NELE PÕLDVER

An experimental exploration of automatic processes in visual perception
DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

52
NELE PÕLDVER

An experimental exploration of automatic processes in visual perception
CONTENTS

LIST OF ORIGINAL PUBLICATIONS ....................................................... 6

INTRODUCTION .......................................................................................... 7
  Overview of the visual system ............................................................... 7
  Automaticity in visual perception ......................................................... 9

DETECTION OF MOTION .......................................................................... 14
  Relative motion principle revisited ...................................................... 15
  Study I: Reaction time to motion onset and magnitude
  estimation of velocity in the presence of background motion ........ 16
  Detecting motion direction changes without the need
  for focal attention .............................................................................. 17
  Study II: Unattended and attended visual change detection
  of motion as indexed by event-related potentials and
  its behavioral correlates ................................................................ 20
  Processing motion information relative to other tasks ..................... 22
  Study III: Visual evoked potentials to change in coloration
  of a moving bar ............................................................................. 23

PROCESSING EMOTIONAL EXPRESSIONS ........................................... 26
  Study IV: vMMN for schematic faces: automatic detection
  of change in emotional expression.................................................. 29

VISUAL MISMATCH NEGATIVITY- AN AUTOMATIC CHANGE-
DETECTION MECHANISM ........................................................................ 32
  Study V: Visual Mismatch Negativity (vMMN): A review and
  meta-analysis of studies in psychiatric and neurological
  disorders .......................................................................................... 35

CONCLUSIONS ............................................................................................ 39

ACKNOWLEDGEMENTS ........................................................................ 41

REFERENCES ............................................................................................... 43

EESTIKEELNE KOKKUVÕTE (Summary in Estonian) ............................. 55

PUBLICATIONS ........................................................................................... 61

CURRICULUM VITAE ................................................................................ 148

ELULOOKIRJELDUS (CV in Estonian) ...................................................... 150
LIST OF ORIGINAL PUBLICATIONS

This dissertation is based on the following original publications, further referred to as Studies by respective Roman numerals:


The author of the present dissertation contributed to these publications as follows:

In **Studies I and II**, the author participated in formulating the research questions and preparing the experimental paradigms, collected and analysed the data, interpreted the results, and wrote the manuscripts as the first author.

In **Study III**, the author collected the data and participated in interpreting the results and writing the manuscript.

In **Study IV**, the author participated in formulating the research questions, preparing the experimental paradigm, collecting and interpreting the data, and writing the manuscript.

In **Study V**, the author participated in initiating the idea of the manuscript, co-wrote two chapters of the review and participated in correcting and reviewing the manuscript.
INTRODUCTION

The main function of the visual system is to represent various dimensions of the surrounding environment that are relevant for the organism orienting in this environment. As the needs and behaviour of organisms evolved and became more sophisticated, new dimensions were added to the increasingly complex representation of the environment in the visual system. As an example, although some ancestors of vertebrates had colour vision – different opsin molecules with different absorbance peaks, meaning they had the ability to differentiate between colours – this capacity became widespread only among mammals and primates (Gehring, 2014). Thus, some visual dimensions appeared earlier and some others on later stages of evolution. It is intuitively obvious, for instance, that it was more urgent to recognize the movement of something without identifying its exact shape or colour. The architecture and functions of the visual system support the idea that some perceptual processes are executed more automatically than others are.

The research presented and discussed in this thesis is an experimental exploration of processes in visual perception, which all display a considerable amount of automaticity. These processes are targeted from different angles using different experimental paradigms and stimuli, and by measuring both behavioural and brain responses. I will first give a short overview of the structure and functions of the visual system and introduce the concept of automaticity. Then I will continue with the topic of motion detection, which is regarded one of the most basic processes shaped by evolution. In the first two empirical studies, I will investigate the effects of the surrounding environment on the detection of motion, as well as show that our visual system is built to automatically process motion information that is outside of our attentional focus. After that, I will proceed with the study on how sudden changes in coloration of a moving object are detected, and address the question of what happens when multiple stimulus qualities are present and varied, especially when both of those are rather salient. Next, I present evidence indicating that we probably have a module for detecting facial expressions quickly and automatically. The roof of this module is based on the visual mismatch negativity (vMMN) and presented evidence allows to assert that vMMN exists and can be used for studying automatic visual processes. Finally, I discuss the cases when the brains’ capability of processing the changes in visual input automatically is altered or intact, and show the practical value of vMMN.

Overview of the visual system

It was noticed relatively early that different visual attributes are analysed by different parts of the brain. An astonishing progress in our understanding of the architecture of the visual system (starting with studying the receptive fields of
cortical cells in the middle of the 20th century) identified several streams of information, which may be responsible for different perceptual capacities (Hubel & Wiesel, 2005). Anatomical and physiological studies indicated that there are separate systems for the analysis of colour, stereopsis, movement, and orientation in the visual cortex (Livingstone & Hubel, 1987, 1988). This specialization begins already on lower levels of the visual system, starting from the input layer of the eye’s retina with different types of photoreceptors (rods and cones, the latter divided into three categories according to their wavelength selectivity). The retinal output layer includes anatomically different parasol (M), midget (P) and bistratified (K) ganglion cells, which project to different cellular layers in the lateral geniculate nucleus (LGN). The M cells project to the magnocellular layers (layers 1–2), the P cells to the parvocellular layers (3–6), the K cells to the interlaminar layers of the LGN1 (Kaplan, 2014). The magnocellular pathway is selective for motion and depth information. It is believed to have developed earlier, and known to operate faster than the parvocellular pathway. The parvocellular pathway is selective to colour (especially red-green comparison) and form information. The koniocellular pathway is thought to take part in colour processing in mostly the blue-yellow comparison dimension. At lower levels, in the retina and in the LGN, cells in these subdivisions differ in their colour selectivity, contrast sensitivity, temporal perception, and acuity (Livingstone & Hubel, 1988; Kaplan, 2014). The projections of the magno- and parvocellular (but also koniocellular) pathways (from the LGN to primary visual cortex, V1) function as the initial inputs to different information processing streams in the visual cortex. Within those streams, there are feedforward, horizontal, and feedback connections (see also p. 33 for a note on terminology) (Lamme & Roelfsema, 2000).

Clinical observations as well as animal studies have suggested that it might be relevant to distinguish between two large cortical information processing systems providing “vision for action” (the dorsal stream) and “vision for perception” (the ventral stream) (Goodale & Milner, 1992; Milner & Goodale, 2008; Goodale, 2011; Whitwell, Milner, & Goodale, 2014)2. The dorsal stream receives mostly magnocellular input – cells from the magnocellular layers of the LGN terminate in V1 layers 4Ca and 4B, and project to the posterior parietal cortex.

---

1 While magno- and parvocellular pathways are excessively studied, less is known about the third, koniocellular, pathway, which retinal source supposedly consists of small bistratified ganglion cells (color sensitive blue-ON/yellow-OFF cells) projecting to the interlaminar layers (or the K layers, sometimes also referred to as the S layers) of the LGN (Kaplan, 2014).

2 The anatomical dorsal and ventral streams have been classically termed as “Where” (processing spatial information) and “What” (processing object information) streams, respectively (Ungerleider & Mishkin, 1982). Goodale and Milner added the goal-relevancy aspect to the dichotomy, redefining the dorsal stream as a “How” system responsible for ecocentrically guided action (e.g. grasping something), while the “What” system (ventral stream) is responsible for object representations.
through visual cortex areas V2, V3 and V5 (also known as MT, middle temporal area). The ventral stream receives input from predominantly parvocellular (LGN projections to layer 4Cβ in V1), as well as magno- and koniocellular (LGN projections terminating in the upper layers of V1) pathways, and projects to inferior temporal cortex through V2 and V4 areas (Livingstone & Hubel, 1987, 1988; Goodale, 2011; Kaplan, 2014).

Dorsal and ventral streams have been found to support different computations involved in vision. For example, it was observed that if a brain ablation made an organism unable to identify shapes and patterns visually, these organisms nevertheless showed a considerable ability to localize objects in space by means of vision (Schneider, 1969). It was also noticed that some human patients with lesions to their primary visual cortex demonstrated residual visual capacity—termed blindsight—, but without acknowledged perceptual awareness (Sanders, Warrington, Marshall, & Weiskrantz, 1974; Kentridge, Heywood, & Weiskrantz, 1999; Stroop & Cowey, 1997; Weiskrantz, 1990). This indicates that some visual processing occurs outside of or can bypass primary visual cortex (see also recent proof on that: Mundinano, Chen, de Souza, Sarossy, Joanisse, Goodale, & Bourne, 2017), and it can function autonomously without the necessary involvement of consciousness or attentional control. One obvious generalization from these observations is that awareness and control is not a requirement for all visual mechanisms. There may be low-level mechanisms that require no attention and control for their execution.

**Automaticity in visual perception**

The observations about the architecture and functions of the visual system indicate that the visual representation is created by multiple visual mechanisms, which differ considerably from one another by their principal characteristics. One of these defining characteristics is the level of automaticity of the process. Automaticity refers to the ability to do things without occupying one’s mind or keeping attention with operations required for the execution of a process. Although automaticity may develop in the result of practice, repetitions, and training, there are processes that can be executed automatically without any deliberate learning or practice.

In a series of papers, Richard Shiffrin and Walter Schneider advanced a two-process theory of human information processing (Schneider & Shiffrin, 1977; Mundinano, 2017).

---

3 It has been shown that there are different routes bypassing the V1 in the visual system. In the main one, related to action (grasping, reaching) and visual attention, signal is transformed from the retina to superior colliculus in the midbrain, where information is sent to parietal cortex via the pulvinar (in the thalamus) and MT. There are also direct projections from the retina to the pulvinar, as well as direct projections from the LGN to MT (Goodale, 2011; Shipp, 2004). The retina-pulvinar-MT route has recently been shown by Mundinano and colleagues (2017) in a case study of a patient with bilateral V1 damage at 9 days of age.
Shiffrin & Schneider, 1977). First, automatic processing refers to activation of a learned sequence of elements in long-term memory that is initiated by appropriate inputs and then proceeds automatically. Automatic processing occurs quickly, without subject’s control, with little or no capacity limitations of the system, without requiring focal attention, is unaffected by task load, and once (perceptually) learned, is difficult to alter or stop. Second, controlled processing is a temporary activation of a sequence of elements that can also be set up easily but demands attention, is usually serial in nature, therefore capacity-limited, and is controlled by the subject (Birnboim, 2003; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977, 1984). The distinction between automatic and controlled processing, although supported by many, has not been convincing to everyone, and is debated over outside the field of perception psychology as well (e.g. Moors, 2016).

Many researchers have proposed largely overlapping lists of characteristics of automatic processes. Bargh (1994) names the characteristics that usually accompany automatic mental processes or behaviour the four horsemen of automaticity – (un)awareness, (un)intentionality, efficiency and (un)controllability.

