Rentschler (2008) in a categorization task. 15 Gabor gratings are organized in three classes – triangle, square and circle. The image is taken from (Jüttner & Rentschler, 2008, fig. 1b)

It remains to be seen how these results from categorizing Gabor patches would relate to categorization of objects based on a characteristic element.

Chapter III. Models and Theories of Perceptual Learning

Modeling places important constraints on explanations about perceptual learning and pushes theoretical accounts to be more quantitative and concrete. Testable behavioral predictions are often derived from simulations.

A well-known, supervised, feedforward neural network model in perceptual learning is suggested by Poggio, Fahle and Edelman (1992). The model is representative of the class of supervised, feedforward neural networks. Perceptual learning, however, can also occur without feedback, attention or task relevance (Watanabe, Náñez, & Sasaki, 2001) although studies show that feedback increases the speed of learning at least in a vernier task (Herzog & Fahle, 1997). A major drawback of the above model of vernier visual hyperacuity is that it depends only on feedback to learn.

Another approach binds perceptual learning to regularities in the visual stimuli. Such networks do not need top-down feedback and usually rely on the change of the network representations taking part in the processing of this input.

Recurrent networks are a common type of unsupervised connectionist models. Recurrent networks rely on physiologically plausible horizontal connections and Hebbian learning (Hebb, 1949). One example of a recurrent network is the model of perceptual learning in a contrast discrimination task (Adini, Sagi, & Tsodyks, 2002). Adini et al. describe two subpopulations of excitatory and inhibitory neurons that are fully connected. The model relies on Hebbian learning – synaptic modifications are activity-dependent. Thus, this model does not require feedback (teacher signal).

One drawback of this model is that it cannot easily simulate the experimentally demonstrated task-specificity of perceptual learning. Another shortcoming is that the model is quite task-specific – it is not clear how it would behave in other tasks, even in low-level tasks like vernier discrimination. In addition, most experiments show that feedback can influence perceptual learning, even if it is possible without feedback. Thus, a model of perceptual learning should be able to include feedback influences without relying only on supervised learning mechanisms. A basic restriction for a perceptual learning model should be its ability to learn both with and without feedback. These the two different scenarios are also a very important source of predictions that a model can make.

A physiologically plausible multichannel reweighting model of perceptual learning (Petrov, Dosher, & Lu, 2005) was developed along the lines of the reweighting theory (Dosher & Lu, 1998, 1999). The model is able to operate both in supervised and unsupervised mode and both types of simulations in an orientation task were compared successfully with experimental data from humans for the same task (Petrov et al., 2005; Petrov, Dosher, & Lu, 2006). In this model feedback is represented as activation of a feedback unit which provides additional input to a Hebbian learning decision making nodes only when teacher signal is available. The model is able to operate without feedback but when feedback is available, it is successfully used by the model to facilitate and speed up the learning of the network.

Theories and models addressing position specificity of perceptual learning

Probably the first more general theory which attempts to explain the experimental results, ranging from full position specificity to complete translation invariance of perceptual learning, was the Reverse Hierarchy Theory (Ahissar & Hochstein, 1997, 2004). According to this theory, perceptual learning first starts to occur at higher perceptual levels, where more general information is extracted and processed. If this, however, doesn’t lead to performance improvement, gradually lower-level structures are included in the learning process in order to allow for more specific information, needed for tasks like vernier discrimination or orientation judgment, to be processed.

The reweighting theory of perceptual learning has addressed the issue of position specificity only very recently (Dosher, Jeter, Liu, & Lu, 2013) by elaborating the reweighting model in an Integrated Reweighting Theory of perceptual learning. The updated reweighting model includes lower-level position-specific units as well as mid-level location independent units. All units are connected to a decision making unit and the experimentally observed variations in perceptual learning tasks are explained with changing the weights (reweighting) between the different types of representation units and the decision making node. The model includes connections between low-level and the mid-level representations, but these connections don’t seem to learn (Hebbian learning is described only between different representation types and the decision making unit, but not between the low-level and mid-level representations). The model is focused mainly on explaining position and object specificity of perceptual learning in orientation tasks but can in principle account for a variety of other tasks and phenomena too (including the role of feedback on perceptual learning). One important aspect which remains unclear, however, is the nature of the connections between low-level location specific and mid-level location invariant connections.

Why, how and when are these connections formed if not through the same mechanism of Hebbian reweighting remains unclear – this aspect might not be so important for a model of orientation judgment, but it is crucial for understanding categorical learning in visual objects.

Chapter IV. Categorical perceptual learning

The formation of conceptual structures through perception was tested in a series of experiments with infants by Quinn and colleagues (Quinn, Schyns, & Goldstone, 2006; Quinn & Schyns, 2003). The authors found that the acquisition of a new visual category by infants can lead to perceptual grouping that is not consistent with the Gestalt principles of good continuation. That is, after repeatedly presented with category exemplars like those on Figure 3B infants learned the "packman" form Figure 3C as a diagnostic part of the objects. As a result of this the children interpreted the ambiguous objects on Figure 3A as containing a packman and a polygon. Evidence for this segmentation comes from the habituation paradigm – after habituation to the objects in Figure 3A, the children reacted to the circle as a novel object.

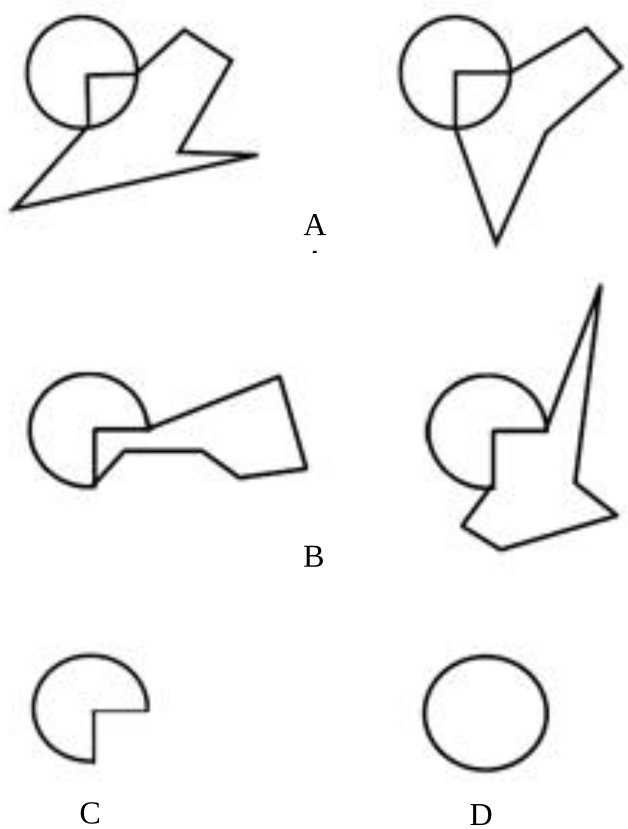


Figure 3. Stimuli from Quinn and colleagues, 2006. Figure based on (Quinn et al., 2006, fig. 1–2)

In contrast, the infants who did not receive prior category training with the objects from Figure 3B interpreted the objects from Figure 3A as containing a circle and polygon – after habituation the novel object for them was the packman in Figure 3C. The segmentation of the ambiguous objects from Figure 3A into a circle and polygon is consistent with the Gestalt principles of perceptual grouping. However, Quinn and colleagues have shown that if a new category is formed even the strong Gestalt principles can be overruled by the conceptual system which begins to interpret the objects on Figure 3A as exemplars of the newly formed category and thus to “see” them not as overlapped circles and polygons, but as objects containing the characteristic packman part.

These results have strong similarity to processes in perceptual learning. The possible mechanisms for the empirical findings above are discussed in detail in the context of the new model of categorical perceptual learning, presented in Chapter VII.

Categorical influence on the perception of characteristic parts of objects was also found by Pevtzow and Goldstone (1994). Participants in the study were presented with distortions of the four objects (A, B, C, D) shown in Figure 4. These four objects were divided in two categories and participants in the study had to learn to classify them correctly as one of the two categories or a third category containing distracter objects. Participants learned the correct categorization with the help of a feedback after each choice that indicated if their responses were correct or wrong. One group of participants were trained that objects A and B belonged to one category and objects C and D belonged to the other (horizontal categorization rule), while another group was trained that A, C belonged to one category and B, D to another (vertical categorization rule).

As seen in Figure 4, the horizontal categorization rule promoted the two parts on the right as characteristic for the two categories, whereas the vertical rule promoted different parts (those on the bottom) as useful for correct categorization.

