Aus dem Institut für Medizinische Psychologie der

Ludwig-Maximilians-Universität München

Vorstand: Prof. Martha Merrow, PhD


# Experiments on the dynamics of attention: Perception of visual rhythm and the time course of inhibition of return in the visual field


Dissertation

zum Erwerb des Doktorgrades der Humanbiologie

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München


vorgelegt von

Hui Li

aus

Anhui, China

2014

Mit Genehmigung der Medizinischen Fakultät

der Universität München

Berichterstatter:   Prof. Dr. Ernst Pöppel

Mitberichterstatter:   Prof. Dr. Benedikt Grothe

Prof. Dr. Stephan Kröger

Prof. Dr. Marianne Dieterich

Mitbetreuung durch den promovierten Mitarbeiter:

Dekan:   Prof. Dr. med. Dr. h.c. M. Reiser, FACR, FRCR

Tag der mündlichen Prüfung:   22. 07. 2014

# Eidesstattliche Versicherung

## Li, Hui

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

## Experiments on the dynamics of attention: Perception of visual rhythm and the time course of inhibition of return in the visual field

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

## Muenchen July 23, 2014

Ort, Datum

## Hui Li

Unterschrift Doktorandin/Doktorand

# Contents

# Abstract

How attention is controlled is one of the challenging topics in cognitive neuroscience and psychology. For spatially represented targets in the visual field it has been shown that some features of visual stimuli like different colors instantaneously 'pop-out', while others require a serial search which is conceived of as an effortful task. It is an open question whether dynamic feature of a stimulus are processed instantaneously without high attentional demand or serially with high demand. This question was studied in experiments on rhythm perception with periodically moving stimuli, and a visual search paradigm was employed. The search display consisted of vertically moving dots with regular rhythms; one dot however moved with a different period, and this dot with a longer or shorter period had to be detected as fast as possible. To make the period of the movement a critical target, amplitudes and phases of the distractors were randomized. It was observed that the perception of a visual rhythm defined only by the period does not lead to a pop-out effect. Apparently, the conjunction of period, equal phase and equal amplitude of movements are necessary for an effortless processing of visual rhythms. Interestingly, a faster rhythm compared to the distractors was detected with shorter reaction times. In additional experiments, it was for instance shown that auditory information supports the extraction of rhythmic visual targets indicating an intermodal mechanism.

In another experimental set-up it was tested whether the attentional machinery is controlled by a common temporal mechanism. Experiments on 'inhibition of return' (IOR) have indicated that attentional control in the peri-foveal region of the visual field underlies a different neuronal mechanism compared to the periphery of the visual field. This eccentricity effect of IOR raises the question, whether attentional control for the visual periphery is characterized by a longer time constant as the peripheral inhibitory control is

much stronger. Experimental evidence indicates, however, that the two attentional systems share the same time window of approximately three seconds. These observations support the notion of a functional subdivision of the visual field which is overcome, however, by a common temporal control mechanism.

# Zusammenfassung

Wie Aufmerksamkeit kontrolliert wird, ist eine der besonderen Herausforderungen in den kognitiven Neurowissenschaften und der Psychologie. Für räumlich repräsentierte Reize konnte gezeigt werden, dass bestimmte Aspekte visueller Reize wie verschiedene Farben sofort hervorstechen („pop-out"), während für andere Reize serielle Such-Strategien notwendig sind, die also mentalen Aufwand erfordern. Es ist eine offene Frage, ob dynamische Merkmale von Reizen ohne besonderen Aufwand verarbeitet werden, oder ob serielle Prozesse erforderlich sind, um sie zu erkennen. Diese Frage wurde in Experimenten über Rhythmus-Wahrnehmung mit periodisch sich bewegenden Reizen untersucht, und ein visuelles Such-Paradigma wurde angewandt. Es wurden auf einem Display vertikal sich bewegende Punkte gezeigt, wobei einer der Punkte sich mit einer anderen Periode, schneller oder langsamer, bewegte, und diese Punkte mussten so schnell wie möglich erkannt werden. Um nur die Periode als kritische Variable zu untersuchen, wurde die Phase und die Amplitude der anderen Reizpunkte randomisiert. Es wurde festgestellt, dass die unterschiedliche Periode allein nicht zu einem „pop-out"-Effekt führt. Damit ein abweichender, sich bewegender dynamischer Reiz erkannt wird, müssen offenbar Periode, Phase und Amplitude übereinstimmen. Reize mit einer kürzeren Periode als die Hintergrundreize wurden deutlich schneller erkannt. In weiteren Experimenten konnte beispielsweise gezeigt werden, dass akustische Information die Extraktion rhythmisch sich bewegender visueller Reize deutlich verbessert, was auf intermodale Effekte hinweist.