1. **(Un)awareness**: A person may be unaware of the mental processes that are occurring. One of the best indicators of awareness is selective attention defined as a tendency of visual processing to be confined largely to stimuli that are relevant to behaviour (Knudsen, 2007; Moore & Zirnsak, 2017). Equally informative are the situations where selective attention is not operational. A good example is a spontaneous brain response to change detection, which was discovered by Risto Näätänen and termed the mismatch negativity or MMN (Näätänen, Gaillard, & Mäntysalo, 1978). MMN, first studied in auditory modality, is traditionally elicited by a discriminable physical change (called a deviant stimulus) in a train of frequently presented homogeneous stimuli (called standard stimuli), and computable from the brain’s electrophysiological activity as a difference between the ERP (event related potential) responses to deviants and standards. (See the more thorough overview of the MMN on pp. 18–19) One of the most central characteristics of the MMN is its independence from awareness. MMN is best elicited when the subject is not attending to the MMN-generating stimuli, but is engaged in an auxiliary task (in auditory MMN experiments often reading a book or watching a silent movie) that consumes most of the attention (Duncan, Barry, Connolly, Fischer, Michie, Näätänen et al., 2009; Näätänen, Paavilainen, Rinne, & Alho, 2007). There are convincing results that MMN is elicited in the brains of very young children who have not yet developed skills to control and regulate their psychological functions (e.g. Cheour, Alho, Čeponienė, Reinikainen, Sainio, Pohjavuori, et al., 1998; Cheour, Martynova, Näätänen, Erkkola, Sillanpää, Kero, et al., 2002; for a review see Cheour, Leppänen, & Kraus, 2000). Independence of awareness is also evident in research showing that the magnitude of the MMN is practically the same whether an individual is awake or sleeping, as
have been shown on adults (Nashida, Yabe, Sato, Hiruma, Sutoh, Shinozaki, & Kaneko, 2000; Sculthorpe, Ouellet, & Campbell, 2009) as well as infants (Cheour-Luhtanen, Alho, Kujala, Sainio, Reinikainen, Renlund, et al., 1995; Stefanics, Háden, Huotilainen, Balázs, Sziller, Beke, et al., 2007). More strikingly, MMN to change detection has been recorded in foetuses pre-birth (Huotilainen, Kujala, Hotakainen, Parkkonen, Taulu, Simola, et al., 2005). MMN is also elicited in clinical patients with reduced or lacking awareness as a predictor of recovery – comatose patients and patients in persistent vegetative state (for a review see Näätänen, Kujala, Kreegipuu, Carlson, Escera, Baldeweg, & Ponton, 2011). These examples of intact change detection processes illustrate that one of the most central characteristic of automaticity is the lack of or reduced awareness.

2. **(Un)intentionality**: A person may not be involved with the initiation of a mental process. A typical means to manipulation is the instruction given in experiments. For example, when the instruction is to perform a certain task (e.g. detecting a stimulus), the researchers often assume the participants ignore some other information simultaneously presented to them. If the participants are unable to ignore it, this may indicate that the process of registering information is automatic and executed even without intention. Vision researchers are familiar with situations in which participants are not able to carry out relatively simple tasks although they were instructed to do so. For instance, this applies in case of various visual illusions. As an illustrative example, Morgan, Hole and Glennerster (1990) presented a modification of the classical Müller-Lyer illusion (the line appearing to be longer if the end points have outgoing arrowheads attached to them, and shorter if the arrowheads are ingoing). Instead of presenting the lines, only the end points of the lines were marked as dots, each surrounded by a cluster of dots in other colour. In the Müller-Lyer illusion spirit, although the target dots were in exact same positions in both cases, in one case the clusters were tilted towards the outer frame of the screen; in the second case, the clusters were tilted towards the centre of the screen. The subjects had to report the distance between the black dots. A target dot seemed to be displaced towards the centre of a surrounding cluster of dots, even though it was clearly discriminable from other members of the cluster by their colour, resulting in subjects estimating the distance between the dots to be shorter, when the clusters were closer to the centre of the screen. The authors argue that this is a demonstration of constraints upon visual processing, which arise when subjects are instructed to carry out a task to which the visual system is not adapted. (Morgan et al., 1990). This was obviously not an intention of the subjects to report the position of the centre of a cloud of dots.

3. **Efficiency or cognitive load**: Automatic mental processes tend to have low cognitive load, not requiring very demanding mental resources. It is well documented that encoding and representing information in short-term memory is capacity limited: only a certain amount of information can be stored (Miller,
Automatic processes are supposed to have no such limitations and they can easily go beyond limited capacity. A good example is the visual pop-out phenomenon, which refers to the situation in which a unique visual target can be detected rapidly and effortlessly among a set of homogeneous distractors (Treisman & Paterson, 1984; Wolfe, 1994). One criterion demonstrating efficiency is the time needed for the execution of a task. Unmistakably, mental processes, which require shorter time for their execution, are better candidates for automatic processes than those whose implementation is more time consuming (Kahneman, 2011). There is considerable amount of evidence suggesting that early visual processing is executed with the help of relatively simple and autonomous mechanisms (Julesz & Schumer, 1981), even if those perceptual phenomena are divided into different processing stages that follow each other. For example, many motion perception phenomena such as induced or rotary motion (Proffitt, Cutting, & Stier, 1979; Wallach & O’Leary, 1985) to say nothing of biological motion (Cutting, Proffitt, & Kozlowski, 1978; Johansson, 1973; Blake & Shiffrar, 2007) seem to require several stages of computation. Gunnar Johansson proposed that first, a vector of the common motion is extracted, and then all particular motions are computed relative to that (Johansson, 1973, 1978). This means that the computation of one local motion vector is put into the context of other motion vectors, which is necessary for the determination of absolute motion. Although there seems to be at least two simultaneous processes – one for common motion and one for relative motion – the minimization of relative motion dominates perception for most stimuli (Cutting & Proffitt, 1982). Even though motion perception goes through several stages of processing, requires several computations and comparison of the results of these computations, it is believed to be largely automatic not requiring excessive cognitive load (this is what already Johansson, 1973, stated for biological motion).

4. (Un)controllability: A person may not have the ability to stop or alter a mental process or behaviour after initiation. Many perceptual functions are organized based on the goal-directed or top-down principles. There is evidence that intraparietal cortex and superior frontal cortex (forming a dorsal frontoparietal network) are involved in visual processing modulating selection and detection of visual stimuli (Carrasco, 2011; Corbetta & Shulman, 2002). However, some (visual) functions are rather autonomous, stimulus-driven and cannot be stopped or altered. For instance, it was proposed that human and nonhuman animal minds contain an autonomous subcortical module, which activates automatically to stimuli that are fear relevant and which are relatively impenetrable to cognitive control (Öhman & Mineka, 2001). Activation of the “fear module” cannot usually be stopped, is relatively unaffected by other cognitive processes, meaning also it does not need conscious representation of the fear-inducing stimuli to be initiated.
It is reasonable to represent visual processing as a hierarchy of mechanisms leading to increasingly complex representations (Hochstein & Ahhisar, 2002). At the bottom of this hierarchical system are largely automatic mechanisms, which are usually cognitively impenetrable (Moors & De Houwer, 2006). The conscious perception begins at a certain stage of the hierarchy and becomes dominating at the top. Thus, awareness and controllability become more likely on higher stages of the hierarchy of visual mechanisms. As recent studies demonstrate, elementary automatic processes may not be entirely different from higher cognitive abilities. They often demonstrate a primitive (sensory) intelligence (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001; Näätänen, Astikainen, Ruusuvirta, & Huotilainen, 2010).
DETECTION OF MOTION

Every day walking to the University during rush hour, I have to cross a busy street with four lanes that does not have a traffic light. When near it, I direct my attention to vehicles approaching from my left on the first lane to estimate their velocity and whether they stop, and make a decision to cross or not to cross the lane. At the same time, there are vehicles on lane two and a bit further, on the last two lanes, coming from the right. As I am still busy crossing the first lane and noticing the movement on the second one, I have not yet directed my attention to the last two lanes, that are further away, but despite that, the movement of the vehicles there has been registered by my visual system. When we think about the functions motion detection has in the environment – from avoiding to become pray in the jungle for our ancestors to orienting in the city jungle nowadays –, it is only logical that this process needs to be rather automatic in its nature. This is supported by the knowledge of the temporal qualities of the neurological mechanisms that are tuned to the detection of motion, described even as a “magnocellular advantage” (Laycock, Crewther, & Crewther, 2007). For example, the conduction velocity of axons in the magnocellular pathway is high compared to the parvocellular pathway (Livingstone & Hubel, 1988; Kaplan, 2014). Response latencies of magno- and parvocellular neurons have been shown to differ from around 10 ms (measured from the LGN, Maunsell & Gibson, 1992) up to 20 ms (measured from V1 layers, Nowak, Munk, Girard, & Bullier, 1995; see also Lamme & Roelfsema, 2000, for a meta-analysis on the response latencies in visual cortex). However, inferring neural timing of events from only response latencies should be addressed with caution. Maunsell and colleagues cleverly showed, by using single-unit recordings from both parvo- and magnocellular LGN neurons, that although the magnocellular response latencies were around 10 ms shorter, this advantage was eliminated when they took into account the number of cortical parvo- and magnocellular recipient neurons (possibly reflecting the actual input as a result of signal summation) (Maunsell, Ghose, Assad, McAdams, Boudreau, & Noerager, 1999).

In the functional division of the visual cortex, middle temporal area (MT/V5) has been shown to be mostly responsible for motion processing (direction and movement selective neurons were shown in the corresponding area in rhesus monkeys by Zeki, 1974; direction selective neuronal populations in human prestriate cortex were shown by Moutoussis & Zeki, 2008). Already the inputs from V1 to MT are strongly direction and speed selective, also tuned to binocular disparity (for a review see Born & Bradley, 2005). It has been argued that one of the functions of the MT area is to make sense of the surrounding environment via segmenting it into centre and surround, which is crucial for differentiating figures from grounds, perceiving objects moving, or registering the 3D shapes from motion (Tadin, Lappin, Gilroy, & Blake, 2003). Bradley and Andersen (1998) used single-cell recordings and showed that the firing rates of MT centre-surround neurons increase when the disparity in the surround
differs from the disparity in the centre. Thus, MT neurons seem to serve the purpose of image segmentation through assessing the salience of at least two stimulus qualities – direction and depth (Bradley & Andersen, 1998).

Many researchers have claimed that our survival depends critically on being able to perceive the movement of objects accurately and quickly without registering other properties (Palmer, 1999, p. 466). At the same time, our visual field is almost constantly filled with different features – edges, colours, objects, moving and stationary parts, etc. Although research in visual perception often means studying and experimentally manipulating with a single specific visual attribute of a stimulus – e.g. orientation, motion, colour, luminance, shape – this does not, however reflect the everyday real-life visual input we normally get. The perception of one visual attribute does not happen in isolation being dependent on what happens in the neighbouring areas (Gibson, 1950; Johansson, 1978). The perceptual centre-surround relationship (Nawrot & Sekuler, 1990) resembles that in the neuronal level, especially in V5/MT area (Tadin et al., 2003). Bradley and Andersen have elegantly phrased it: “How do we know, for example, when two regions of contrast on the retina correspond to a single object or to separate objects? Although colour and texture may vary over different parts of an object or surface, the direction and speed tend to be the same on all the parts. Therefore, relative motion is a useful cue for parsing an image into its separate components.” (Bradley & Andersen, 1998).