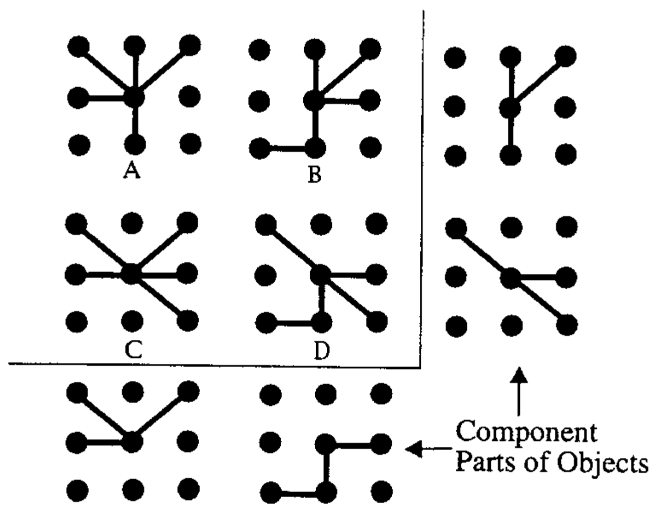


Figure 4. Stimuli used by Pevtzow and Goldstone (1994). The four objects, A, B, C, and D, could be categorized according to two different categorization rules. When A and B are placed in one group, and C, D are placed in the other, the component parts on the right are diagnostic. When A and C are placed in one group, and B and D are placed in the other, then the parts on the bottom are diagnostic. The figure is taken from (Pevtzow & Goldstone, 1994)

The participants in the study developed conceptually influenced detectors only for the diagnostic parts of the objects. This was demonstrated by the participants’ reaction times in the consequent task to find whether a part (diagnostic vs. control non-diagnostic) is present or absent in an object (part-whole task). Participants responded faster to the parts that were diagnostic for the previous categorization task than to those that were not.

These studies as well as many others (Goldstone & Barsalou, 1998; Goldstone, 1998) strongly suggest that perceptual learning is not a process isolated from the general cognitive system but modifies and is modified by other processes like conceptual learning. Mounting evidence from psychophysics and neuroscience demonstrate the same principle of complex intertwining of top-down and bottom-up processes in perceptual learning over and over again even for the very simplest of tasks in perceptual learning (Sagi, 2011).

Deliberate focus on higher-level object categorization from the point of view of perceptual learning is needed to better understand this interaction between visual categorical learning and perceptual learning in terms of quantitative computational mechanisms and underlying neural substrates.

Chapter V. Research Goals and Methodological Approach

Perceptual learning in psychophysics is focused mainly on several tasks like vernier acuity or line orientation, since these seem to involve very basic, low-level learning, different from other forms of learning (Fahle & Poggio, 2002; Sagi, 2011). At the same time learning to categorize visual objects has also been argued to involve perceptual learning (Goldstone, 1998).

The main goal of this dissertation is to bring together the two fields – the cognitive science approach to understanding complex cognitive processes like categorization which are basic for humans and the psychophysical approach which attempts to systematically study the process of perceptual learning and map it to the corresponding brain structures.

In order to achieve the above goal, a classical categorization problem – forming visual categories based on characteristic elements - is studied through the prism of perceptual learning. The main questions this research tries to unravel are:

1. What are the underlying mechanisms which determine visual categorical learning?
2. Can these mechanisms and the resulting changes in performance be simulated with a physiologically plausible computational model?
3. What are the lowest-level learning structures, involved in visual categorical learning?
4. Can categorical learning of this type indeed be considered to involve perceptual learning?
5. How does the nature of the task (i.e. not a "low-level" task) interact with learning?
6. Are there explicit (conscious) strategies during categorical learning of simple objects and if so, what is the interaction between these strategies and perceptual learning?

In order to answer the first two questions, a new computational model of visual categorical learning will be developed. The new model is based on earlier simulations (Goldstone, 2000) some of which are replicated with the original modelling framework and the assumptions, advantages and drawbacks of the earlier approach are studied and discussed in detail.

The restrictions for the new model are: to be as physiologically plausible as possible; to be capable of simulation both supervised and unsupervised categorical learning; to remain simple enough so that its mechanisms can be studied, understood and related at least schematically to information processing in the brain. The simplicity restriction means that the model's basic assumptions and predictions (in its intended scope) should be clearly visible and therefore – falsifiable.

To answer questions 3, 5 and 6, several experiments are conducted which study the positional specificity of visual categorical learning of simple objects. Typical tasks in visual perceptual learning are shown to elicit position specific learning which doesn't transfer to new locations on the retina (Dill, 2002). While there are conditions when perceptual learning could be transferred successfully (Xiao et al., 2008), the positional specificity or possible translation invariance of the learning remains a benchmark for any learning task which might involve low-level perceptual learning – if visual categorical learning indeed involves perceptual learning, then it should exhibit at least to some extent the position specificity found in other perceptual learning tasks. On the other hand – if this learning is fully invariant with regard to position on the retina, then it might involve only higher brain structures and therefore not involve the same type of perceptual learning which is observed in tasks like vernier acuity or orientation judgment.

Chapter VI. Physiologically Plausible Modeling and CPLUS Replication

The models tested and replicated in the dissertation, as well as the final modeling framework, which is put forward, rely on two basic, physiologically plausible mechanisms – Hebbian learning (Hebb, 1949) and Competitive learning (Rumelhart & Zipser, 1985).

Hebbian learning (1) states that the change in the connection between two connectionist units $\Delta W_{i,l}$ is a function of the degree of activity of the two units - A_i and A_j. In the equation below α is a constant that describes the learning rate.

$$(1) \qquad \Delta W_{j,i} = \alpha A_i A_j \quad ,$$

This learning rule is considered to be quite physiologically plausible because it closely mimics the changes in real neuronal circuits. For this reason, the Hebbian learning rule is still frequently used in connectionist modeling including in some of the most recent and influential theories and models of perceptual learning (Dosher et al., 2013; Petrov et al., 2005).

Competitive learning was introduced by Rumelhart and Zipser (1985) as a Hebbian-based form of learning. The basic idea behind this learning is that the units compete with each other for their right to learn. When a pattern is presented, the most active unit "wins" and its weights evolve in such a way that this unit will be activated even more by the same pattern next time it is presented. The less active units "lose" and their weights remain the same for this trial.

The simple possible learning rule (as originally described by Rumelhart and Zipser, 1985) for a competitive unit *i* was:

$$(2) \qquad \Delta w_{ij} = \begin{cases} 0 & \text{if unit } i \text{ loses on stimulus } k \\ g\dfrac{c_{jk}}{n_k} - g w_{ij} & \text{if unit } i \text{ wins on stimulus } k \end{cases} \quad ,$$

where Δw_{ij} is the degree of weight change in the connection between unit *i* and the lower level input unit *j*, c_{jk} is the activation of input unit *j* for stimulus pattern S_k – it is equal to 1 if the unit is active and zero otherwise; n_k is the number of active units in pattern S_k and w_{ij} is the current weight between units *j* and *i*. There are possible variations of this rule, where the losing unit also learns, although slower than the winning unit.

A model of conceptual and perceptual learning by unitization and segmentation (CPLUS) has been put forward by Goldstone (2000) and later improved (Goldstone, 2003). The model simulated the results from Pevtzow and Goldstone (1994) and its later version is able to extract perceptual building blocks from the given stimuli. The model is based on modified completive learning which is influenced by a top-down categorization signal.

A replication of the first version of CPLUS suggested that successful detector formation by the categorization network strongly depended on the number of competitive units and on the chosen input space which was problematic for the model and for its theoretical predictions.

The principles used in the segmentation network assumed hard-wired perceptual constraints corresponding to Gestalt principles of "goodness" of parts like continuity, connectedness, etc. Goldstone himself, however, previously demonstrated that these principles of "goodness" are not inherent to the objects as previous assumed by Palmer (Palmer, 1977, 1978), but can be overruled by (perceptual) learning (Pevtzow & Goldstone, 1994). Therefore, an important challenge for a perceptual learning model of categorization would be to show how these principles of "goodness" actually could emerge, rather than hard-coding them in a model.

A new model, based on the general ideas of CPLUS, was developed and studied.

Chapter VII. A new model for perceptual and conceptual learning

The new model for perceptual learning consists of two main layers and an artificial input retina (Figure 5). The first layer is based on the competitive learning paradigm (Rumelhart & Zipser, 1985). However, units compete only for a small part of the input - that is, each unit has a receptive field and competes only with other units with the same receptive field (units from the same inhibitory cluster). In the current implementation of the model there is no overlap between receptive fields. Competing units are organized in inhibitory clusters - two units with the same receptive field cannot be active at the same time. Only the winner for this receptive field is active. Therefore, only one unit from a particular inhibitory cluster can be active at a time. A competitive unit is connected with horizontal (lateral) Hebbian connections to all units from the other inhibitory clusters. The horizontal Hebbian connections link the parts of an input pattern in terms of coactivation of the competitive units that are specialized to those parts.

The activation of a competitive unit is computed in two time-steps according to equations (3) and (4):

$$A_{i,k}^{d}(t) = \sum_{j=1}^{n} I_{j,k}^{d} W_{i,j}^{d} \tag{3}$$

$$A_{i,k}^{d}(t+1) = A_{i,k}^{d}(t) + \eta \sum_{\substack{p=1 \\ p\neq d}}^{c} \sum_{l=1}^{s} W_{i,l}^{d,p} A_{l,k}^{p}(t) \ , \tag{4}$$

where $A_{i,k}^{d}(t)$ is the activation of unit *i* from cluster *d* in moment *t* when input pattern *k* is presented, $I_{j,k}^{d}$ is the activation of input pixel *j* from receptive field *d* for pattern *k*, $W_{i,j}^{d}$ is the weight of the connection between unit *i* and pixel *j*, $A_{l,k}^{p}(t)$ is the activation in moment *t* of competitive unit *l* from cluster *p* for pattern *k*, $W_{i,l}^{d,p}$ is the weight of the horizontal connection between unit *i* from cluster *d* and unit *l* from cluster *p*, *n* is the number of pixels in receptive field *d*, *s* is the number of competitive units from cluster *p*, and *c* is the number of clusters.