In einer weiteren Studie wurde untersucht, ob die neuronale Aufmerksamkeits-Maschinerie gemeinsamen zeitlichen Prinzipien gehorcht. Versuche zum Phänomen des „Inhibition of Return" (IOR, Hemmung der

Aufmerksamkeits-Wiederkehr) haben ergeben, dass die Mechanismen der Aufmerksamkeits-Steuerung im perifovealen Bereich anderen Gesetzen gehorchen als in der Peripherie des Gesichtsfeldes. Dieser „Ekzentrizitäts-Effekt" wirft die Frage auf, ob die zeitlichen Prozesse der Aufmerksamkeits-Kontrolle in der Peripherie durch längere Zeitkonstanten gekennzeichnet sind, da die inhibitorische Kontrolle dort ausgeprägter ist. Es zeigt sich allerdings, dass die beiden Aufmerksamkeits-Systeme das gleiche Zeitfenster von etwa drei Sekunden nutzen. Diese Beobachtungen stützen das Konzept der funktionellen Inhomogenität des Gesichtsfeldes, die aber durch einen gemeinsamen zeitlichen Mechanismus in eine kognitive Einheit gebracht wird.

# 1. General Introduction

## 1.1 Visual attention

### 1.1.1 What is visual attention?

Imagine the following scenes: Awakened by an alarm in the morning, you open your eyes, trying to find it and press down the button. You pick out a blue shirt from your wardrobe. Before leaving the house, you find your keys on a messy table. You sit in a cafe, waiting for a friend. Then a bunch of people walked in, you wondered whether your friend was one of them. One by one, you look at these people and check them. Suddenly, you recognize your friend, and she has probably been waving at you for a while.

In each scene described above, the overall information presented to our visual system is too rich for our brain to be completely processed at once. To deal with this excess of input, the visual system uses attention mechanisms to select relevant subsets of the scene for more extensive processing while ignoring others according to the behavioral goal at the moment. The subset could be an array of locations (e.g., the possible locations of the key on your table), a certain feature (e.g., the color of your shirt), an object (e.g., your alarm clock), or a small group of objects (e.g., people in the café) (Goldsmith, 1998).

As one of the central topic in cognitive psychology and neuroscience, visual attention has been studied for decades. The present thesis is based mostly on the literature in this domain, and is focused on the dynamic aspects of attentional processing. In the following sections, two fields with the most relevant concepts will be presented.

## 1.1.2 Studies on visual search

Visual search is a vital ability for non-human animals, since their survival depends highly on finding food and avoiding predators. Over one century ago, the British zoologist Edward Poulton wrote down his theoretical speculations of visual search in the book "The colors of animals". In his opinion, the phenomenon known as polymorphism, such like the co-existence of different colors within a species, is a benefit to the species that make it more difficult to be detected by predators (Poulton, 1890). For example, it is more difficult for a bird to search for multiple types of larvae at one time than to search for a single type.

Fifty years later, during the Second World War, theories of search were developed in secret by a group of American mathematicians, aimed to help the US Navy to locate enemy ships and to find its own lost personnel as well at sea. Those works were declassified and published by Benard Koopman in the 1950s (Koopman, 1956a, 1956b, 1957). In these articles, he theoretically developed many fundamental principles of visual search, such as the distribution of attention and the criteria for termination, which still have significant influences for contemporary theories of search. Although the early efforts from biological and operation researches have no empirical evidences, many aspects of search behavior they discussed can be found in present researches of vision, cognitive psychology and machine intelligence.

Visual search paradigm has been extensively used to study the deployment of visual attention (Treisman & Gelade, 1980; Theeuwes, 1992). In a typical experiment, a feature is defined as the search target, and the participants have to look for the target among several distractors composed of task-irrelevant features: for example, a red letter among green ones, or a horizontal bar among vertical ones. Often reaction times (RTs) increase as the number of distractors (set size) increases. However, it has been found that

search for some features is easier and more efficient than for others, in which case the function of RT against set size shows a slope near zero ms/item when the target is present. This is the typical 'pop-out' effect observed in search features such as color, size, and orientation (Treisman & Gelade, 1980; Wolfe & Horowitz, 2004), and such search process is termed "efficient search" or "parallel search".