**Relative motion principle revisited**

It was already Galileo Galilei who denied absolute motion. His famous boat example from the “Dialogue concerning the two chief world systems” in 1632 proposed that to speak about an object we necessarily need to specify a frame of reference relative to which we determine all of the object’s displacements (Galilei, 1967, a translation by Stillman Drake). When Karl Duncker presented his description of the induced movement phenomenon in early 20th century (Duncker, 1929, referred in Holmgren, 1973; Cutting & Proffitt, 1982; Wallach, O’Leary, & McMahon, 1982; Reinhardt-Rutland, 1988), it was almost agreed upon that motion cannot be perceived with no relation to other objects or a specified background which serves as a frame of reference. The most common example of the induced motion phenomenon is the case where physically stationary object is perceived to move in the direction opposite to another object, which is moving near the former or surrounding it. This type of a condition is termed also “heterokinesis” (Nawrot & Sekuler, 1990) or “simultaneous motion contrast” (Tynan & Sekuler, 1975). An opposite result, where a stationary object is perceived to move in the same direction as an inducing surround or object, is a phenomenon of motion capture or assimilation (Rama-chandran, 1987), also referred to as “homokinesis” (Nawrot & Sekuler, 1990). Psychophysical and electrophysiological studies have shown that surround
moving in the same direction as the target decreases neuronal activity and behavioural sensitivity (Ido, Ohtani, & Ejima, 2000). It became a standard theory that motion is always perceived relative to some other objects (Gibson, 1950; Holmgren, 1973; Johansson, 1978; Wallach et al., 1982). The relativity principle of motion perception was later supported by accumulating research. It was noticed that beside the object-relative frame of reference the observer can use her or his body as a reference for perceived motion. Thus, cues for motion perception can be discriminated into subject-related cues that provide information regarding the motion of the object relative to the observer, e.g. her/his retina, and relative cues (sometimes referred to as object-relative cues e.g. in Wallach et al., 1982) regarding the motion of objects relative to each other (Reinhardt-Rutland, 1988). See the Introduction of Study I for a more detailed overview on relative motion.

**Study I: Reaction time to motion onset and magnitude estimation of velocity in the presence of background motion**

Study I was designed to address the question of how the presence of motion is detected in the company of background motion. Abovementioned previous studies seem to suggest that it is difficult or even impossible to ignore the background, which serves as an automatic frame of reference relative to which the target motion is calculated.

The stimulus was a sine-wave grating (see Study I Methods section for specific parameters) extended to the whole computer screen. A round area in the centre of the screen (and in the centre of the visual field) served as a target and the subjects’ task was to react to the motion onset of this area (Experiment 1, reaction time (RT) being a dependent variable) or estimate on a scale from 1 (slowest) to 10 (fastest) how fast the target is moving (Experiment 2, velocity estimation being a dependent variable). In Experiment 1, we manipulated with the size of the target, the size of the gap between target and the background area (i.e. no gap versus gap), target and background stimulus-onset asynchrony (SOA) and velocities. In Experiment 2, only the small target size was used, with either a gap or no gap separating it from the background, or no background at all. Two different conditions for target motion duration were used – fixed ($t_{T} = 300$ ms) and variable duration ($t_{T} = V_{T}^{2/3}$, where the slower the stimulus moved, the longer it was presented).

Surprisingly, the results of Study I show no support for the relative motion principle. We found that in both experiments, and comparing the RT results with the baseline RT results (in case of no moving background), when there was a background effect, it seemed to worsen the perception of target motion onset.

---

4 This study has been previously published, but viewed from a different angle, in another doctoral dissertation: Raidvee, 2012.
This resulted in increased reaction times, and in the apparent decrease in the perceived velocity as shown by velocity magnitude estimations. These delays were even present when the background was moving in the opposite direction compared to the target – a situation, which should, as we expected based on the relative motion literature, create motion contrast and facilitate the perception of target motion. It seems that the observers could ignore the background motion for most of the occasions as if it did not exist at all. To put it in the historical relative motion terms, it seems that observers can use subject-relative cues – most likely their retina – as a stable enough frame of reference relative to which object motion is detected. With this task in mind, it is possible at least to some extent ignore other displacements or motions even in a close spatial vicinity of the target. Thus, the visual system is surprisingly accurate in separation of absolute motion – motion relative to human body as a frame of reference – from relative motion, which can occur between objects themselves. The results of Study I suggest that there seems to be no need for reference to other nearby objects, therefore it is very likely that motion detection is more like an automatic process which can be executed without time and energy consuming comparisons.

Detecting motion direction changes without the need for focal attention

Even if we are deliberately focussed on only a small part of the visual field – the velocity of a car approaching, a hole in the pavement to be avoided, faces of people on the sidewalk to see whether we know them – the surrounding environment is also constantly monitored by the visual system in our brain. While these processes are difficult to study via only behavioural reactions, it is possible to do so with the help of brain imaging. In the current thesis, the focus is on scalp-recorded electroencephalography (EEG) measurements, which are non-invasive and relatively easy to acquire, and from which event-related potentials (ERPs) – averaged EEG responses that are time-locked to presented stimuli – are computed. The latencies of the ERP components and their sequence is a track record of the underlying neural activity of stimulus processing in milliseconds, therefore with great temporal resolution. ERP amplitudes reflect the amount of cognitive recourses used in processing (Luck, 2012). The support for our brain’s ability to automatically register the events outside the focus of our attention, that violate regularity, comes from the research regarding the ERP component MMN (Näätänen et al., 1978) and its magnetoencephalographic (MEG) equivalent MMNm (Näätänen, 2001; Tervaniemi, Sinkkonen, Virtanen, Kallio, Ilmoniemi, Salonen, & Näätänen, 2005).
MMN is an early\textsuperscript{5} negative-going component amongst cognitive ERPs that indicates change detection. Traditionally, it has been determined as a difference between the averaged ERP responses to repetitive standard and rare deviant events in an oddball-type of stimulus presentation paradigm (the probability of deviants in a train of standard stimuli usually not exceeding 20%). However, to dissociate MMN from another early and often overlapping N1 component, it has been suggested that responses to deviant stimuli from an oddball paradigm should be compared to control stimuli from an equiprobable presentation paradigm (e.g. Jacobsen & Schröger, 2003; Astikainen, Lillström, & Ruusuvirta, 2008; for a review see Kimura, Katayama, Ohira, & Schröger, 2009; for further discussion see Kimura, 2012). Also, it has been suggested to compare physically the same stimuli presented as standards in one series and as deviants in another (Stefanics, Csukly, Komlósi, Czobor, & Czigler, 2012).

The MMN emerges to changes in stimulation in different sensory modalities. The most excessive research has been done in auditory modality, where the MMN usually peaks between 150–250 ms after the change onset (for aMMN reviews see Näätänen et al., 2007; Näätänen, Kujala, Escera, Baldeweg, Kääriäinen, Carlson, & Ponton, 2012). There are solid findings also in visual domain (for recent vMMN reviews see Kimura, 2012; Stefanics, Kremláček, & Czigler, 2014), where the latency range has been found to be wider – up to 400 ms (Kimura, Schröger, & Czigler, 2011). The vMMN usually has an occipito-parietal scalp location (for an overview and discussion, see Kimura, 2012), frontal components/generators have been less reported (but see examples from Study V (Table 1), e.g. Urban, Kremláček, Masopust, & Libiger, 2008, which show that the frontal component might be related to several clinical conditions). In addition, there is growing evidence in somatosensory (sMMN; Kekoni, Hämäläinen, Saarinen, Gröhn, Reinvikainen, Lehtokoski, & Näätänen, 1997; Shinozaki, Yabe, Sutoh, Hiruma, & Kaneko, 1998; Akatsuka, Wasaka, Nakata, Inui, Hoshiyama, & Kakigi, 2005; Restuccia, Zanini, Cazzagon, Del Piero, Martucci, & Della Marca, 2009; Strömmer, Tarkka, & Astikainen, 2014; Strömmer, Pöldver, Waselius, Kirjavainen, Järveläinen, Björksten, Tarkka, & Astikainen, 2017) and some in olfactory (oMMN; Krauel, Schott, Sojka, Pause, & Ferstl, 1999; Pause & Krauel, 2000; Sabri, Radnovich, Li, & Kareken, 2005) modalities. Hence, the MMN has proven to be a general indicator of detecting regularity violations, not specific or limited to the processing in the auditory cortex.

\textsuperscript{5} Sussman and colleagues (Sussman, Chen, Sussman-Fort, & Dinces, 2014) argue that the claim “MMN is an early auditory process” should be redefined to “MMN can be elicited after attention exerts influence”, since “early” in the MMN field means that MMN emerges prior to non-modality-specific attention-related ERP components. On one hand, MMN of course peaks quantitatively early (beginning from around 150 ms) if we think about the cognitive processes and for example motor execution of behavior. On the other hand, by the time of MMN occurring, exogenous brainstem and mid-latency responses have already taken place (Luck, 2012).
It is important to note, that what is presented as a standard or deviant event, is not necessarily defined by a physical change between the stimuli in an oddball type of presentation sequence, as was thought in the early days of the vMMN research (Pazo-Alvarez, Cadaveira, & Amendedo 2003). It can also be a change in a complex stimulus category (for a review of findings in visual modality, see Czigler, 2014) as well as a change in an abstract inter-stimulus presentation regularity (Stefanics, Kimura, & Czigler, 2011). Risto Näätänen with colleagues (Näätänen et al., 2001; Näätänen et al., 2010) have referred to the MMN as “primitive intelligence”, as our brain is already on the sensory level anticipating what is happening next in a sequence of events, and forming complex models of environmental representations.

Automaticity is one of the key characteristics of the (v)MMN. Knowing that motion perception carries high level of automaticity (Cavanagh, 1992), stimuli incorporating motion information (e.g. motion direction changes) are an excellent candidate for eliciting the vMMN. Also, it has been proposed that using stimuli eliciting dominantly magnocellular pathway activation targets the underpinnings of several neurodevelopmental and neuropsychiatric conditions (autism, schizophrenia), which rely on the magnocellular-dominated cortical feedback projections (Laycock et al., 2007). Therefore, it is somewhat surprising that motion stimuli have not been used more in vMMN research. Before Study II was published in 2013, there were nine papers, all but one from two different research groups, investigating motion direction changes in visual oddball paradigms. The studies incorporating healthy participants all showed that vMMN was elicited to motion direction changes (Pazo-Alvarez, Amenedo, & Cadaveira, 2004; Pazo-Alvarez, Amenedo, Lorenzo-López, & Cadaveira, 2004; Kremláček, Kubová, Kuba, & Langrová, 2006; Amenedo, Pazo-Alvarez, & Cadaveira, 2007). The same can be concluded if we look at the control group results from the next studies that used motion-elicited vMMN in special populations. These studies also reported vMMN deficits in schizophrenia (Urban et al., 2008), longer duration metamphetamine abuse (Hosák, Kremláček, Kuba, Libiger, & Čižek, 2008), developmental dyslexia (Wang, Bi, Gao, & Wydell, 2010), aging (Lorenzo-López, Amenedo, Pazo-Alvarez, & Cadaveira, 2004), and no pathological vMMN in spinocelebellar ataxia type 2 (SCA2, Kremláček, Vališ, Masopust, Urban, Zumrová, Taláb, et al., 2011). The results of the last mentioned studies are discussed in detail in Study V. Since 2013, there are three additional papers using very different paradigms. In one of them vMMN emerged and was shown to be independent of attentional manipulation (expansion-contrasting direction change in radial motion as a deviant, Kremláček, Kuba, Kubová, Langrová, Szanyi, Vit, & Bednář, 2013); in another, vMMN to peripherally presented motion stimuli did not emerge on ERP level, although deviance detection was indicated by visual mismatch oscillatory response (vMOR, Tugin, Hernandez-Pavon, Ilmoniemi, & Nikulin, 2016). The third study, with using saccade and fMRI measurements, differentiated between early (up to 150 ms) and late (150–250 ms) deviance detection, where the degree of
mismatch between predicted and incoming events is mainly processed at the later stage (Malekshahi, Seth, Papanikolaou, Mathews, Birbaumer, Verschure, & Caria, 2016).