The output layer is fully connected to the competitive layer with feedforward Hebbian connections (there is no top-down spread of activation coming from the output layer).

The activation of an output unit is computed according to formula (5)

(5) $$A_{o,k} = \sum_{i=1}^{m} A_{i,k} W_{o,i},$$

where $A_{o,k}$ is the activation of output unit *o* when input pattern *k* is presented, $A_{i,k}$ is the activation of competitive unit *i* for pattern *k*, $W_{o,i}$ is the weight of the Hebbian connection between output unit *o* and competitive unit *i* and *m* is the total number of competitive units, which is equal to the product of the number of competitive units in a particular cluster (*p*) and the number of clusters (*c*): $m = pc$.

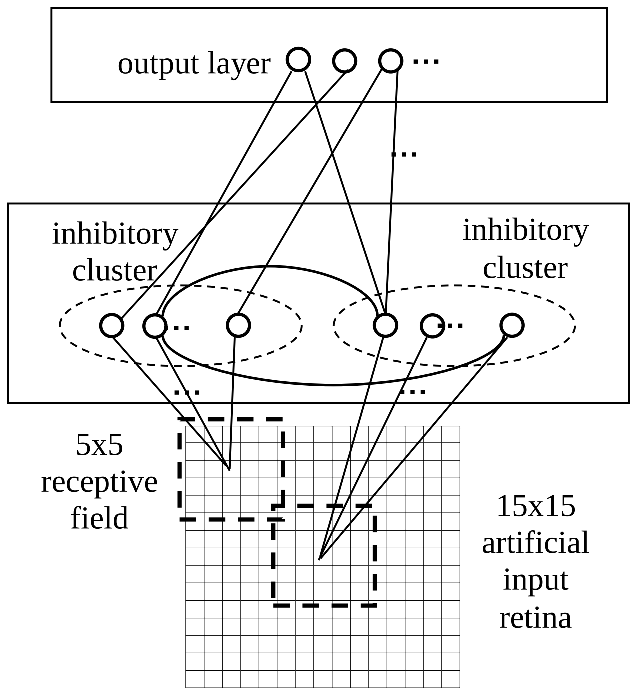


Figure 5. The new model for perceptual learning. Only some of the connections are shown for visualization purposes. See the text for full details.

In the following simulations, *s* is the same for all clusters, that is, the number of competitive units in the different clusters is constant. The parameter *η* is set to 0.1 and represents the smaller contribution of the horizontal connections compared to the bottom-up activation. The winner from each cluster is determined as the most active unit inside the cluster. The output units have sigmoid activation functions.

Learning for the connections between an input receptive field and the competitive units from the corresponding inhibitory cluster follows the classical formula:

(6) $$\Delta W_{i,j}^{d} = \begin{cases} M(I_{j,k}^{d} - W_{i,j}^{d}) & \text{if unit } i \text{ loses on stimulus } k \\ L(I_{j,k}^{d} - W_{i,j}^{d}) & \text{if unit } i \text{ wins on stimulus } k \end{cases},$$

where *L* is the learning rate for the winning unit (0.1 for all simulations), *M* is the learning rate for the losing unit – it is set to 0.001 for all simulations. $I_{j,k}^{d}$ is the activation of the retina pixel *j* from receptive field *d* when input *k* is presented, and $W_{i,j}^{d}$ is the weight between pixel *j* from receptive field *d* and competitive unit *i*. The stimuli are presented as activation patterns on the retina, where each pixel is either 1 (active) or 0. Activation of competitive units is normalized so that the winning unit's activation is 1 and all the losing units from the cluster sharing the same receptive field are inhibited to have zero activation. The horizontal Hebbian weights learn according to the Hebbian rule:

(7) $$\Delta W_{i,l}^{d,p} = \alpha A_i^d A_l^p - D,$$

where α is the learning rate, is the activation of unit i from cluster d, is the activation of unit l from cluster p, and D is the decay rate of the weights.

The Hebbian weights between the competitive layer and the output layer learn according to the same rule as the horizontal connections, with the exception that they have different decay and learning rates. All Hebbian weights were set to zero at the beginning of a simulation.

The network learns after each pattern is presented. The competitive layer corresponds to lower-level, position specific cells with smaller receptive fields that cover only parts of an input, while the output units correspond to higher-level perceptual structures with larger receptive fields. The model doesn't include a decision making unit, rather than that, the top-down signal from higher-level structures (such as decision making units) is represented by top-down feedback for the output layer (effectively reduced to simple "clamping" one output node or another).

Two types of simulations are possible with the described model. The first type corresponds to learning without feedback. In this operational mode, the output layer is activated at random since no teacher signal is available. In other words, this is unsupervised learning of the competitive layer, based only on the characteristics of the input space. When feedback is available, a particular pattern of activation appears on the output layer as a teacher signal. This signal represents the influence of higher-level conceptual processes on learning.

The unsupervised learning of the competitive layer alone was simulated with stimuli similar to those used in Quinn and Schyns (2003) and Quinn et al. (2006) and presented in Chapter IV (Figure 3).The experimental results strongly suggested that unsupervised learning is capable of overriding Gestalt laws of organization such as good continuation if the prior learning history supports an alternative organization. Using an unsupervised model to simulate such empirical results from infants seems like a natural correspondence, given that infants in the first few months of life do not receive instruction on how to organize their visual experiences.

The current model can provide a computational account for these empirical findings. The competitive layer is capable of extracting elements and statistical dependencies from the input structure even if no feedback is available. Thus, the gestalt law of continuity was simulated with presentation of simple forms at different positions on the retina. Ten such patterns (three vertical lines, three horizontal lines, and four circles) were presented in random order for 2000 cycles. This pre-training phase simulated the infant's perceptual experience prior to arrival at the laboratory and conceivably corresponds to the experiences of young infants as they encounter visual patterns in the environment. We were interested in the ability of the model to acquire perceptual constraints from commonly occurring patterns instead of explicitly building in the good continuation principle. This could also be interpreted as the evolved representation of naturally occurring statistics in visual patterns (Olshausen & Field, 1996).

The input retina consisted of 225 pixels organized in a 15x15 square matrix. There were 9 non-overlapping square 5x5 receptive fields with 8 units in an inhibitory cluster competing over each of the receptive fields, which makes for a total of 72 nodes in the competitive layer. The learning rate of the horizontal Hebbian weights was 0.05 and the decay rate was set to 0.009. After the pre-training phase, some of the competitive units specialized for parts of lines, while others specialized for arcs of a circle. Then an ambiguous pattern (Figure 6A) was presented. This portion of network training and testing corresponded to the first familiarization test phase in the study with infants, when similar patterns each consisting of an overlapping circle and a polygon were presented, which led to the segmentation of the circle and the polygon by infants. The ambiguous pattern given to the model activated four “arc” and two “line” nodes from the competitive layer, thus forming a good, continuous circle and some parts of a polygon which was consistent with the infants’ behavior. The activation pattern over the competitive layer is visualized in Figure 6B with the following algorithm – each pixel represents the weight between this pixel and the competitive unit multiplied by the competitive unit’s activation.

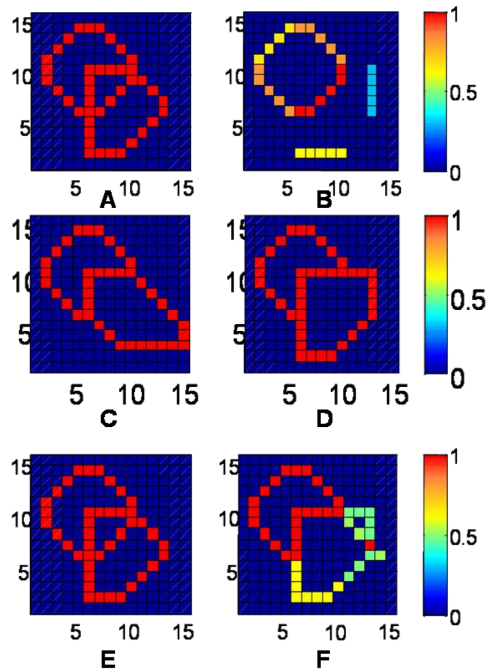


Figure 6: Unsupervised learning simulation

The same network was fed for 200 cycles with two patterns containing “pacman” shapes (Figure 6C, Figure 6D) and again was presented with the ambiguous pattern Figure 6E. This corresponded to the two-part procedure (Chapter IV) in which the infants were first presented with pacman shapes and subsequently with circle shapes. Once again the model behavior was very similar to what the experimental results suggested. This time the pacman shape was strongly active and some polygon segmentation appeared but was less active than the pacman (Figure 6F). The pacman shape was actually represented by three competitive units specialized for arcs and one specialized for an angle. The “arc” units were initially connected to the fourth arc unit which completed the active circle from Figure 6B; however, after the patterns containing the pacman shapes were repeatedly shown to the network, the angle unit became more active than the arc unit over the same receptive field (because of the strengthening of the lateral connections “holding together” the co-occurring elements of the packman), which led to the angle unit winning for this receptive field. This could be interpreted as a spontaneous formation of a virtual pacman shape detector that is constructed from smaller low-level representations of three arcs and one angle segment.