On the other hand, visual search that produces a 'RT-to-set size' slope greater than 20-25 ms/item is often called "inefficient search" or "serial search" (Wolfe et al., 2011). The disparity in search performance between efficient and inefficient search is believed to reflect different requirements for attentional resource involved in feature processing: When a search is efficient, the target feature requires relatively little attentional resource to process, whereas, in an inefficient search the feature requires greater attentional resource (Broadbent, 1958; Treisman & Gelade, 1980; Wolfe, Cave & Franzel, 1989). According to Wolfe's studies of visual attention deployment, features such as color, size, and orientation are defined as 'guiding attributes', for they can efficiently guide attention to find the target (Wolfe & Horowitz, 2004).

Although many guiding attributes have been identified (Wolfe & Horowitz, 2004), only a few studies have investigated dynamic visual features such as visual motion (Horowitz et al., 2007; Dick & Ullman, 1987; Mcleod & Driver, 1988) or repetitive visual flicker (Spalek et al., 2009) as potential attributes. It is still unclear whether visual rhythm represented by a periodic object movement such as the movement of a bouncing ball, is a guiding attribute. This question is addressed in the first publication.

## 1.1.3 Studies on inhibition of return (IOR)

In 1984, Posner and Cohen accidentally discovered a phenomenon when they were studying the orienting of attention using a spatial cuing paradigm. In

their experiment, participants were asked to fixate on a center box which was flanked with a box on its left and right sides. An abrupt brightening of the outline of one of the peripheral boxes that was randomly selected was used as a spatial cue. A target then appeared at one of the peripheral boxes randomly after a varied time. The time between the occurrence of peripheral cue and the target was called stimulus onset asynchrony (SOA). What they found was that, the spatial cue will facilitate later response to targets appearing at the same location. However, when the SOA exceeds 300ms, the facilitation will be replaced by an inhibitory effect. They called this phenomenon "inhibition of return" (IOR), because they believed it reflects an inhibitory attentional control which can prevent participant from reorienting back to a stimulus they previously attended to.

Since the initial observation of IOR, extensive studies have been conducted to investigate its spatial characteristics. However, most of the existing studies focused on the spatial distribution of IOR around the cued location, whereas the differences between IOR in central/perifoveal and far peripheral regions of visual field are left under-investigated.

Evidences from early neurobiological and psychophysical studies suggested some functional subdivision in the human visual field. For example, the light sensitivity is higher for the central and perifoveal visual field compared to the peripheral areas which has a plateau of constant sensitivity (Pöppel and Harvey, 1973). The oculomotor system can execute only one saccade for the stimuli appeared at the perifoveal, whereas it needs two saccades when the stimuli appeared beyond 10 degree eccentricity or so (Frost and Pöppel, 1976). A study on residual vision also showed that, the patient can discriminate the location of targets within their blind area only when the targets were presented beyond about 10 degree eccentricity (Pöppel et al., 1973).

In view of the functional in-homogeneity of these two regions, Bao and Pöppel (2007) researched into the differences on attentional control as

measured with IOR between central/perifoveal and far peripheral visual fields. Results showed that the IOR effect at far peripheral regions is much stronger when compared to central and perifoveal regions, indicating a functional dissociation of attentional control between these two visual fields. On the basis of the dynamic characteristic of attention control, we further asked whether the temporal course of IOR will show different patterns in the two regions of visual field. This question is addressed in the second publication.

## 1.2 Rhythm perception

### 1.2.1 What is rhythm?

Rhythm is a ubiquitous temporal pattern in this world. It occurs at all time scales. For example, we have biological rhythms such like breathing and heartbeat, which lasts for a lifetime. The motions of the celestial bodies, the change of the seasons, the morning and evening tides, the flow of day and night, they all rise and fall rhythmically occurring at rates much lower than our heartbeat. The flapping of bee's wings, the oscillations of a string, the vibration of our vocal chords; these are also rhythmic phenomena occurring at rates much faster than the heartbeat.