**Study II: Unattended and attended visual change detection of motion as indexed by event-related potentials and its behavioral correlates**

In Study II, we addressed the question of how motion information is processed under two different attentional conditions – either without direct focal attention or while attended. We presented a novel continuous whole-display experimental setup for eliciting the vMMN, where horizontal motion direction change of a grating (with the same parameters as in Study I) in the (peripheral) background area was a deviant. Similar gratings (but not extended to the whole screen) have been previously used in vMMN research as both stationary (van Rhijn, Roeber, & O’Shea, 2013) or moving (Kremláček et al., 2006; Hosák et al., 2008; Urban et al., 2008; Kremláček et al., 2011). Presenting stimuli across the whole visual field is ecologically more valid than presenting one or a few stimuli alone, usually without a background, since the former situation is something we naturally experience in everyday situations. The specific stimulus configuration of Study II was prepared based on the results of Study I, where we showed that the behavioural reactions to the central (target) area were least affected by the peripheral background if the target was large and clearly separated from the background (condition D in Study I). In this way we could create a situation where the background could be ignored by the subject, when he or she was directing attention only to the central target area, performing a demanding task of reacting to motion onset there (“Ignore” condition in Study II). To be reminded, one of the characteristics of the (v)MMN is its independency of attention. Näätänen and colleagues (Näätänen et al., 2007) emphasize that MMN is best elicited in case of directing attention away from MMN-eliciting stimuli, while it also emerges (but with attenuated amplitudes) in case of attention. Attention-related components can interfere with the emergence of the MMN component (Sussman, 2007; Stefanics et al., 2014). Interestingly, only a few studies directly compare attended or unattended conditions, or conditions with different task load (which is related to the amount of attention possibly available for the task-irrelevant stimulus detection) in vMMN paradigms. For example, Berti (2011) showed that vMMN is equally elicited under attention, as well as in the latency range of the attentional blink (the emergence of attentional blink was confirmed by behavioural detection task). Another study presented the vMMN eliciting sequences in the binocular rivalry paradigm, showing similar vMMNs under different attention conditions (van Rhijn et al., 2013). Pazo-Alvarez and colleagues (2004) and Kremláček and colleagues (2013) showed no task difficulty effect on vMMN; however, Kimura and Takeda (2013) found a vMMN latency delay with the
increase in task difficulty (but no effect on the vMMN amplitude). Therefore, as a comparison, we also presented an attention demanding condition (“Attend” condition) in Study II. The subject was asked to attend to both areas of the screen – central target and peripheral background – and evaluate whether their motion directions are the same or opposite. All subjects were presented with both experimental conditions. We recorded the subjects’ EEG (see the details of EEG measurement and data analysis in the Methods section of Study II), as well as registered their manual reactions (RT-s) to the target motion.

The EEG results on 40 subjects showed two posterior vMMN components in the “Ignore” condition – one reflecting early deviance detection (peaking in 125–150 ms latency in occipital and 150–175 ms latency in parietal areas) and another emerging from 250 ms post-stimulus. There was only a later posterior vMMN component for the “Attend” condition, starting from 275 ms post-stimulus, plus a positivity in Frontal region, that possibly reflects P3 and attention-related task-activity. Task relevance seems to interfere with the early deviance detection. As we discuss in Study II: “The fact that an early vMMN is not seen in “Attend” condition might reflect the executive attention process in visual modality. When the features of standard and deviant stimuli (i.e. motion direction) are actively processed for conducting a difficult primary task (as was the case in our experiment), the visual top-down attention might suppress the automatic change-detection mechanism responsible for the emergence of vMMN.” (Kuldkepp, Kreegipuu, Raidvée, Näätänen, & Allik, 2013). There is also a possible overlap with the MMN and another negative-going component N2b that is evoked by active deviance detection (Sussman, 2007).

The results of Study II strongly support the notion that motion processing is automatic. Detecting regularity violations in motion direction was quick (as indicated by the vMMN waveform first peaking around 150 ms after stimulus onset), uncontrollable, unintentional (as target events appeared in a different time-sequence with standard and deviant events), did not require focal attention and required little cognitive load, since we registered vMMN concurrently with a demanding behavioural task. Changes in the nervous system preceed the changes manifested in behaviour, as shown by even a simple comparison between the response latencies of visual stimulation in the cortex (Lamme & Roelfsema, 2000) and the response times from various psychophysical experiments (Tovée,
including the behavioural measurements presented in the studies of the current thesis. This means that in everyday situations we can easily go on with whatever we are doing, only to know that the surroundings are under constant monitoring, and that whenever needed, we can quickly react.

**Processing motion information relative to other tasks**

In Study I, we dealt with changing the spatial relationship between the target stimulus and its surrounding area, while they both had similar stimulus qualities. We found no relative motion effect, indicating that target motion was automatically processed despite of what happened in near vicinity. What happens when we attend to, or predict something to happen to one quality of a stimulus, but another quality of that same stimulus is present or changing as well? This is of utmost interest especially in case of such stimulus attributes that are supposed to be dominantly processed by different pathways in the visual system – i.e. motion information in the magnocellular, colour in the parvocellular pathway (although such completely discrete dissociations obviously cannot be done).

Researchers have been influenced by and attracted to discoveries of multiple visual areas each of which is specialized for processing particular stimulus attributes (Maunsell & van Essen, 1983; Livingstone & Hubel, 1987, 1988; Zeki, Watson, Lueck, Friston, Kennard, & Frackowiak, 1991; Grill-Spector & Malach, 2004; Orban, Van Essen, & Vanduffel, 2004; Wandell, Brewer, & Dougherty, 2005; Tobimatsu & Celesia, 2006). For example, areas coding various attributes of colour in macaque extrastriate cortex were found to be insensitive to direction (Conway, Moeller, & Tsao, 2007), while direction selective neurons have been found in high concentration in area V5/MT (Zeki, 1974; Born & Bradley, 2005). Supposing that different visual areas have different activity time courses, Semir Zeki conjectured that it would be possible to find visual demonstrations showing disparity between visual delays. For example, if an object alters its colour and motion direction at the same time, the onset of these two changes may not be perceived as simultaneous (Moutoussis & Zeki, 1997, 2008; Lo & Zeki, 2014). A perceptual advantage for colour change has been extensively reported (e.g. Moutoussis & Zeki, 1997, see Murd, 2014 for a longer overview). The whole logic behind Zeki’s experiment is based on an assumption that motion and colour are processed by two separate sets of neurons, which are essentially uninformed about the results of the other processing unit (Zeki et al., 1991; Lo & Zeki, 2014; Moutoussis & Zeki, 2008). However, the picture is not as clear-cut. It seems that motion information can interfere with colour processing, but it is related to velocity, stimulus type and also type of the task (Werner, 2007; Murd, 2014). More recently, it has been shown with human functional magnetic resonance imaging (fMRI) using stimuli that change their colour or motion that an unexpected change in one feature renders the whole
object unexpected, meaning the predictive coding of anticipated independent features spreads to other features (Jiang, Summerfield, & Egner, 2016). Logothetis and Wandell (2004) argue that although the dominating neurons in cortical areas V4 (that is predominantly associated with colour processing) and V5 (associated with motion processing) may deliver long-range signals responsible for representing different stimulus features, these areas are also heavily interconnected. This means that activity in one area may modulate the sensitivity of the other. This also explains the discrepancies between the results of single-cell recordings and fMRI blood oxygen level-dependent (BOLD) signal measurements, since single-cell recordings mostly focus on one brain area, and modulatory activity between different areas cannot be shown under those circumstances (Logothetis & Wandell, 2004).

Although it has been almost a dogma for a while that most visual attributes are processed separately, it is theoretically and even technically impossible to abstract one attribute completely from all other visual attributes. It is more likely that there is an exchange of information when two or more visual attributes are processed at the same time. For instance, if a moving object was to change its colouration, it is impossible to imagine that the colour system has no information about the constantly changing position of the object. Therefore, it is informative to study the processing of one visual attribute in relation of processing some other visual attribute. Study III was designed to investigate how the processing of different visual attributes is automatically or deliberately combined.

**Study III: Visual evoked potentials to change in coloration of a moving bar**

In Study III, we presented a setup where subjects were instructed to detect a colour change in stationary or a (slow or fast) moving stimulus. The colour change happened in a different location in each trial (in one of 10 possible locations in the middle third of the screen). This was partly a replication of a previous study conducted in our laboratory where it was demonstrated that the detection in change of colouration is not independent of the movement parameters of the stimulus (Kreegipuu, Murd, & Allik, 2006). However, in addition to reaction time, we also recorded visual evoked potentials (VEPs) to colour changes (for further details of measurement and data analysis see the Methods section of Study III). We showed again (as did Kreegipuu et al., 2006), that it takes less time to detect the change in the object’s coloration when the object is

---

6 Study III was a part of series of studies conducted in our laboratory investigating the effect motion has on perceiving other features of a stimulus. These studies as well as a more thorough overview on the literature concerning color and motion processing are summarized in a recent doctoral thesis (Murd, 2014) and will not be rewritten in the current thesis. Note also that Study III was included in the abovementioned doctoral dissertation.
moving, especially when the speed of motion is fast. These findings are not in accordance with the findings of preferential processing of colour over motion. The facilitation effect motion has on a colour-related task has been found in other studies, e.g. improving colour constancy (Werner, 2007). Werner concluded, and this is a plausible explanation for the results of Study III as well, that even when attention is directed to a colour-related task, motion is a highly salient feature that captures attention and activates a high-level motion system (Cavanagh, 1992) that influences colour processes.

There was also an effect of the position of the colour change, resulting in shorter RTs when the change took place at later positions along the movement trajectory. This means that when a stimulus was visible for a longer duration, the response was more effectively executed, which might reflect less uncertainty of the observed event. This is in line with the theory of normalization long known from vision perception (e.g. Sekuler, Sekuler, & Sekuler, 1990). Sekuler and colleagues showed shorter reaction times in longer viewing periods (starting from around 500 ms). They concluded that after a certain integration time (in their results it was 500–700 ms, but other integration periods, likely dependent on stimulus parameters, have been found, see Sekuler et al., 1990 for references), a change in the stimulus motion direction is normalized to motion onset. This theory also explains why on short stimulus presentation periods, detecting changes takes longer time than detecting onsets. It seems that this normalization process is not limited to only directional changes in the visual stimulus, but could also explain the findings in Study III, where overcoming the uncertainty of the colour change happens in the later positions of the movement trajectory.

The effect of motion on the RT results in Study III was also reflected in the VEP results. Interestingly, the amplitude of VEPs (except for N1) elicited by the colour change of a faster moving object were reduced compared to changes of slower moving and stationary objects. This is not what we initially expected (see the Introduction of Study III). We have outlined several possible explanations for such results (see the Discussion section of Study III), one of them being that different pathways (parvo- and magnocellular) organised into ventral and dorsal cortical streams are differently observable (with possible overlapping confounds) through brain electrical potentials. It has for example been argued that even when motion-onset VEPs are under investigation, the emerging P1 peak could reflect pattern-related changes and parvocellular system activity instead (Kuba, Kubová, Kremláček, & Langrová, 2007). Different cortical processes can be recorded with different accuracy by the VEPs (Tobimatsu & Celesia, 2006). For instance, seeing motion onset results in more pronounced synchronized activity of several subpopulations of direction-selective neurons, manifested in a more pronounced N2 VEP, than in case of motion reversal (Henning, Merboldt, & Frahm, 2005).