The new model was also able to replicate qualitatively the simulation results from Goldstone (2000) without relying on built-in perceptual constraints. As mentioned earlier, the competition in the new model is for small parts of an input inside a receptive field, instead of competition for the whole input. This leads to a somewhat different interpretation of a detector – in the present model a detector is composed of several smaller competitive units from different receptive fields that form together a coherent shape detector over the whole input retina.

In the following simulations, the formation of such detectors was influenced not only by the input properties as in the unsupervised learning but also by a conceptual teacher signal that led to the formation of categorization-relevant detectors at the output layer of the network. This happened through reweighting the feedforward Hebbian connections between the competitive and the output layer. A teacher signal was directly presented as a pattern of activation on the output layer during the supervised training. This was done for simplicity since the influence of higher-level categorization or judgment structures can be simulated in different ways (e.g. see Chapter VI, equation (3) for the mechanism that was used by Goldstone, 2000)

A 256 square 16x16 pixel retina was used; competitive units' receptive fields were square 8x8 non-overlapping matrices, which yielded a total of four receptive fields. Each inhibitory cluster consisted of 4 units competing with one another. The output layer had two units. Learning rate for the output Hebbian weights was set to 0.1 and the decay rate was 0.04. The horizontal Hebbian connections had the same learning and decay rates as in the previous simulation.

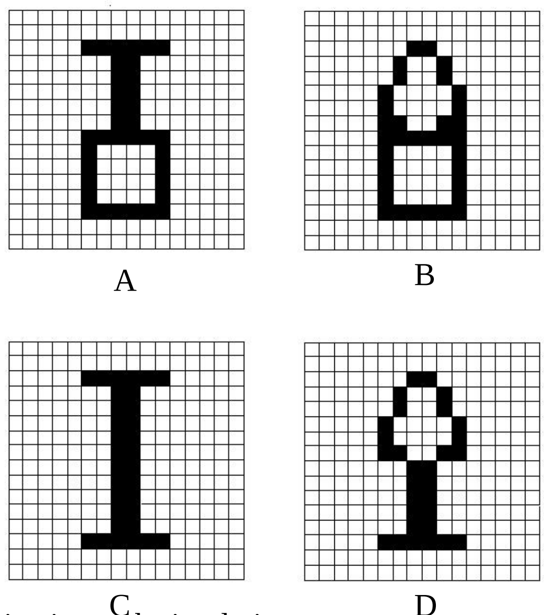


Figure 7. Inputs for the categorization task simulation

The number of units in the network and the size of the receptive fields was changed in order to test a more difficult scenario of presentation of the stimuli (each stimulus falls in four different receptive fields). The network which was used in the unsupervised simulation led to the same results, as the presented here.

Four input patterns were presented to the network (Figure 7). First, feedback was given to the network that Figure 7A and Figure 7B belong to one category (1, 0) and Figure 7C, Figure 7D belong to another (0, 1). With this horizontal categorization rule, 50 cycles were run with the four input patterns presented in a random order during each cycle. The mean squared error of the output units displayed a rapid decrease (Figure 8B). The network learned to distinguish patterns A and B as members of one category from C and D belonging to another. That is, when Figure 7A or Figure 7B were presented, output unit 1 was active and unit 2 was not. On the contrary, when Figure 7C or Figure 7D were presented, output unit 2 was active and unit 1 was off. The two output units can be considered detectors for the two categories. The learned weights of the connections between the competitive layer and each of the two output units are shown in Figure 7A. Only two of the competitive units had positive weights to output unit 1 and the other two had positive weights to output unit 2. Thus, the output units had learned to ignore the responses of those lower-level nodes that were not relevant for categorization and combined together those parts which were relevant, forming diagnostic shape detectors (Figure 7C, Figure 7D).

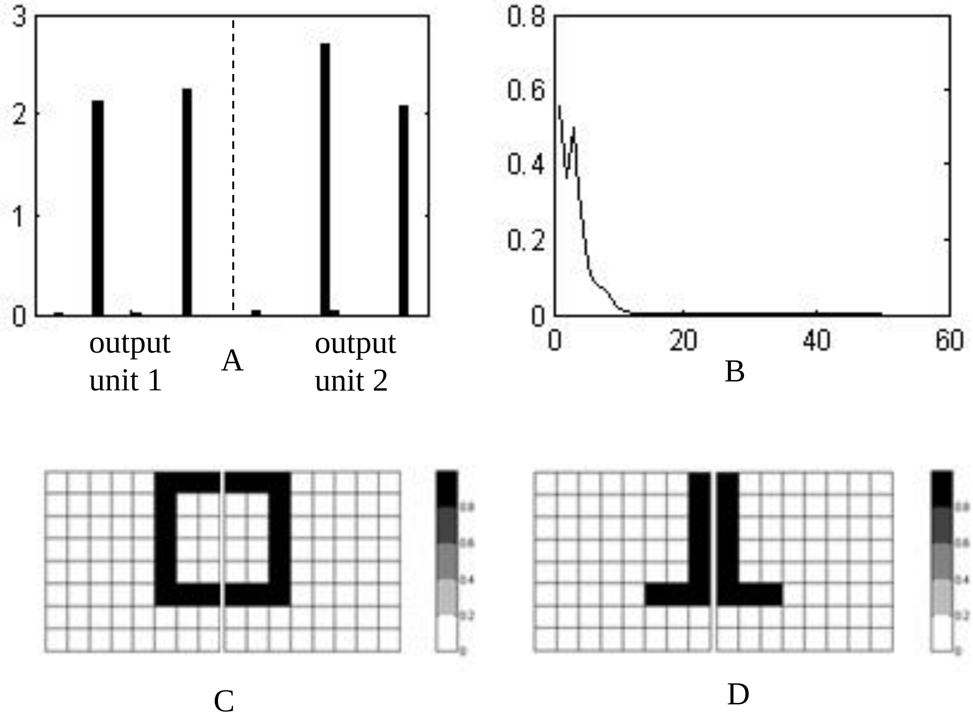
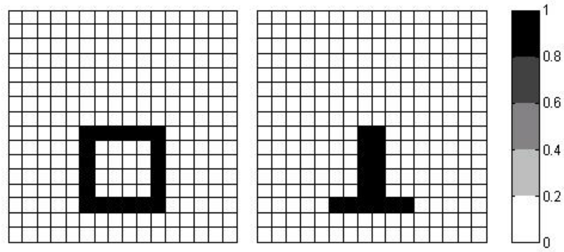


Figure 8. Panel A – weights between the competitive layer and the two output nodes. Panel B – mean square error for the output nodes. Panel C – the pixel-to-unit weights for the two competitive units with positive weights to output unit 1. Panel D – the pixel-to-unit weights for the two competitive units with positive weights to output unit 2.

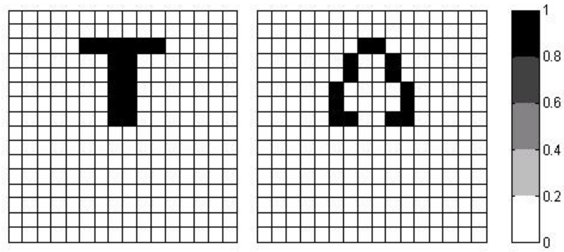
The formation of the detectors was not influenced by the number of lower-level competitive units that participated in the shape representation. The result was the same with smaller 4x4 receptive fields. This change led only to the same diagnostic shape detectors being composed of four instead of two competitive units. The competitive units participating in a detector's representation were specialized for small input patterns contained within their receptive fields. The global representation activated by the whole input pattern, however, was a continuous shape honoring the Gestalt principle of Good Continuation.

In a second simulation, a vertical categorization rule was applied to a network with identical parameters. This time patterns Figure 7A and Figure 7C were from the same category (1, 0) while patterns Figure 7B and Figure 7D were from the other (0, 1).

The results from the second simulation are compared to the outcomes of the first simulation in Figure 9. For visualization purposes the output layer weights are multiplied by the competitive layer weights, which represent the participation of each pixel in the diagnostic shape detectors that were formed at the output layer.



Horizontal categorization rule – AB, CD



Vertical categorization rule – AC, BD

Figure 9. Detectors built according to a horizontal and vertical categorization rule.

The same patterns led to the formation of different detectors when the vertical categorization rule was applied. This result was very stable over simulations and replicated the type of results reported by Pevtzow and Goldstone (1994).

Inspection of all specialized competitive units showed that there was no difference in their representation after the vertical and horizontal rule simulations. This means that the general structure of the input space was captured every time by the competitive units. Correct categorization was due to the formation of a diagnostic shape detector at the output layer through reweighting the “read-out” Hebbian connections between the competitive and the output layers.

One conclusion from the simulation results is that there are automatic low-level changes that capture the structure of visual stimuli irrespective of the given task. However, when feedback is available, a more complex shape representation is constructed at a higher-level to accommodate the task requirements

Another interesting conclusion comes from the unsupervised behavior of the network. The simple mechanism of competitive learning, reinforced by the horizontal Hebbian connections, is able to extract perceptual categories that are statistically present in the input space. This strongly supports empirical findings that Gestalt principles of perceptual organization can at times be overruled by category learning. The model also suggests a way in which even certain Gestalt principles like continuity can be learned, rather than built-in, as a consequence of experience with a learning environment that includes visual patterned stimulation (Quinn & Bhatt, 2005; Spelke, 1982).