We perceive rhythmic patterns at different time scales in different ways. The repetitions occurring at comparatively slower rates are conceived via long term memory. The repetitions occurring at rates near heartbeat are perceived directly as rhythm, and often induce a beat with which we can tap along. At the fastest rates, the repetitions blur together into a steady perception like "pitch" in the auditory modality or a "flicker fusion" in the visual modality. The present work will focus on rhythm perceptions under the second circumstance mentioned above. Before embarking, a further elaboration of definitions and research review will be presented in the following.

## 1.2.2 Studies on visual rhythm perception

Rhythm is typically associated with auditory stimuli such as music or speech, but it can also refer to regular temporal patterns exhibited in any modality (Pöppel, 2009). For example, a visual rhythm may consist of repeated flashes or periodic object movements; the latter seems much closer to the essence of the concept, as the original Greek word of rhythm (rhythmos) means "any regular recurring motion". In our everyday life, periodic movements, such as the swinging of a pendulum or the bouncing of a ball, can induce a rhythmic percept in the observer, which assembles the 'beat' in music. Moreover, different types of human locomotion, such as waving, walking, running, dancing etc, often form rhythms with different periods, thus induce senses of 'beat' with different 'tempi' when observed visually. Despite the ubiquity of visual motion rhythms in our daily environment, studies of its perception are surprisingly rare in the research field of visual attention. Therefore, the deployment of attention and other mechanism we employ in a dynamic visual scene containing concurrent streams of rhythmic movement, still remains unspecified.

In cognitive psychology, rhythm perception has been assumed to be modality-specific and has mainly focused on the auditory modality. Compared to auditory rhythms, visual rhythms are perceived and discriminated more poorly, and are more difficult to be synchronized (Grahn 2012; Kosonen and Raisamo 2006; Kolers and Brewster 1985; Patel et al. 2005; Repp and Penel 2002). However, in these previous studies, the rhythmic stimuli typically used is stationary one, like light flashes. The stationary nature of stimuli does not match on a phenomenal level the manner in which natural rhythms occur in the visual world where motion cues play an important role as seen in regular periodic movements of an object.

## 1.2.3 The role of attending in rhythm perception

Much of the research in the area of rhythm perception focused on the percept of a beat. Every adult without formal music training can induce the beat from music and tap along with it. In recent studies, infants and even newborns were shown to be capable of 'feeling the beat' (Hannon & Johnson, 2005; Winkler, Haden, Ladinig, Sziller & Honing, 2009). These findings suggest that beat perception is a fundamental and maybe innate process to human. However, does the intuitive impression of effortless process of beat imply that beat perception is a pre-attentive process that needs no attention to involve in?

Amongst many theories of rhythm perception, the *dynamic attending theory* (DAT) proposed by Large and Jones (1999) is the most influential one. The theory was originally used to explain how we attend to temporally varying auditory events in a sequence. There are two essential assumptions in their model: One is the existence of self-sustaining 'internal oscillations', termed *attending rhythms*, which generate expectancies in the listener and thus enable the anticipation of future occurrences of events on the basis of an underlying periodicity. The other posits that the rhythm of an external stimulus drives the attending rhythms in the listener, such that the attending rhythms become entrained, or tuned, to the periodicity of the external rhythm. Thus, attention at the particular points in time - as entrained by the external rhythm - becomes enhanced, which facilitates further processing of the stimuli occurring at those time points. In this view, beat perception thus can be seen as regular fluctuations in attentional energy over time (Drake, Jones & Baruch, 2000; Large & Jones, 1999). The DAT theory has received supports from both behavioral and neuroimaging studies (Barnes & Jones, 2000; Chapin et al., 2010; Coull, Frith, Buechel, & Nobre, 2000; Lakatos et al., 2005; Schroeder, Lakatos, Chen, Radman, & Barczak, 2009).

However, in DAT, no interpretation was given to the relationship between beat perception and selective attention. In most of the related researches mentioned before, participants always directed their attention towards the rhythmic stimuli during task. Therefore, it is difficult to draw any conclusions regarding the question whether we can develop beat percept without attending, in other words, whether selective attention is a prerequisite for beat perception.