The disharmony between behavioural and physiological data is not entirely unusual. One of the explanations for the discrepancy between the behavioural
and psychophysiological results could rise from the fMRI experiments. For example, it has been shown that the BOLD signal that has been found to be proportionally related to neuronal activity (Logothetis & Wandell, 2004) is not necessarily in accordance with behavioural responses, i.e. what the observers subjectively perceive in a visual scene (e.g. Whitney, Goltz, Thomas, Gati, Menon, & Goodale, 2003). A higher mean BOLD signal accompanies the difficulty of the task requiring more processing resources over time, reflected in the cumulative activity (e.g. in increasing working memory load, Nyberg, Dahlin, Stigsdotter Neely, & Bäckman, 2009). Weaker mean BOLD signal, in turn, might indicate that less processing resources are necessary for a task that is easier. Changes in the neural level as a function of task difficulty are good criteria for evaluating the efficiency (and hence automaticity) of a perceptual process. In the context of our results, which showed shorter reaction times (i.e. more effective or easy detection) to faster moving stimulus, this might explain the discrepancies with the diminished VEP amplitudes. It is also possible that certain stimulus aspects are represented in the brain responses but the observers do not have awareness of or cognitive access to these brain representations. This is supported by the results of Study II (“Ignore” condition) as well, which showed that the events that are registered and processed by our visual system (as indicated by the EEG results) do not necessarily yield changes in behaviour.

An important lesson provided by Study III is that although participants were asked to register only changes in the target’s colouration (i.e. ignoring the target motion) it was very difficult to fully follow this instruction. The observer’s inability to follow exactly the given instruction is usually a sign of involuntary and automatic processes. Target velocity determined nevertheless the time required to detect the change in colour, which further supports the idea that the visual system does not process object’s features independently, but uses the information about one feature to predict events in another. Such results disprove the underlying assumption that colour and motion are processed in two different areas, which do not exchange information between each other. Rather the results of this study indicate that the detection mechanism considers the target motion even if it is irrelevant to the task that was given.
PROCESSING EMOTIONAL EXPRESSIONS

In the previous sections we saw that visual motion detection is a highly automatic process, even affecting the detection of other features like colour change in Study III. The importance of motion detection obviously has its roots in evolution, as it is crucial for survival; therefore, we have automatic and cognitively impenetrable mechanisms for doing this job. But moving objects are clearly not the only input that might signal behaviourally important information (e.g. threat, but why not also pleasant outcomes) – humans are social animals and have developed similarly automatic processes for the perception of other stimuli such as faces, including facial expressions of emotion. It is the expressional quality that is most likely causing the preferential processing of faces (Palermo & Rhodes, 2007).

Face detection is a quick and automatic process. Electrophysiological measurements allow assessing that with a very precise temporal resolution. The categorization if a stimulus is or is not a face happens already in the first 100 milliseconds after stimulus onset, as has been shown with human MEG and ERP recordings and animal single-unit recordings. Also, differentiating if a face carries emotional information, and whether that information has a positive or negative valence, can happen already around 100 ms (see Palermo & Rhodes, 2007, for an overview). As a comparison with other stimulus categories, Pegna and colleagues showed that while face-related significant ERP activity started already at 100 ms, activity related to objects and words started around 180–200 ms, which is considerably later (Pegna, Khateb, Michel, & Landis, 2004). It has been determined that there even exists a specific ERP component – N170 – that is sensitive to face-stimuli (Bentin, Allison, Puce, Perez, & McCarthy, 1996; note that N170 is also sensitive to bodies, see e.g. Hietanen & Nummenmaa, 2011). This appears to be the time range where information about the specific identity of a face can already be extracted.

Without going into much detail about the evolution of emotion expressions (see e.g. Shariff & Tracy, 2011 for a longer discussion on that), I bring out that many current theorists support the “two-stage model” of the evolution of emotional expressions. This idea has its roots already in Charles Darwin’s “The Expression of Emotions in Man and Animals” originally published in 1872 (2013)7. In the two-stage view, emotional expressions have two functions8 –

---

7 It must be noted, though, that Darwin himself (although it has been falsely attributed to him) did not consider expressions to be adaptive or carry communicative meaning. Rather he thought that emotional expressions are vestiges from once useful gestures associated and practiced with emotions that now have no use, but still keep appearing with those said emotions.

8 Barrett (2011) argues, that emotional expressions might not carry a signaling value (more in line also with Darwin’s original ideas), but instead are socially learned symbols. This notion is also supported by studies on congenitally blind subjects, who display only a limited
first, they evolved to prepare the organism for reacting adaptively by eliciting a series of physiological responses (e.g. widening the eyes in case of surprise to see the surprise-inducing stimulus better); second, in a later stage, they developed to provide the means for social communication, by both expressing the emotions/intentions and interpreting those of others (Shariff & Tracy, 2011; Öhman, Lundqvist, & Esteves, 2001). Emotional expressions are not restricted to human species (think about an angry dog uncovering its fangs for example), which speaks for their important communicative role.

Taking the evolutional significance of survival into account, it seems to be only logical that threat-inducing signals, including threatening faces, are preferentially processed and elicit a sympathetic autonomous response and mobilise the organism for an appropriate behaviour (in a very robust approach, either fight or flight, as stated by Walter Cannon already in 1930’s). Threat-inducing signals are associated mostly with fear, anger, and disgust amongst basic emotional expressions (Ekman, 1999). Seeing a fearful face signals us that there is a source of danger in the surrounding environment; disgust might give information about, for example, inedible and contaminated food; an angry face could mean a potential verbal or physical attack to us. The threat/anger superiority hypothesis (orienting attention preferentially to facial gestures that signal threat) in detecting emotional faces was formulated already thirty years ago (Hansen & Hansen, 1988) and has been supported by a large number of empirical studies (see Lundqvist, Juth, & Öhman, 2014 for references on experiments using the visual search paradigm; Stefanics et al., 2012, as an example of the negativity bias from the vMMN field). It has also been shown to persist even if the whole-face upright configuration is missing – in case of only presenting eyes (with eyebrows) as stimuli (Fox & Damjanovic, 2006), in case of inverted faces (Öhman et al., 2001; but opposite results are also found, e.g. Tanaka & Farah, 1993; Fox & Damjanovic, 2006; Calvo & Nummenmaa, 2008), or even when only low-level features such as V shapes (resembling eyebrows in such a direction that are considered to signal threat) are presented (Larson, Aronoff, & Stearns, 2007; Larson, Aronoff, Sarinopoulos, & Zhu, 2009). However, there are opposite results, as for example Weymar and colleagues showed that the repertoire of facial expressions. However, even the learned symbolic communication is evolutionally significant, as it seems to serve the purpose of perceiving the intentions of other members of the species.

Interestingly, although Hansen and Hansen’s work is thought of as one of the pioneer studies supporting the angry superiority in face detection (an angry face popping out among happy ones faster than vice versa), it has later been criticized for being methodologically challenged. Purcell and colleagues argued that Hansen and Hansen, while modifying the original Ekman and Friesen emotional expression stimuli, created several contrast artifacts (e.g. dark or light blobs attracting more attention) that account for their results (Purcell, Stewart, & Skov, 1996). By using the same Ekman and Friesen pictures in an original grayscale version, Purcell and colleagues found that happy faces were found faster in angry crowds, and that the processing appeared to be serial.
threat advantage (as indicated by a larger N2pc ERP component amplitude related to spatial selective attention) emerges only to the whole-face configuration in the angry-happy schematic face comparisons, not to “eyes with eyebrows” or “eyebrows only” stimuli (Weymar, Löw, Öhman, & Hamm, 2011). On one hand, this is in line with the notion that faces are holistically processed (Tanaka & Farah, 1993). On the other, the results showing also preferential processing of certain extracted parts of a face that might convey emotional meaning, speak in favour of a more general system that is tuned to detecting various threat-inducing input, not only faces. One of the promising approaches to the mechanisms of how threatening stimuli – including for example, angry faces – are preferentially, quickly and automatically processed, proposes a “module of fear”, a subcortical rapid information processing route to amygdala (Öhman & Mineka, 2001; Öhman, 2002).

However appealing (with the backup by a vast amount of research) the idea of angry/threat superiority might seem, there is also much research speaking in favour of detecting happy (i.e. not fear-eliciting) stimuli faster and/or better than angry or fearful targets (e.g. Juth, Lundqvist, Karlsson, & Öhman, 2005; Calvo & Nummenmaa, 2008 (showing also that sad expressions were most poorly detected); Becker, Anderson, Mortensen, Neufeld, & Neel, 2011). The happy face advantage probably has communicative importance, signalling social cues for acceptance and affiliation (Becker et al., 2011). One of the explanations for the heterogeneous results on either angry or happy superiority in face detection differentiates between emotional and perceptual factors contributing to preferential processing of stimuli (e.g. Calvo & Nummenmaa, 2008). In this view, it is not necessarily the affective meaning that is automatically differentiated in faces, but specific physical features are faster processed than others are because of their distinctiveness. Calvo and Nummenmaa (2008) for example argue that the happy face is more visually salient, because of the distinctive features (especially in the mouth region), resulting in shorter detection times, and only after that, decisions about stimulus identity can be made. However, perceptual factors alone cannot account for the processing of emotional faces; there is also much research supporting the dominating role of emotional, not perceptual factors (see the discussion on that and references in Lundqvist et al., 2014). For example, Becker and colleagues showed that happy faces were detected better than angry ones even when the former were smiling with closed lips (as opposed to the classical Ekman and Friesen face pictures where the preferential detection of happy faces could be due to the exposed white teeth) and even when the lower half of the face was removed (Becker et al., 2011). It is worth noting that the threatening face processing advantage in a visual search task seems to become more unstable with stimuli other than schematic faces (Juth et al., 2005; Lundqvist et al., 2014). The negativity bias is also dependent on task type – for example while visual search tasks, where detecting negative expressions is a task, usually report it (Lundqvist et al., 2014; Carretié, 2014), categorization tasks show that negative expressions are more slowly and less accurately categorized.
(see the overview on that in Schlaghecken, Blagrove, Mantantzis, Maylor, & Watson, 2017). Also, as a recent review and meta-analysis on exogenous attention experiments with emotional stimuli (including faces) showed, although over 90% of the reviewed studies found the negativity bias, only under half of those presented positive stimuli as well (Carretié, 2014). Such counterbalancing is, however, necessary, to say something about the valence effects. Recently, Lundqvist and colleagues proposed (backing it up with re-analysing their own previously published data, as well as available data from other research groups) that the discrepancies found in whether threat inducing or non-inducing stimuli are preferentially processed, might be explained by the arousal value of emotional stimuli, instead of only their valence (Lundqvist et al., 2014). The arousal elicited by positive or negative faces could be in the same magnitude, even if the detection task results show different positive-negative valence and attention relationships. This has also received support in the meta-analysis of Carretié (2014), who reports more arousing positive pictorial stimuli (i.e. with sexual content) contributing to the positivity bias. Research regarding ERP responses to affective pictorial stimuli (not specifically faces) have also shown the different effects the two categories – valence and arousal – have on ERPs (for a review see Olofsson, Nordin, Sequeira, & Polich, 2008). Valence effects, thought to represent automatic selective attention capture in case of salient stimuli, are present in earlier ERP latencies (100–250 ms), while arousal influences later ERP components (starting from 200 ms post-stimulus, including P300) which indicates automatic attention to motivational stimuli that serves the subsequent memory encoding (Olofsson et al., 2008).