The presented new model puts forward three simultaneously acting learning mechanisms. Competitive learning is essential for building lower-level local representations while horizontal connections lead to the formation of meaningful parts composed of several competitive units. Thus, the horizontal connections are responsible for the formation of more global perceptual representations based on the environmental regularities and co-occurrences. The output layer in the model is interpreted as a mid-level perceptual layer which integrates activation from the lower-level units though selective read-out from lower-level representations. This is achieved through feedforward Hebbian connections – a mechanism very similar to the one used in the Reweighting Theory. This layer can receive top-down signal that influences learning. Again this is very similar to the Integrated Reweighting Theory (Dosher et al., 2013; Petrov et al., 2005), however the interpretation here for the output layers is rather as a mid-level perceptual representation than a higher decision making structure.

The neural substrates for this model are difficult to pinpoint. There is evidence from single-cell recordings that there are cells tuned for complex representations of parts and wholes in the Inferotemporal cortex of monkeys (Baker, Behrmann, & Olson, 2002). Therefore the output layer of the network could very well correspond to structures in the Inferotemporal cortex. The competitive layer is more difficult to connect to a particular level of processing. The level of abstraction used in the model doesn't allow for more than speculating it is probably above V1 and V2, but the focus of the model is rather on simulating plausible mechanisms than reconstructing the visual pathways. Still, the question of the level of changes in categorical learning of the type simulated here is very important. If such learning is achieved through reweighting of connections between fixed representations in the visual cortex and higher brain areas, can this learning be considered the same phenomenon as performance improvement in tasks like vernier acuity or line orientation?

The following chapter attempts to address this question, by studying the position specificity in the visual field of the same type of categorical learning that is simulated with the model.

Chapter VIII. Experiments on Position Transfer in Categorical Perceptual Learning

In order to study the level of processing for categorical learning of simple objects, one way is to explore whether this type of learning is position specific to the training location in the visual field, as many of the typical perceptual learning tasks are (Chapter II). It has been argued that perception of characteristic parts for stick-line object categories is facilitated by the formation of "perceptual detectors" (Goldstone, 2000; Pevtzow & Goldstone, 1994). While there is strong evidence from neuroscience that similar detectors for parts could indeed be formed through learning in the inferotemporal cortex of monkeys (Baker et al., 2002), there aren't many experiments on categorization of objects in the classical approach for studying perceptual learning. Whether this particular type of categorical learning (categorization of simple objects based on characteristic elements) is invariant or position specific has never been tested to the best of my knowledge.

In order to systematically study the position specificity (or translation invariance) in the visual field of categorical perceptual learning, several experiments were conducted. Two pilot experiments showed that the task was quite difficult for participants when they didn't know the categorization principle (characteristic elements). Those who succeeded in learning to categorize, always knew explicitly about this principle and used it as a strategy. Those who couldn't finish the pilot experiment (i.e. to learn to categorize) usually had misleading strategies. In order to provide similar experimental setting to all participants and test the simplest possible scenario, an explicit categorical learning experiment was conducted.

Experiment 1 – An easy, explicit perceptual learning task

This experiment was conducted as the simplest possible translation invariance study on categorization of simple figures with categories defined as characteristic parts of the objects. The goal of this experiment was to serve as a baseline for further studies and experiments, using similar categorical learning tasks. The learning task here was relatively simple (completing the experiment took between several minutes and half an hour); learning was fully explicit (manipulated by the instructions and tested in a post-study questionnaire). If this explicit categorical learning involves only higher areas, learning should be fully translation invariant.

Method

In this experiment performance was tested as a function of the distance from the training location. Two transfer locations (at distances of 2 and 4.5 degrees of visual angle) were compared to the training location (control condition). Performance was measured as the percent of correct responses in a block of 20 trials and the response time for giving an answer (within-subject dependent variables). The order of presentation of the control and two training conditions was fully counterbalanced. Participants were randomly assigned to two different categorization rules (for the same stimuli set), thus there were overall 12 experimental setups implemented as separate E-prime e-run scripts (2 categorization rules x 6 possible orders of stimuli). Participants were randomly assigned to one of the 12 setups.

The training position was located 2.5 degrees to the top and to the left from the fixation cross. The tested transfer locations were shifted by 2 and 4.5 degrees of visual angle from the training position. All three positions were at the same distance from the fixation cross and were not symmetric in any way.

Participants

Participants in the study were undergraduate students at New Bulgarian University, who were recruited with print ads spread around the university and were paid for participation in the study. 24 participants finished successfully the training phase and were tested for translation invariance of learning. One outlier was excluded from the analysis because of a very low score (55%) in the control position. All other participants scored above 85% in the control position.

Apparatus

For the translation invariance paradigm we used a Tobii 1750 remote eye tracker to ensure that participants were looking at the fixation cross in the center of the screen. The Tobii 1750 eye tracker provided a gaze sample every 20 ms and an E-prime 1.2 script checked in real time if the gaze was indeed at the fixation cross. Participants had to look continuously at the cross for 200 ms for the stimulus presentation to be triggered. The eye tracker also checked that the viewing distance was between 58 and 62 cm. When these rules were satisfied, the cross disappeared from the screen and at the same time the stimulus appeared for 100ms at the appropriate location. Participants' responses and reaction times were collected using a Serial Response Box (SRBox), which guaranteed very low error in reaction time estimation (<1ms). Stimuli were presented on a 17' TFT screen (with a matrix response time of 25 ms) built in the Tobii 1750 eye tracker. We used the native screen resolution - 1280x1024.

Stimuli

The stimuli were constructed in a fashion similar to Pevtzow and Goldstone's (1994) stimuli – they were composed of black line-segments connecting dots on a 3 by 3 grid (Figure 10). The line segments were 2 pixels wide (approximately 0.05 degrees of visual angle) and the points were with diameter 3 pixels. The overall stimulus was 80 pixels high and 80 pixels wide (corresponding to approximately 2 degrees of visual angle). There were four basic stimuli which were distorted by adding a random line to the stem. The random lines were added to all possible places except if the added line resulted in any closed parts or detached lines. Categories were defined by a pair of overlapping lines shared by category members (Figure 10). Participants were given at random one of two different categorization rules in order to avoid possible effects that were specific for the selected characteristic element. Participants in the experiment always saw one object at a time, composed of 5 lines (a four-line basic object plus a randomly added line). For every participant there were two categories and the same categorization rule remained throughout the whole experiment. The stimuli were presented on a white screen.

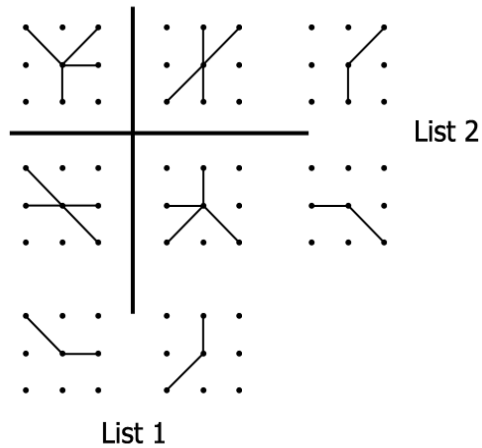


Figure 10. Stimuli for Experiment 1 - four basic stimuli and their characteristic elements when horizontal categorization rule is used (List 1) and when vertical categorization rule is used (List 2)

Procedure

Participants were seated in front of the eye tracker (at 60 cm viewing distance from the screen) and were told that they would take part in an experiment studying cognitive processes during learning. They were given instructions about their task (categorization) and were shown four sample stimuli with the characteristic elements marked (the marked elements were different from the actual ones used throughout the study). The participants were told that during the experiment they would have to find by themselves which correct characteristic elements determined the categorization, but that the elements would be the same throughout the whole experiment.

After reading the instructions, participants were briefly familiarized with the stimuli and the way of presentation – participants were asked to look at the fixation cross in the center of the screen and when their gaze was at the cross, a stimulus was presented for 100ms at the training position and then disappeared (there was no mask after it). During the familiarization, participants did not try to categorize the objects. They observed them passively. After the familiarization, participants were told they would have to categorize each object into either Category “1” or “2” by pressing the labeled buttons on the SRBox. After pressing the desired button, they received immediate feedback as to whether their choice had been correct or wrong. They were also shown after each trial their cumulative percent correct answers. Feedback was color-coded (correct answers in blue, incorrect in red) in order to make it easier for participants.

Trials were organized in blocks of 20. Participants had to score at least 90% correct answers during the 20 trials in order to continue to the next phase of the experiment. The percent of correct answers was computed after each trial and was displayed as feedback. After 20 consecutive trials, the participants were presented an additional message with their overall results for the last 20 trials. This message marked the end of a block and the beginning of the next one (the percent of correct answers was computed for each block).

When participants reached 90% correct answers (or above) for a block of 20 trials, they were congratulated and were told they had passed the first part of the experiment and that they would continue the same task with the same rules and objects, but the objects were to be presented at different places on the screen. After this instruction there were three blocks of 20 trials each, every block was presented at one of the three locations (training position, 2 degrees transfer and 4.5 degrees transfer). The order of presentation was fully counterbalanced across subjects – for 3 positions there were 6 possible combinations and every participant was randomly assigned to one of the 6 combinations.

At the end of the experiment, the participants were asked about their strategies and if they said they were looking for a diagnostic part, they were asked to draw it.

Results

The post-experiment questionnaire showed that all but one of the participants were able to explicitly define their categorization strategy - looking for the presence of a diagnostic part in the objects and draw the characteristic element for which they were looking. Thus, the learning was indeed explicit.