Knowing that information about emotional expressions is analysed quickly in a rather effortless manner (for example, detecting emotional faces among non-emotional), and having a tool (i.e. the vMMN) that allows to study early change detection, it seems only logical to combine the two in studying automatic processing.

**Study IV: vMMN for schematic faces: automatic detection of change in emotional expression**

Study IV addressed the question of automatic detection (indicated by the vMMN) of emotional content. We used emotional schematic faces (neutral, angry, happy) in two different visual MMN paradigms – “classical” oddball and an optimal (or optimum) multi-feature paradigm (Näätänen, Pakarinen, Rinne, & Takegata, 2004). Schematic faces as stimuli have been shown to have high ecological validity (Öhman et al., 2001; Lundqvist et al., 2014), and are easily controllable with for example no gender or age attributes, or low level perceptual artefacts as confounding variables.

Based on previous studies which demonstrated that threatening facial expressions can be detected faster and before other expressions (Hansen & Hansen,
we initially proposed that our brain is more ready to form an internal model of a fear-inducing, i.e. angry, stimulus. Because the MMN technique is based on the idea that our brain is building such models (see the more thorough insight into that on pp. 33–34 of the current thesis), which help us automatically process environmental information, we expected that angry (i.e. possibly fear-inducing) expression has a special status reflected in the MMN signals. Recently, Kovarski and colleagues provided a compact overview of the vMMN studies using facial expressions as stimuli (Kovarski, Latinus, Charpentier, Cléry, Roux, Houy-Durand et al., 2017). Interestingly, in the vMMN field, more research has been done with fearful and even sad negative stimuli than angry ones (see Table 1 in Kovarski et al., 2017), which makes Study IV especially valuable.

In addition, one of the explorative aims of Study IV was to compare the “classical” oddball vMMN paradigm with the optimal (or optimum) multi-feature paradigm in a within-subjects experimental design that had never been done before in the visual domain. Importantly, this was the first vMMN study using an optimum paradigm (although there were studies using two deviant stimuli presented equiprobably, see the Introduction of Study IV for more details). Four conditions with 85% standard stimuli (S), 12.5% deviant stimuli (D) and 12.5% target stimuli (scrambled “face”, i.e. a non-face object) were presented in the oddball experiment: a) neutral S, angry D, b) neutral S, happy D, c) angry S, neutral D, d) happy S, neutral D. In the optimum experiment, three conditions were used, each featuring a different standard (either neutral, happy or angry) and other stimuli (including their inverted versions, as well as S presented as a deviant, and an inverted version of a S, not analysed in the Study) presented as deviants, as well as 12.5% non-face targets. Note that the conditions were counterbalanced in both experiments. Behavioural measurements included manual RT-s to target detection (showing high performance of the subjects in both experiments) as well as a post-experiment questionnaire, where the subjects rated the valence, arousal and attention-capture of stimuli, labelled them and rated their strategy.

The main results of Study IV showed that emotional stimuli (regardless of their valence) were preferentially processed in the occipital and parietal scalp locations compared to the neutral ones, confirming that emotion information is extracted from stimulus configurations relatively fast and without the need for conscious effort. These results, together with the results of subjective stimulus ratings showing no significant valence differences, are in line with the “arousal hypothesis” recently proposed by Lundqvist and colleagues (Lundqvist et al., 2014). There were equally good vMMN results in both experimental paradigms – oddball and optimum –, supporting the use of the optimum multi-feature paradigm in the future. The vMMN results did not solidly confirm a proposal made by Arne Öhman and colleagues (Globisch, Hamm, Esteves, & Öhman, 1999; Öhman et al., 2001; Öhman, 2002; Schupp et al., 2004) that threatening
stimuli are faster and easier detected than others. However, in case of the angry deviant stimulus, we see that the processing negativity was earlier and was not dependent on the type of the standard stimulus, as was the processing of the happy deviant. Also, the results of the post-experimental questionnaires showed that the angry face that the subjects were not specifically directing their attention to (as it was presented in the vMMN-eliciting sequences, while the subjects performed a demanding target-detection task) attracted more subjective attention than the happy face, even almost the same as the target. This, again (as in Study III), demonstrates that ERPs correspond to only one aspect of an internal representation and that decisions about responses are not necessarily based on this aspect.
VISUAL MISMATCH NEGATIVITY- AN AUTOMATIC CHANGE-DETECTION MECHANISM

The auditory MMN was discovered in 1978 by Risto Näätänen and colleagues (Näätänen et al., 1978). By now, it is widely used in cognitive neuroscience in basic research as well as in clinical populations (see Näätänen et al., 2011 and Näätänen et al., 2012 for reviews). Establishing the existence of a similar mechanism in visual modality evolved in the nineties and took some time. Although many researchers searched for a visual analogue of the auditory MMN and even claimed it existed (Cammann, 1990), most of the MMN community at first remained rather sceptical. The first review by Pazo-Alvarez and her colleagues made the existence of the vMMN plausible (Pazo-Alvarez et al., 2003), but as the authors of the review themselves stated 15 years ago: “In most cases, the authors have only checked for the emergence of the characteristics of the auditory MMN without considering the peculiarities of the visual system”. So, several unresolved issues remained, allowing at that time to be sceptical concerning the final verdict on the existence of the vMMN.

The dominating view on how changes are detected in the visual environment mostly emphasises the need of focal attention (for a review see Rensink, 2002). Phenomena such as change blindness (inability to detect changes in a scene or an object; Rensink, O’Regan, & Clark, 1997), attentional blink (failure to allocate attention to the next stimulus and perceive it in a rapid sequence, when the ISI between the first and the next stimuli is between 200–500 ms; Raymond, Shapiro, & Arnell, 1992), or inattentional blindness (failure to perceive unexpected objects when attention is strongly focussed elsewhere; Mack & Rock, 1998, cited in Simons, 2000; the most famously known experiment being the “invisible” Gorilla experiment, Simons & Chabris, 1999) seem to agree with the view that without visual attention, changes in the environment can escape the viewer’s awareness. On the other hand, even if the person is not aware of the change, it could still affect the processing of visual information and his/her (behavioural) reactions. This is evident for example from brain lesions research (e.g. blindsight; Sanders et al., 1974), as well as studies on priming (Tulving & Schacter, 1990). It would also be unreasonable to assume that we are able to allocate all our attentional resources to just detecting and processing all the changes/events in the environment, since the amount of incoming information via our visual system in parallel with our other senses is enormous and happening at the same time we are engaged in various activities. On the contrary – most of the processing of the surrounding environment and changes happening there goes on automatically with an ability to quickly direct one’s attention to events that need a reaction (orienting response, Sokolov, 1963) and a possibility to ignore events that do not require reacting at that particular time. One might say that our brain is constantly monitoring the surrounding environment and processing the incoming information without us even realizing that most of the
time. By now, this is also evident from vMMN research, which has quickly grown in the last decade. According to the Web of Science © (Clarivate Analytics) database (search conducted on December 7th, 2017), 141 publications (articles, reviews and editorial materials, excluding conference/meeting abstracts) are related to the keywords “visual mismatch negativity”, “visual MMN” or “vMMN” from 2008–2017; the number being only 23 before 2007 (and included)\(^{10}\), i.e. before the planning of the empirical work of Study II presented in this thesis began. The results of Study II and Study IV also confirm that vMMN emerges to sudden changes in different visual input, doing so without necessarily directing attention to those changes.

There is no question whether MMN exists in visual modality anymore. Instead, the research and debate has moved forward to determining its underlying mechanisms. There have been several views (see a thorough overview of these in Kimura, 2012), starting from the sensory memory trace explanation in oddball paradigms (first excessively supported in auditory modality, see Näätänen, 1992), evolving to the regularity-violation explanation (Czigler, 2007) and abstract sequential rule violation explanation (Stefanics et al., 2011). What is important to note is that none of those could uniformly explain all the results, which has left the search for a uniform underlying mechanism open. Recent views on the underlying mechanisms of vMMN move more and more towards the agreement that the vMMN indicates the brain’s ability to predict the statistical regularity of events and detect disturbances in this regularity (a prediction-error explanation, Stefanics et al., 2014). The basis for that is perceptual learning (tracing back to James Gibson, e.g. Gibson & Gibson, 1955) – the environmental regularities (from the presentation of frequent and repeating standard events in the MMN paradigms) are implicitly learned and based on them, predictions are generated. If a rare deviant event violates a prediction, a prediction error occurs, which is manifested in the mismatch response. This process, as visual perception in general, relies on forward and backward\(^{11}\) connections operating together in the hierarchical structure of the visual system.

---

\(^{10}\) This is just a superficial example of the difference in article quantities, not intended to find all the studies related to MMN in visual modality. For example, when we look at the studies found from the time period up to 2007 by this type of basic search, and compare the results with Pazo-Alvarez et al. (2003) and Czigler (2007) vMMN reviews, it is clear that many studies are missing from the search results. These are for example intermodal studies (Cammann, 1990) or studies referring to processing negativity, not vMMN (see Pazo-Alvarez et al., 2003 Table 2).

\(^{11}\) Traditionally (Maunsell & van Essen, 1983; see also the “reverse hierarchy theory” by Hochstein and Ahlísar, 2005) the connections from lower to higher level are termed „feedforward“, and the connections from higher to lower level „feedback“. Friston (2005) terms the connections “forward” and “backward”, respectively. As Friston (2005) argues, the purpose of predictive models in the brain is to generate sensory data from high-level causes, meaning that „forward connections provide feedback by conveying prediction error to higher levels“. For the clarity, term „feedback“ is not used here in the terminology of the dual division.
According to the predictive coding view, backward connections carry the information about predictions that either match with the incoming sensory input (in which case the input is “explained away”) or do not match, which results in the prediction error.

The predictive coding view on cortical responses (Friston, 2005) emphasizes the role of plasticity in neural connectivity that mediates the perceptual learning taking place in (v)MMN generation through backward connections. More and more research supports the role of the NMDA (N-methyl-D-aspartate) glutamate receptors in this plasticity (Coyle, Tsai, & Goff, 2003), as they are found mostly in the supragranual cortical layers where the mediating backward connections terminate, and as they are voltage-sensitive and with slow temporal dynamics (Friston, 2005). The NMDA receptor system functioning has been associated with the functioning of long-term memory, as well as working memory, including estimating stimulus familiarity/novelty (Javitt, Stein Schneider, Schroeder, & Arezzo, 1996). While the empirical studies incorporated in this thesis do not investigate the neural underlying mechanisms of vMMN, this topic is of importance at least briefly to mention, when addressing the relationship between the vMMN and various clinical conditions. Based on drug challenge studies, it is suggested that the MMN deficits in the auditory modality are mediated mainly by the NMDA receptor hypofunction. For example, MMN amplitude is decreased or even abolished by NMDA receptor antagonists ketamine, MK-801, or PCP (phenethylidione) (see Umbricht & Krljes, 2005; Näätänen & Kähkönen, 2009; Näätänen et al., 2011, or Todd, Harms, Schall, & Michie, 2013 for reviews). NMDA receptor hypofunction is also hypothesized to play a role in schizophrenia, as the pharmacological NMDA antagonists induce sensory and cognitive deficits that resemble those seen in the disease (for the NMDA/glutamate hypothesis of schizophrenia, see e.g. Javitt et al., 1996; Javitt, 2009; Javitt, Zukin, Hersco-Levy, & Umbricht, 2012). The reduced aMMN amplitude in schizophrenia (to especially duration and frequency changes) is a widespread result (Umbricht & Krljes, 2005; Näätänen & Kähkönen, 2009; Damaso, Michie, & Todd, 2015). Recently, the relationship between glutamatergic neurotransmission and duration MMN amplitude was also shown without any pharmacological manipulations on healthy subjects (Kompus, Westerhausen, Craven, Kreegipuu, Pöldver, Passow et al., 2015). In the visual modality, there are also reports on the role of the NMDA receptor system in sensory processing deficits (Javitt, 2009). Studies investigating the relationship between vMMN and glutamatergic neurotransmission are probably something we will see in the near future to help disentangle the underlying mechanisms of vMMN.

By now, vMMN is also used in clinical research, and predictably more and more reports on that are yet to come. In addition to the fact that vMMN enables to study automatic visual processing without the necessary involvement of the subject’s focal attention, or investigate change-detection processes in case for...
example auditory modality cannot be accessed, there are also clinical conditions, where visual processing in particular is disturbed, making vMMN a useful tool in investigating such disturbances. Visual impairments have been reported in several neurodevelopmental and psychiatric disorders (Laycock et al., 2007). To bring out just a few examples, it has been suggested that impaired temporal processing in the magnocellular pathway contributes to developmental dyslexia (e.g. Talcott, Hansen, Willis-Owen, McKinnell, Richardson, & Stein, 1998; Wang et al., 2010); early visual processing is impaired in schizophrenia (Javitt, 2009). The first comprehensive review on vMMN in clinical research was published in 2016 by a group of vMMN researchers, including the author of the current thesis.

**Study V: Visual Mismatch Negativity (vMMN): A review and meta-analysis of studies in psychiatric and neurological disorders**

Study V gives an overview and presents the meta-analysis of the current research of vMMN in clinical populations, in case of substance abuse (alcohol, nicotine, metamphetamine), as well as in different age groups to assess the potential changes in vMMN during maturation and ageing. The clinical conditions covered by the studies reviewed are: schizophrenia, mood disorders (major depressive disorder and bipolar disorder), neurodegenerative disorders (Alzheimer’s disease, mild cognitive impairment, SCA2), developmental disorders (autism, dyslexia, mental retardation), deafness, panic disorder, and hypertension. A very general conclusion from Study V is that the vMMN is altered in all those conditions under observation. This means either the expected vMMN reduction (as is the case in, for example, schizophrenia) or enhancement (as is expected, for example, in case of nicotine consumption, which is supposed to sharpen the primary processing of unattended sensory input and therefore enhance cognitive performance). Similar findings have been reported in aMMN (see Näätänen et al., 2011, 2012). When taking a closer look at the 33 studies investigated in Study V, we see that there is a great methodological variety in which stimuli and in which paradigms have been used, how vMMN is calculated/presented, whether standard and deviant (or control and deviant) responses have been statistically compared, whether and which control groups have been used, and whether vMMN has been correlated to any other (clinical) measurements (e.g. test scores, condition severity, age). Without rewriting the entire Study V here, I will bring out just a couple of examples. Results regarding vMMN in mood disorders seem to be vague, as the statistical importance of the differential response between standard and deviant stimuli (i.e. the vMMN) has not been assessed in the papers, and there seems to be a dependency on the stimulus type. Also, no clear tendency can be drawn from the vMMN studies in Alzheimer’s disease or in aged subjects, where there are heterogeneous results. On the other hand, there are rather homogenous results in case of, for example, schizophrenia,
where patients of all four studies demonstrate smaller or no vMMN response. The meta-analysis of effect sizes and power conducted in Study V has promising results, as the median power calculated across 27 studies that had the necessary parameters presented, was 0.77 (0.61; 0.85 — the first and third quartiles, respectively).

Since the time the review paper was published, there are five additional reports indexed in the Web of Science © (Clarivate Analytics) database (online search conducted on December 7th, 2017) regarding the vMMN used in clinical and special populations, generally in line with the heterogeneous results reported in Study V. A study using emotional schematic facial stimuli confirmed the reduced vMMN of the patients with schizophrenia compared to healthy control subjects (She, Li, Ning, Ren, Wu, Huang et al., 2017). Two of the studies investigated age-related changes in vMMN and again, as in Study V, no homogeneous conclusions about aging effects can be drawn. Gaál and colleagues showed that in some cases aging might improve the processing of deviants via the mechanism of longer stimulus integration time (Gaál, Bodnár, & Czigler, 2017). The second study using numbers and letters as stimuli (Kieffaber, Okhravi, Hershaw, & Cunningham, 2016), showed diminished vMMN amplitudes (amongst other auditory and visual ERPs) in aged compared to young subjects. An interesting study compared daily nicotine smokers with non-smokers (Stothart, Maynard, Lavis, & Munafò, 2016), presenting them with the warning signs found on cigarette packages. The results showed no differences in the change detection magnitude (as indicated by the vMMN amplitude) between the two subject groups, although for the smokers, the vMMN latency was delayed. Neither were there differences in early perceptual (as indicated by P1) or attentional (as indicated by P3) responses, but only in higher level emotional processing (as revealed by the LPP that started considerably later and had a lower amplitude in the smoking subject group). The only other vMMN study investigating nicotine effects available and reported in Study V (showing vMMN enhancement) was that of Fisher and colleagues (Fisher, Scott, Shah, Prise, Thompson, & Knott, 2010), but since nicotine was acutely administered in the latter, the two studies cannot really be compared. Liu and colleagues (Liu, Xiao, & Shi, 2016) studied detection of facial expressions in adolescents. The adolescent group had larger fronto-central vMMN amplitudes in an early latency range (100–200 ms) compared to the adults. Also, early vMMN amplitudes were larger only in fronto-central than occipito-temporal locations in the adolescent group, possibly reflecting the heightened sensitivity and enhanced prefrontal activity to affective stimuli in adolescence. Stimulus emotional valence (happy or sad) was significant only in the adult group. As we argue in Study V, the vMMN maturates differently for separate visual functions, and it seems that for emotional expressions, this maturation is not complete at adolescence (the mean age of the group was 14.6).

As we see in Study V, measuring the vMMN in clinical or special populations calls for more rigorous testing of different experimental protocols to
have a more unified understanding and agreement on methodological approaches. Both Study II and Study IV have improved the field in that matter. Study II is one of the few papers in the field where the occurrence of vMMN is investigated in two different attentional settings using the same vMMN-eliciting stimulus sequence – a condition where the vMMN eliciting stimuli are ignored, as well as a condition, where attention is directed to exactly the same stimulus sequence. Such direct juxtapositionings are necessary to confirm that vMMN is independent of attention, which is one of the established prerequisites of the component (Näätänen et al., 2007). In addition, we propose (based on the results of Study I as well) to use a stimulus configuration, where the attention-capturing target detection task is presented in the center of the visual field, while the vMMN-eliciting stimuli are presented in the periphery and the stimulus presentations are independent of each other in time. (See the Discussion pp. 7–8 of Study II for a discussion on the significance of the target location both in space (i.e. different locations on the screen) and time (i.e. either in the same sequence with the vMMN-eliciting standard and deviant stimuli, or separate from the latter).) Recently, Stefanics and colleagues reviewed different experimental protocols used for diminishing the attention effects on the vMMN (Stefanics et al., 2014). The authors concluded that “even if the level of attentional control in vMMN studies is highly variable, the results of the various studies have been remarkably similar, since their overwhelming majority has reported negative-going deviant minus standard ERP components with posterior scalp distribution in the ~100–400 ms range (Stefanics et al., 2014). However, as we also discuss in Study V (in section 2.6. Aging and maturation), the non-efficient attention control in some of the studies might explain the poor vMMN results.

Study IV presents several methodological gains. Most importantly, it is the first study comparing the oddball and optimum paradigms in the visual domain. The results showed similarly solid vMMN for angry and happy faces in both paradigms. This is a significant encouragement for using the optimum measurement paradigm, where more deviant stimuli can be presented in a shorter amount of time. Also, as can be seen in the vMMN literature, there are different ways of presenting the standard/deviant (or control/deviant) difference, i.e. what the vMMN is. Most commonly the vMMN waveform is a difference wave between (mean) deviant and standard/control stimulus activity, presenting voltage changes (amplitude) as a function of time. It is possible to look at the area under curve (AUC, e.g. Urban et al., 2008), the highest amplitude (voltage) peak in the (usually before chosen) expected latency range (Fisher et al., 2010), the mean amplitude in a certain (before determined) latency range (either choosing one time window around the highest peak of the vMMN curve, e.g. Astikainen et al., 2008; or defining several time windows, e.g. Amenedo et al., 2007), or oscillatory activity (Stothart & Kazanina, 2013). In Study IV, two differently calculated vMMN estimates (from the same data) were compared: a) maximal negative peak of the waveform (determined as the mean value of the peak and two of its neighbours); and b) mean amplitude in five time intervals, which
were determined by visual inspection of the mean waveforms, as well as the literature. We showed that the two estimates are very similar to each other, as their mean linear product-moment correlation was 0.961 (minimum 0.943, maximum 0.971), which at least lets us take it a bit more easy on choosing the “right” parameter for vMMN calculations.
CONCLUSIONS

The research presented in the current thesis used different behavioural and brain imaging methods, to target the underpinnings of automatic processing in visual perception. We focused on two examples of visual input that are thought to be processed in a rather automatic way and carry high value for a person to survive in the physical and social environment – motion and emotional expressions. In addition, we investigated the emergence of visual mismatch negativity that reflects the brain’s automatic change-detection ability, and showed that largely involuntary processes exist, which monitor visual input, categorize it, and detect deviations from a “standard” input.

This thesis can be summarized by the following statements, which recapitulate the novelty and main findings obtained in the five presented studies:

1. Although it is widely believed that no motion can be perceived without establishing a frame of reference with other objects and/or motion on the background, Study I demonstrated that the onset of motion is detected based on the retinal motion, which is not related to the context of other objects and motions into which the target is embedded. This finding speaks in favour of a simple and automatic process of detecting motion, which is largely insensitive to the surrounding context.

2. Cortical electrical potentials elicited by the change in peripheral motion in Study II were similar irrespective of whether observers attended or not to the change in the motion parameters. Again, this suggests there is an automatic mechanism for the detection of motion, which cannot be interfered by the overt attention.

3. Although observers were instructed to detect the colour change, they were not able to ignore the velocity with which the targets were moving in Study III. The fact that velocity facilitated the detection of changes in colour shows that processing of these two attributes – motion and colour – is not entirely isolated from one another. Also, it indicates that the observer can be unable to fully follow experimental instructions, and might rely on automatic mechanisms with a prefixed visual attribute instead.

4. Participants had a predisposition to detect emotional facial expressions faster and more easily compared with neutral facial expressions, with a tendency towards more automatic attention to angry stimulus in Study IV. These results confirm the existence of an automatic process for analysing the stimulus configuration and its match to a specific combination of visual elements.