The mean percent of correct responses in all conditions was above 90% (Table 1). Still, a Repeated Measures ANOVA showed there was a significant difference between conditions: F (2, 46) = 7.813; $p<0.01$; $\eta_p^2 = 0.254$;

A Pairwise comparison of the means (with Bonferroni adjustment) showed that there was no difference between the control position and the 2 degrees transfer position ($p=0.647$). There was difference between the control and 4.5 degrees transfer positions ($p < 0.05$) and between the 2-degrees-transfer position and the 4.5-degrees-transfer position ($p < 0.05$).

Difference between the conditions was also found in the reaction times. Repeated Measures ANOVA showed slower reaction time for the 4.5-transfer-position: F (2, 46) = 3,476; $p < 0.05$; $\eta_p^2 = 0.249$.

Position	Mean, %	Std. Error
Control	97.1	0.8
Transfer 2 degrees	95.6	1.3
Transfer 4.5 degrees	91.9	1.7

Table 1: *Mean percent of correct responses per condition*

A Pairwise comparison (with Bonferroni adjustment) between the means shows that only the difference between the 2-degrees and 4.5-degrees conditions was significant: p<0.05; see Table 2 for mean RTs.

Position	Mean, ms	Std. Error
Control	830	54
Transfer 2 degrees	791	54
Transfer 4.5 degrees	919	66

Table 2: Mean reaction time per condition.

d' was also computed as a more reliable, criterion-independent measure than the percent correct responses – the pattern of results was the same for d'.

Results from Experiment 1 showed partial transfer of learning for the 4.5-degrees condition - there was a significant difference (in terms of both correct responses and response time) between the control position and the further transfer position (4.5 degrees shift). This could be interpreted as evidence that lower-level visual structures play a role in the changes that occur during learning even in this simple category learning scenario.

Another explanation, however, could be connected to selective spatial attention (Dill, 2002). Two possible mechanisms for attention effects were considered – the first one would be a fast, "spotlight" movement of attention. Such confounding was eliminated by fully counter-balancing the order of the training and transfer location. A Repeated measures ANOVA with the order of the positions as a between-subject factor, showed no effect of the order ($p = 0.86$) on performance or interaction between the factors "position" and "order" ($p = 0.7$). The significant difference between the positions was preserved in this model. Therefore, there is no reason to think that a certain order of testing the three positions contributed to the observed effect. Still, when attention is discussed in connection with position specificity of perceptual learning, usually what is meant is relatively long-lasting facilitation for any visual processing at the training location (Dill, 2002).

In order to control for spatial attention as well as for other possible confounding factors (like non-categorical learning) a second control experiment was conducted.

Experiment 2 – a control experiment without a categorical learning task

Experiment 2 was identical to the previous one with only the instruction changed - participants were shown the characteristic element in the instruction and were given the simple task to press Button 1, if the element was present in an object and Button 2, if it was not (a part-whole judgment task). This task imitated the explicit strategy that the participants in Experiment 1 had reported but did not involve any categorical learning.

Different characteristic elements were given to different participants at random. Everything else remained the same as in Experiment 1. If spatial attention or any other non-categorical learning related factor was the explanation of the results from Experiment 1, then the same pattern should be observed in Experiment 2.

Method and Participants

22 participants were recruited from the same pool as in Experiment 1.

Results

Participants achieved 90% correct answers very quickly (usually in the first training block). Therefore the task involved little or no learning and performance was near-perfect from the start. There was no significant difference between the control and transfer positions (Table 3).

Position	Mean, %	Std. Error
Control	96.8	1.07
Transfer 2 degrees	98	0.78
Transfer 4.5 degrees	95	1.14

Table 3. Mean percent of correct responses per condition in Experiment 2

The overall model (with Greenhouse-Geiser correction) showed a trend: F (2, 42) =2.558; *p=0.109*, η_p^2 = 0.109 but only because of the difference between the two transfer positions (the control being in the middle). Indeed, the Pairwise comparison (with Bonferroni adjustment) showed no difference between either of the transfer positions and the control one: *p = 0.512* for the difference between the control and the 2-degrees-transfer position and *p = 0.741* for the key comparison between the control and 4.5 degrees transfer, which was significant in Experiment 1. The difference between the two transfer positions was not significant (p = 0.184).

The reaction times in Experiment 2 were much faster, but showed a similar pattern to the results in Experiment 1 (Table 4).

Position	Mean, ms	Std. Error
Control	644	32
Transfer 2 degrees	643	28
Transfer 4.5 degrees	744	35

Table 4. *Mean response time per condition.*

Repeated Measures ANOVA showed slower reaction time for the 4.5-transfer-position: F (2, 42) = 15.248; *p < 0,001*, η_p^2= 0.42

The results from Experiment 2 showed no statistically significant difference for the percent of correct responses (unlike Experiment 1) but significant delay in the response time for the 4.5 degrees transfer (as in Experiment 1) - Figure 11.

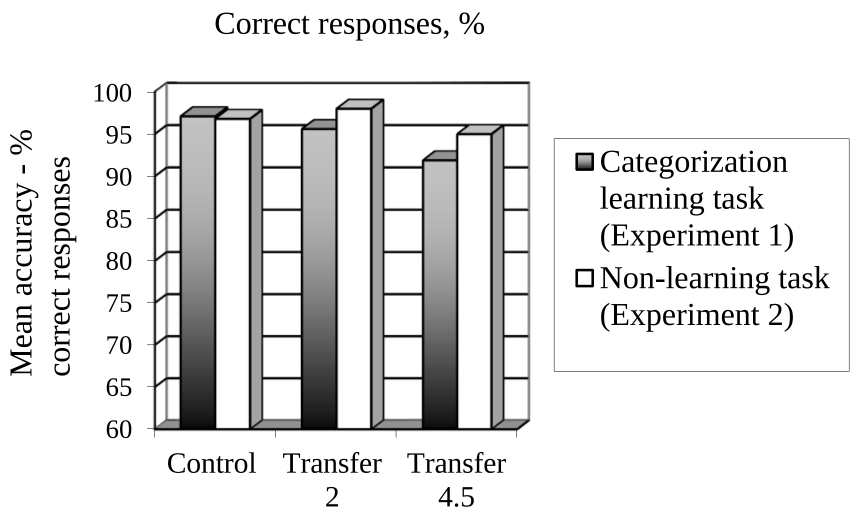


Figure 11. Percent correct answers for the control and two transfer positions - Experiment 1 and Experiment 2

The overall response times were much faster in the second experiment (see Figure 12), which indicates that the task was easier than the categorization in Experiment 1. Although the task imitated the strategy reported by participants in Experiment 1 (i.e. when a characteristic part is present press Button 1, when it is not – press Button 2) the overall slower response times in Experiment 1 show that the categorization task probably involved some additional cognitive processing.

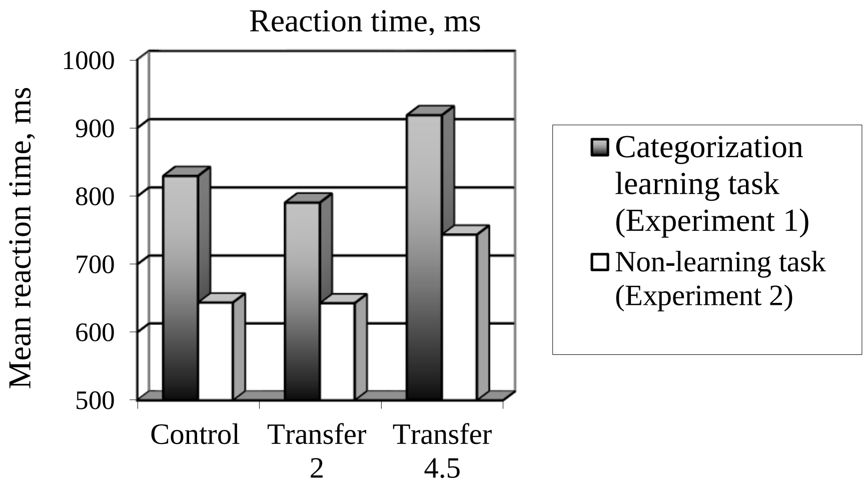


Figure 12. Reaction times (in ms) for the control and two transfer positions, for Experiment 1 and Experiment 2

Discussion of Experiments 1 and 2

The first experiment showed some evidence for involvement of low-level structures in the learning of simple object categories. Selective spatial attention or non-categorical learning could be another explanation for these results. The second experiment, however, showed that there was no effect of attention on position performance (in the percent correct responses) for a non-learning but otherwise identical situation. The delay in response time for the 4.5 transfer was similar to the observed in Experiment 1 and this effect could indeed be caused by selective spatial attention or another phenomenon. The percentage of correct responses, however, did not match the pattern from Experiment 1, indicating that in addition to attention, there is some spatial restriction for categorical perceptual learning. One possible interpretation of these results is that the learning process actually occurs simultaneously at different perceptual stages (as the model assumes as well).

Experiment 1 explored the question of whether explicit category learning of simple objects involves only strategic, high-level decision-making processes, or lower-level changes to how stimuli to be categorized are perceived. Experiment 2 served as a control for Experiment 1, using the same stimuli and a non-categorical task. The results in Experiment 1 showed almost full transfer of categorical learning (performance above 90% correct answers in all conditions) but significantly worse and slower performance at the farther transfer location than the control (training) one. Experiment 2 showed similar pattern in the reaction times for a non-categorical learning part-whole judgment task, but no cost in performance for moving from the training location to the transfer one.