5. The MMN is one of the most objective and efficient methods for analysing automatic processes in the brain. Well validated for the acoustic domain, a visual analogue of the MMN remained elusive for some time. Study II and Study IV contributed to proving the existence of vMMN, as well as proposed several methodological gains.
Recently, as a member of a group of authors, we suggested that vMMN could be used for the analysis of automatic processes and their alterations in the visual modality that can be more sensitive to various pathological changes. Study V is an important contribution to the vMMN field as it is the first comprehensive review and meta-analysis of the vMMN studies in psychiatric and neurological disorders.
ACKNOWLEDGEMENTS

The research included in this dissertation has received support from the following sources: the Estonian Ministry of Science and Education and the Estonian Research Council (Institutional Research Grant IUT02-13 and SF0180029s08), the former Estonian Science Foundation (grants #8332, #8231 and #6244), the European Social Fund (Primus grant #3-8.2/60 and the financing of the Doctoral School of Behavioural, Social and Health Sciences), European Regional Development Fund (University of Tartu ASTRA Project PER ASPERA: Doctoral School of Behavioural, Social and Health Sciences), and the “Norwegian financial mechanism 2009–2014” through the Norwegian-Estonian Research Cooperation Programme (project EMP180).

How many supervisors does it take to supervise one PhD student? In my case, the number is three and I have benefitted enormously from this arrangement, because each one of them contributed with their own knowledge and specialty. Jüri Allik has been the one guiding my first and ongoing steps in the field of visual perception and psychophysics, which is a fascinating topic of research. In addition to appreciating base research and experimental work, he has also taught me a lot about scientific writing and the rules and tricks of the academic world. Risto Näätänen has been an irreplaceable source of guidance and knowledge in the mismatch negativity research. With Kairi Kreegipuu I have worked the closest for the past 10+ years of doing research. Her bright ideas and sharp thinking as well as methodological thoroughness have been truly motivating. Thank you all for supervising me!

I have had the opportunity to work in the warm and friendly collective of the Institute of Psychology. I therefore thank all my colleagues for creating this work environment. My special thanks goes to the friends, and present and former colleagues from our small research group of experimental psychology, especially Maria Tamm, Carolina Murd, Aire Raidvee, Ainika Jakobson, Andero Uusberg, Helen Uusberg, Kristiina Averin and Richard Naar, with whom I have done research together, but also shared the ups and downs of graduate studies. Carolina and Andero, thank you also for your thoughtfull comments on this thesis. Special thanks to Tiit Mogom for solving all the technical issues, Mai Toom for programming help, Hels Hinrikson and Katrin Kern for language advice, and Katre Sakala and Ene Karja for their work and help in the sometimes inscrutable world of bureaucracy. There are several (scientific) friends, colleagues and co-authors who have shaped my views in different scientific questions and have been inspiring throughout my graduate studies – Juho Strömmer, Piia Astikainen, Jan Krémlacek, Mari-Liis Kaldaja, Kadri Kõiv, Act O’Leary, Kelli Lehto, and Kristiina Kompus. Thank you all!

I have always felt the support of my family (especially you, mom and Kaie!) and friends in pursuing my PhD and choosing the academic career, and for that, I am grateful to each one of you. To some of you, though, I would like to say – yes, working as a researcher is a real job.
Mati, thank you for your love and support, and for gently forcing me to mark down the thesis-writing days in the calendar and stick to them. I could not have done this without you. And my dear little Madli, without even knowing it yet, you have put things in perspective for me, and for that I am truly grateful. I love you both so much!

Last, but not least I thank everyone who has participated in my experiments throughout the years.
REFERENCES


P300, and N400. Clinical Neurophysiology, 120(11), 1883–1908.


Näätänen, R. (2001). The preception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*(1), 1–21. doi:10.1111/1469-8986.30810


Näätänen, R., & Kähkönen, S. (2009). Central auditory dysfunction in schizophrenia as revealed by the mismatch negativity (MMN) and its magnetic equivalent MMNm: a review. *International Journal of Neuropsychopharmacology, 12*(1), 125–135. doi:10.1017/S1461145708009322


Tovée, M. J. (1994). Neuronal Processing: How fast is the speed of thought?. *Current Biology, 4*(12), 1125–1127. doi:10.1016/S0960-9822(00)00253-0


EESTIKEELNE KOKKUVÕTE

Nägemistaju automaatsete protsesside eksperimentaalne uurimine


55

liiguvad üksteise suhtes samas või vastupidises suunas. Lisaks reaktsiooniaegadele mõõttime katses osalejate aju elektrilist aktiivsust elektroentsefalograafia (EEG) meetodiga. EEG tulemused näitasid, et nägemissüsteem töötleb väga edukalt ka seda liikumisinformatsiooni, millele vaatleja teadlikult tähelepanu ei pööra. Taustal toimuv suunamuutus avastati eholahknevusesnegatiivsus tekkis mõlemas eksperimendis – nii juhul, kui inimene oli hõivatud tahelepanu nõudva põhüülseandega (eksperiment 1), kui ka juhul, kui inimene suunas tähelepanu nii eesmärkstiimuli kui ka taustale (eksperiment 2). Huvitav on see, et esimeses eksperimendis, kus katses osalejatel oli palutud taustal toimuvat ignoreerida, toimus erinevuse avastamine ajaliselt varem. Samas ei mõjutanud taustal töimuva töötlemine ekraanil esitatud eesmärkstiimuli samaaegset töötlemist, mida näitasid reaktsiooniajad. See tähendab, et samal ajal, kui inimene on mõne tähelepanu hõlmava tegevusega ametis, suudab tema aju ümberringi toimuvaid sündmusi automaatselt registreerida.


Esimened kolm empiirilist uurimust käsitlesid liikumisinformatsiooni töötlust, mis on evolutsiooni käigus kujunenud olulise isendin. Lisaks liikumisele on aga veel teisi, mis on eluühised ja keskkonnas hakkamasaamise seis-

57

paradigmades ning vajalikke metodoloogilisi täiendusi. Kuulmismodaalsuses on
lahknevusnegatiivsus uuritud väga paljudel erinevatel kliinilistel valimitel ning
järeldatud, et lisaks sensoorse eristamise defitsiidile, tähelepanu kontrolli langu-
sele või tajuprotsesside muutumisele seostub MMN-i häirumine ka kognitiivse
olevat kasulik abivahend kliiniliste valimite uurimiseks, näiteks seisundi ja ravi
efektiivsuse hindamiseks. Uurimus V on esimene kogu siiani ilmunud temaatilist
teadustööd hõlma ülevaatartiikkel ja metaanalüüs visuaalselt lahknevusnegati-
iviivusest psühhiatriiliste ja neuroloogiliste haiguste korral, mis panustab olu-
liselt visuaalse lahknevusnegatiivivuse uurimise ja kasutamise arengusse. Lisaks
psühhiatriiliste ja neuroloogiliste haigustele (skisofreenia, meeleoluhäired
(nt depressioon), paanikahäire, neurodegeneratiivsed haigused (nt Alzheimeri
tõbi) ja psühhoogilise arengu häired (nt autism)) on Uurimeses V käsitletud
teadustöödes uuritud lahknevusnegatiivivuse muutusi ka kõrgvererõhutõve,
kurtuse, mõnainete (alkohol, nikotiin ja metamfetamiin) tarvitamise korral ning
eri vanuses inimestel. Kokku käsitles Uurimus V 33 tolleks ilmunud tea-
dustööd. Väga üldistatult võib öelda, et kõikide vaadeldud haiguste, ainete tarvi-
tamise ja ealiste iseärasuste korral toimus aju automaatses erinevuse avastamise
võimes muutus. See muutus võis olla negatiivne ja väljenduda vMMN-i ampli-
tuudi vähemises, nagu näiteks skisofreenia või vananemise korral, kuid teatud
juhtudel oli muutus ka vastupidises suunas, näiteks nikotiini suurendas vMMN-i
amplituudi. Teadusartiklite metaanalüüs näitas, et uurimustest statistiline me-
diaanvoimsus oli 0.77, mis on oluliselt kõrgem, kui neuroteaduse vaidkonnas
on muidu leitud. Uurimuse V põhjal võib öelda, et ka visuaalses modaalsuses on
lahknevusnegatiivivuse suur potentiaal kliiniliste valimite uurimisel, mille
üheks eelduseks on katseparadigmade ja kasutatava stiimulmaterjali suurem
ühtlustamine.
PUBLICATIONS
CURRICULUM VITAE

Name: Nele Põldver
Date of Birth: 21.10.1982
Citizenship: Estonian
University address: Institute of Psychology, University of Tartu, Näituse 2, Tartu 50409, Estonia
E-mail: nele.poldver@ut.ee

Education
2008–2018 University of Tartu, Psychology, doctoral student
2005–2007 University of Tartu, Psychology, Master’s degree (MSc)
2001–2005 University of Tartu, Psychology, Bachelor’s degree (BSc)

Employment
2017–… University of Tartu, Institute of Psychology, project manager
2014–2016 University of Tartu, Institute of Psychology, research fellow (experimental psychology)
2007–2013 University of Tartu, Institute of Psychology, laboratory assistant
2005–2007 University of Tartu, Department of Psychology, specialist in academic affairs

Fields of research
Experimental psychology, cognitive and affective neuroscience: visual perception (incl. motion perception), auditory perception (incl. language processing), mismatch negativity, emotion processing

Membership in professional organizations
European Society for Cognitive and Affective Neuroscience, Union of Estonian Psychologists, Doctoral School of Behavioural, Social and Health Sciences, Estonian Psychology Students’ Association

Publications


Nimi: Nele Põldver  
Sünniaeg: 21.10.1982  
Kodakondsus: Eesti  
Aadress: Psühholoogia instituut, Tartu Ülikool, Nääituse 2, Tartu 50409, Estonia  
E-post: nele.poldver@ut.ee

**Hariduskäik**

2008 – 2018 Tartu Ülikool, psühholoogia eriala, doktorantuur  
2005 – 2007 Tartu Ülikool, psühholoogia eriala, magistrikraad (MSc)  
2001 – 2005 Tartu Ülikool, psühholoogia eriala, bakalaureusekraad (BSc)

**Teenistuskäik**

2017 – … Tartu Ülikool, psühholoogia instituut, projektijuht  
2014 – 2016 Tartu Ülikool, psühholoogia instituut, eksperimentaalpsühholoogia teadur  
2007 – 2013 Tartu Ülikool, psühholoogia instituut (end. osakond), eksperimentaalpsühholoogia laborant  
2005 – 2007 Tartu Ülikool, psühholoogia osakond, õppekorralduse spetsialist

**Uurimistöö põhisuunad**

Eksperimentaalpsühholoogia, kognitiivne ja afektiivne neuroteadus: nägemistaju (sh liikumistaju), kuulmistaju (sh keeletöötlus), lahknevusnegatiivsus, emotsioonide töötlus.

**Kuulumine erialaorganisatsioonidesse**

Euroopa Kognitiivse ja Afektiivse Neuroteaduse Selts, Eesti Psühholoogide Liit, Eesti Käitumis-, Sotsiaal- ja Terviseteaduste Doktorikool, Eesti Psühholoogiaüliõpilaste Ühendus


32. **Margus Kanarik.** Inter-individual differences in vulnerability to depression: regional brain energy metabolism, serotonergic function and behaviour in animal models. Tartu, 2011, 239 p.


37. **Karin Täht.** The cross-cultural view on students’ motivation to learn. Tartu, 2012, 137 p.


44. **Kelli Lehto.** Depression- and anxiety-related gene variants: effects on personality traits and health-related behaviour. Tartu, 2015, 104 p.


46. **Renate Rutiku.** Refining the methodology for investigating the neural correlates of consciousness. Tartu, 2015, 166 p.

47. **Uku Vainik.** Towards a comprehensive framework for the psychological mechanisms of obesity and overeating. Tartu, 2015, 152 p.