A limitation of these experiments stems from the task itself – more complex stimuli and implicit learning of the characteristic elements which are difficult to search explicitly could have allowed for more elaborate controls. Here, the second experiment possibly suffers from a ceiling effect in performance. Still, the main goal of Experiment 1 was to test a simple categorical learning task, similar to Pevtzow & Goldstone (1994) learning task and similar to the simulations in the model. The results suggested that even in these "high-level" conditions, there might be a lower-level perceptual learning element.

According to the model, the formation of perceptual detectors through categorization is achieved by reweighting between lower-level, slow-learned representations and mid-level perceptual representations. Clearly, the experiments show more complex processing than that due to high-level reasoning – the existence of an explicit strategies for learning and the modification of these strategies. The formation of perceptual detectors for characteristic parts is a slower process which takes place together with higher-level reasoning (and possibly guided through top-down signals). Evidence for these could be found in the incomplete transfer in Experiment 1. Experiment 2 demonstrated that this effect is not simply an artefact from the experiments procedure and that Experiment 1 cannot be simply reduced to finding the correct strategy for categorization via higher-level reasoning.

To explore further the interplay between task difficulty and implicitness a third experiment was conducted. Using the original stimuli (made of 6 lines instead of 5, and with 3-line characteristic elements instead of two-line) and three categories instead of two (Pevtzow & Goldstone, 1994), should lead to participants' performance relying more on automatic processing and less on a simple strategy.

Experiment 3 – a more complex, implicit categorical learning task

The goal of experiment 3 was to explore the result from Experiment 1 in a more difficult, implicit categorization task. Difficulty was increased by using the original stimuli from (Pevtzow & Goldstone, 1994): object consisted of 6 lines instead of 5, characteristic elements consisted of 3 line segments instead of 2 and there were 3 categories instead of 2 – the third category did not have a characteristic element, therefore participants had to learn both categories 1 and 2 and select category 3 if objects did not belong to either of the other categories.

In Experiment 3, the participants were not given explicit instructions about the principle of categorization. They had to learn to categorize entirely by themselves. This attempted to induce implicit categorical learning – that is, to stop participants from using a reduced explicit strategy (looking for a particular part in an object). Possibly in that way more automatic bottom-up learning would be promoted - "knowing how" without "knowing what" (Fahle & Poggio, 2002).

Implicitness of learning was explored further by including another experimental condition which included presentation of stimuli with the characteristic element slightly thicker (i.e. marked) during the training phase. Since the stimuli were presented only briefly in parafoveal vision, it was reasoned that participants might not notice this, but still the marked elements might be used by bottom-up automatic processes to facilitate categorization through faster formation of perceptual detectors based on the marked characteristic parts.

Method

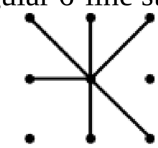
In this experiment more complex stimuli were used and the instruction did not explicitly explain the principle for categorization – it only said that it was a categorization task and that the categorization principles will not change during the experiment. Three categories were used instead of two, the third category being constructed of stimuli that belonged to neither category 1 nor category 2. Again the same two transfer locations as in experiments 1 and 2 were compared to the training location (control condition). We measured performance as the percent of correct responses and the response time for giving an answer (within-subject dependent variables).

The order of presentation of the control and two training conditions was always the same – after the training phase, transfer position 2 was tested first, then transfer at 4.5 degrees of visual angle was tested, and finally the control (training) position was tested. This was the same order as in Pilot Experiment 1. It was reasoned, that if "spotlight" type selective attention was involved or if there was forgetting after the training phase, this would be the most difficult scenario for finding difference between the control and the transfer locations. Since the control was always last, any declines in performance, would affect it to the largest extent. At the same time, there was always a switch of position before the control (in Experiment 1 there wasn't switch of position in 2 of the 6 possible orders, since the control position is the same as the training position). Position 4.5 was always in the middle of the presentation order.

The locations of the three positions (training and two transfer positions) were the same as in the previous two experiments.

The stimuli were constructed of 6 lines instead of 5 (as in experiments 1 and 2) and in half of the cases the characteristic element was slightly emphasized (Figure 13). This was done in order to vary the degree of "implicitness" of the task, assuming that the emphasized stimulus might be easier to learn and more explicit, while the regular stimulus – more difficult to learn. In the emphasized condition, emphasized stimuli for category 1 and 2 appeared only in some of the trials (half of them).

A regular 6-line stimulus



The same stimulus, with the categorical element emphasized

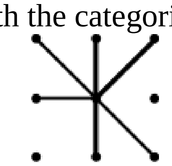


Figure 13. Regular and "marked" stimuli used in Experiment 3

Participants

Participants in the study were students at New Bulgarian University, who were recruited with print ads spread around the university and were paid for participation in the study. The payment was higher, if they needed more time to learn so that participants would be motivated to finish the difficult experiment.

Unfortunately, as in Pilot Experiment 1, many participants refused to try completing the experiment and dropped out after several sessions because the task was too difficult. Out of 45 participants who started the experiment, only 17 completed it successfully.

Apparatus

The apparatus was the same as in the previous experiments

Stimuli

The stimuli were entirely based on Pevtzow and Goldstone's (1994) stimuli – the only difference was the presence of "marked" stimuli for half of the participants. Stimuli were composed of black line-segments, connecting dots on a 3 by 3 grid. The line segments were 2 pixels wide (approximately 0.05 degrees of visual angle) and the points were with diameter 3 pixels. The "marked" stimuli (with the categorical element emphasized) had thicker categorical line segments – 4 pixels wide. The overall stimulus was 80 pixels high and 80 pixels wide (corresponding to approximately 2 degrees of visual angle). There were four basic stimuli which were distorted by adding a random line to the stem. The random lines were added to all possible places except if the added line resulted in any closed parts or detached lines. Categories were defined by three overlapping lines shared by category members as in Pevtzow and Goldstone's original experiment. Participants were given at random one of two different categorization rules in order to avoid possible effects that were specific for the selected characteristic element. Participants in the experiment always saw one object at a time, composed of 6 lines (a 5-line basic object plus a randomly added line). For every participant there were three categories and the same categorization rule remained throughout the whole experiment. The rule was: category 1 consisted of stimuli with characteristic element 1, category 2 consisted of stimuli with characteristic element 2 and category 3 consisted of all the rest of the stimuli (without element 1 or 2). The stimuli were presented on a white screen.

Procedure

The procedure was the same as in the previous experiments

Results

The patterns of results for the regular and for the emphasized stimuli where almost identical – for the regular stimuli group the average accuracy was 90,0%, 85,0% and 77,5% respectively for the control, transfer 2 and transfer 4.5 locations, while for the emphasized stimuli group, the average accuracy for the same respective locations was 85,7%, 85,7% and 78,6%. There was no interaction between the within subject factor Location and the between-subject condition – Implicitness of the stimuli (p>0.1) and the two groups were collapsed in the final analysis. There wasn't a main effect of Implicitness of the stimuli either (p>0.1)

The results from all 17 participants showed the same pattern as in Experiment 1. The overall accuracy was lower, probably because the task was harder; however, there was still full transfer of learning to the closer position and significantly lower performance in the farther location (4.5 degrees of visual angle).

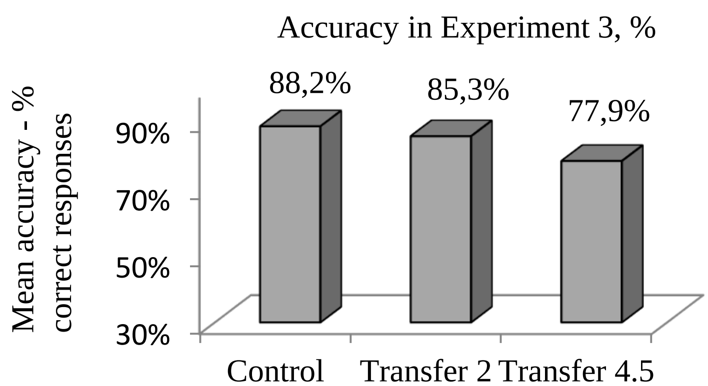


Figure 14. Percent correct answers for the control and two transfer positions in Experiment 3

Repeated Measures ANOVA showed that the overall effect was significant: F (2, 32) = 3,317; $p<0.05$; $\eta_p^2 = 0.172$;

A Pairwise comparison of the means (with Bonferroni adjustment) showed that there wasn't a significant difference between the control position and the 2 degrees transfer position ($p > 0.1$) and between the 2 degrees transfer and the 4.5 degrees transfer positions ($p > 0.1$). There was marginal difference only between the control and 4.5 degrees transfer positions ($p < 0.1$).

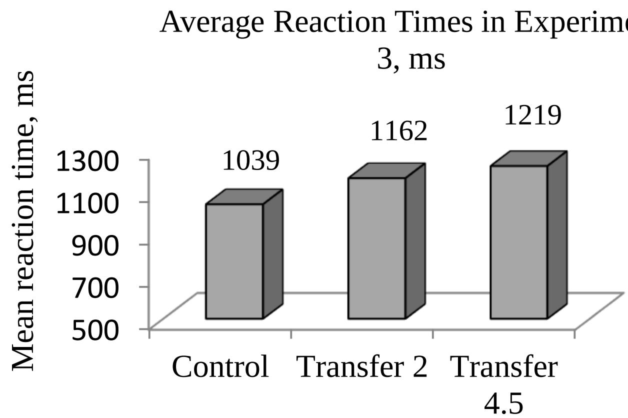


Figure 15. Reaction times (in ms) for the control and two transfer positions – Experiment 3

Difference between the conditions was also found in the reaction times: F (2, 32) = 5,593; $p < 0.01$; $\eta_p^2 = 0.259$. A Pairwise comparison (with Bonferroni adjustment) between the means shows that only the difference between the control possition and the 4.5-degrees condition was significant ($p<0.05$).

Discussion of Experiment 3

Experiment 3 confirmed once again the pattern of the finding from Experiment 1 – there is partial position transfer in categorical perceptual learning for a new position at distance 4.5 degree visual angle from the training location and full position transfer of the learned for the 2 degree position.

The experiment showed that this task is quite difficult for the participants – only 1/3 of the people who started the experiment actually completed it. Difficulties came from having three categories instead of two and from not explaining the participants the exact categorization principles in the instruction. Implicitness was further explored by including stimuli with the characteristic element slightly thicker for half of the participants. There were no trends whatsoever for changes in the pattern of results due to the varied implicitness of the task.

The main focus of the experiments was to explore whether categorical learning based on characteristic elements was position specific or translation invariant for these particular types of categorical stimuli (constructed from line segments) – the experiments showed that there is almost full position transfer, with only the further location showing partial, significantly lower in terms of accuracy and slower in terms of reaction time transfer of learning (but still very high, compared to lack of transfer in similar tasks).

Chapter IX. General Discussion

The research presented here focuses on understanding the role of perceptual learning in categorical learning of simple stimuli which has been argued to involve the formation of perceptual detectors (Goldstone, 2000; Pevtzow & Goldstone, 1994). These phenomena are approached here on the one hand through a new computational model, tested with stimuli and tasks from real experiments with infants and adults, and on the other hand the locus of learning is studied through the well-known in the perceptual learning field paradigm of position transfer (Dill, 2002). Both approaches are aimed at understanding whether this type of categorical learning involves perceptual learning elements; what are the possible mechanisms for perceptual learning to take place in visual categorical learning and finally, what is the actual level of these mechanisms in visual processing.

This dissertation claims (and provides evidence) that perceptual learning exists in higher-level tasks as well as in the typically studied lower-level tasks and that deliberate focus on some of these higher-level tasks (like for example visual categorical learning based on characteristic elements) could help understand perceptual learning better, not to mention establishing more clearly its role in human cognition.

The presented simulations have shown that it is computationally possible to account for both supervised and unsupervised perceptual learning without using built-in primitive features at the level that is eventually diagnostic for categorization. This was achieved by a fairly simple structure and by plausible mechanisms. The suggested model for perceptual learning is a first step toward a more global approach to learning that intends to bring together concepts and perception.

The experiments on translation invariance of learning have shown full position transfer of perceptual learning for a new position at 2 degrees of visual angle distance from the control position and almost full transfer at the second position, at 4.5 degrees of visual angle distance from the training location. This pattern of results is found in several experiments which involve categorical learning of simple objects with varying degree of implicitness. The effects are not found, however, in a simpler part-whole judgment task which arguably doesn't involve perceptual learning (or involves less learning than the categorization tasks). These empirical findings are interpreted as evidence for perceptual learning in common, everyday tasks as visual categorization. These results challenge existing models and theories in the field to account for partial position specificity in visual categorical learning.

Chapter X. Translation invariance in the model and future work

There are certain limitations in every model and translation invariance is particulary difficult to achieve for a Hebbian-based model like the presented in this dissertation. The goal of the translation invariance experiments was to explore the level of the changes which take place during categorical learning and to explore whether such learning exhibits similar properties to the classical perceptual learning, observed in low-level tasks like vernier or line orientation.

In order for translation invariance to occur in the current implementation of the model, the retinotopical input should be available for the neighbouring microdetectors as well, not only for the microdetectors whose receptive fields cover the training location.

Introducing the same input to different competitive units could be implemented in a variety of ways, including lateral Hebian connections between the input nodes which are partly of fully learned throughout the individual visual experience. Another way would be to have recurrent connections or overlapping receptive fields for the microdetectors which vary in size (as there are such varying receptive fields in the primary visual cortext). This could eventually lead to translation invariant learning for the output layer of the model.

In terms of locus of learning, the model explains the incomplete translation invariance effect as changes of weights between low-level and mid-level representational structures. This doesn't mean that the lowest level (micro-detectors) have stopped learning – in the current task, where well familiar stimuli are used, the micro-detectors don't need to learn, because the low-level representations needed for processing of these stimuli are already developed and suited for the task. In tasks like vernier discrimination or line orientation, however, even the micro-detectors need to learn, which is a local process, strongly dependent ot the input. This could explain why perceptual learning in vernier discrimination or in other low-level tasks for example is position specific.

Chapter XI. Contributions

Theoretical contributions

1. This dissertation demonstrates how higher cognitive tasks can (and should) be explored in the same way as low-level "sensory" tasks in the domain of perceptual learning in order to study and understand the level and nature of the perceptual components of such tasks.
2. A new model of categorical perceptual learning claims that perceptual learning is achieved by both reweighting and changes in representations (low-level as well as mid-level). The interplay between these processes is demonstrated computationally in a clear, falsifiable model.
3. Several experiments demonstrate that categorical learning of simple visual objects tends to show the same basic effects (position specificity of learning) found in typical perceptual learning tasks. This provides strong evidence that perceptual learning is not a phenomenon found only in "sensory" or "implicit" tasks but an important part of all tasks which involve sensory processing.

Empirical and methodological contributions

4. A novel Hebbian-based neural network mechanism is developed and explored. While different parts of the model rely on well-known mechanisms such as Hebbian and Competitive learning, the particular structure of the new model is novel and very promising.
5. Two types of categorical learning are replicated through simulations and studied in detail in terms of processes with the new model of categorical perceptual learning.
6. The position specificity paradigm, used often in studies of perceptual learning, is applied for the first time to categorical learning of simple objects and the results are reported in great detail. Tasks like categorical learning of simple objects open many new possibilities of perceptual learning research, for example varying the degree of implicitness of the task.
7. Partial position transfer of learning is demonstrated for a "higher" task like visual categorical learning of objects.

A list of author's publications, related to this dissertation

Gerganov, A., Grinberg, M., & Goldstone, R. L. (2009). Partial Position Transfer in Categorical Perceptual Learning Translation invariance – theoretical rationale. In *Procceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (Vol. 1, pp. 1828–1833).

Gerganov, A., Grinberg, M., Quinn, P. C., & Goldstone, R. L. (2007). Simulating Conceptually-Guided Perceptual Learning. In *Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society* (pp. 287–292). Austin, TX: Cognitive Science Society.

Cited by 4 (6 citations by coauthors are excluded from the list below):

Chartier, S., & Giguère, G. (2008). Autonomous perceptual feature extraction in a topology-constrained architecture. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1868–1873).

Giguère, G. (2009). Une structure associative bidirectionnelle d'auto-encodage permettant l'apprentissage et la catégorisation perceptuels.

Stöckle-Schobel, R. (2012). Perceptual learning and feature-based approaches to concepts--a critical discussion. *Frontiers in Psychology, 3*.

Stoeckle-Schobel, R. V. J. (2014). Concept learning challenged.

Goldstone, R. L., Gerganov, A., Landy, D., & Roberts, M. E. (2009). Learning to See and Conceive. In L. Tommasi, M. A. Peterson, & L. Nadel (Eds.), *Cognitive biology: Evolutionary and developmental perspectives on mind, brain, and behavior. Vienna series in theoretical biology* (pp. 163–188). Cambridge, MA.: The MIT Press.

Cited by 9 (3 citations by coauthors are excluded from the list below):

Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, *16*(1), 2–38.

Blunden, A. G., Wang, T., Griffiths, D. W., & Little, D. R. (2014). Logical-rules and the classification of integral dimensions: individual differences in the processing of arbitrary dimensions. *Frontiers in Psychology, 5*.

Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., Özgen, E., & Davies, I. R. L. (2012). Neural correlates of acquired color category effects. *Brain and Cognition*, *80*(1), 126–143.

Halberstadt, J., Winkielman, P., Niedenthal, P. M., & Dalle, N. (2009). Emotional conception how embodied emotion concepts guide perception and facial action. *Psychological Science*, *20*(10), 1254–1261.

Jennings, B. J., & Martinovic, J. (2014). Luminance and color inputs to mid-level and high-level vision. *Journal of Vision*, *14*(2), 9.

Knittel, A. (2012). Learning feature hierarchies under reinforcement. In *Evolutionary Computation (CEC), 2012 IEEE Congress on* (pp. 1–8).

Meiser, T. (2013). Experimental psychology: A case for methodological rigor and theoretical relevance.

Mroczko-W\kasowicz, A., & Nikolić, D. (2014). Semantic mechanisms may be responsible for developing synesthesia. *Frontiers in Human Neuroscience*, *8*.

Weiskopf, D. A. (2013). Observational concepts. *Concepts: New Directions, MIT Press, Cambridge MA*.