This question has been recently addressed by several neuroimaging studies. For example, using mismatch negativity (MMN) ERP component as an index of metrical expectations, Winkler et al. (2009) conducted a study on newborns and found that violating the beat of an auditory rhythm is detected by the brain of sleeping newborns, thus support the pre-attentive beat perception. Using the same stimuli, Ladinig et al. (2009) showed that participants with no formal music training exhibited pre-attentive perception of auditory rhythm. Additional evidence for pre-attentive beat perception was provided by Geiser et al. (2010) in a similar MMN paradigm. However, other studies concerning the same question have yielded conflicting results. Chapin et al. (2010) showed in an fMRI study that attention is necessary to recruit basal ganglia, a subcortical structure whose activation is linked to beat perception, when listening to complex rhythms. Geiser et al. (2009) also provided evidence against pre-attentive beat perception, while using the same stimuli used by Geiser et al. (2010), who showed support to pre-attentive beat perception. In sum, researches concerning the relation between beat perception and selective attention have not resulted in consistent conclusions.

## 1.2.4 Cross-modal correspondences between visual and auditory features

In everyday scenes, objects and events often produce many correlated inputs to several sensory modalities. For example, when we watch a

basketball player dribbling, the movement of the ball would enter our visual system while the 'tok tok' sound of its hitting the floor would enter our auditory system. In order to disambiguate and gain an estimation of the incoming information, the perceptual system needs to combine unisensory signals referring to the same source while keeping those signals belonging to different sources separate. The question is, how does our brain decide which signals to combine?

The multisensory perception has been a hot area of cognitive neuroscience research for several years. Researchers try to understand and model how the spatial and temporal factors modulate multisensory integration. It is now generally agreed that the temporal coincidence is a usual cue for integration, which means, the closer the stimuli in different modalities are presented in time, the more likely the integration would occur (Jones & Jarick, 2006; van Wassenhove, Grant & Poeppel, 2007). Spatial coincidence could also facilitate cross-modal integration under some conditions (Frens, Van Opstal, & Van der Willigen, 1995; Slutsky & Recanzone, 2001).

Would there be any other cues the brain rely on to determine a multisensory integration? A bunch of research on multisensory information processing showed that the brain may also rely on a feature correspondence between the inputs of different sensory systems (see Spence, 2011 for a review). The impact of cross-modal correspondences on human information processing was demonstrated by an early study of Bernstein and Edelstein (1971). In their study, visual targets were presented diagonally on either side of fixation (upper left vs. lower right for some subjects, or lower left and upper right for the rest), the participants had to discriminate the location of the visual targets as quickly as possible while a task-irrelevant auditory stimulus was presented either simultaneously with the visual stimulus or slightly (up to 45 ms) later. They found that subjects respond more rapidly when both frequency and vertical position were similar (both high, or both low) than when they were

opposite (one high, the other low), showing a cross-modal congruency effect. In the following years, more and more studies investigated cross-modal correspondences using cross-modal congruency effect as an index. For example, researchers have shown that people responded more rapidly to the size of a visual stimulus (either large or small) when the task-irrelevant sound presented on each trail is congruent in pitch (when a low-pitched tone is presented at the same time as a large target) than when the sound is incongruent (when a high tone is presented with a large target) (Marks et al., 1987). To this day, cross-modal correspondences have been demonstrated between both pitch and loudness in the auditory system and lightness and brightness in vision (Hubbard, 1996; Martino & Marks, 1999), auditory pitch and visual size (Gallace & Spence, 2006), auditory pitch and visual elevation, size, and spatial frequency (Evans & Treisman, 2010), and have been demonstrated to influence people's performance in a wide range of different paradigms.

Whether a cross-modal congruency effect can be observed between auditory rhythm and visual motion rhythm is an open question. No such investigation has been done so far on this problem.

## 1.3 The present work

My doctoral research is devoted to study the dynamics of attention. The first part of my work investigated the perception of a temporal feature - rhythm. The second part of my work examined the time course of IOR in different regions of the visual field.

### 1.3.1 Part I: Perception of visual rhythm

This part of my work aimed to answer three questions about the perception of visual rhythm. Four experiments were designed using a visual search paradigm to address these questions: 1) Does visual rhythm perception of periodically moving stimuli need attention? (Experiment 1); 2) Is there a cross-modal congruency effect between auditory and visual motion rhythms? (Experiment 2 and 3); 3) Is there a cross-modal adaptation effect between auditory and visual motion rhythms? (Experiment 4).
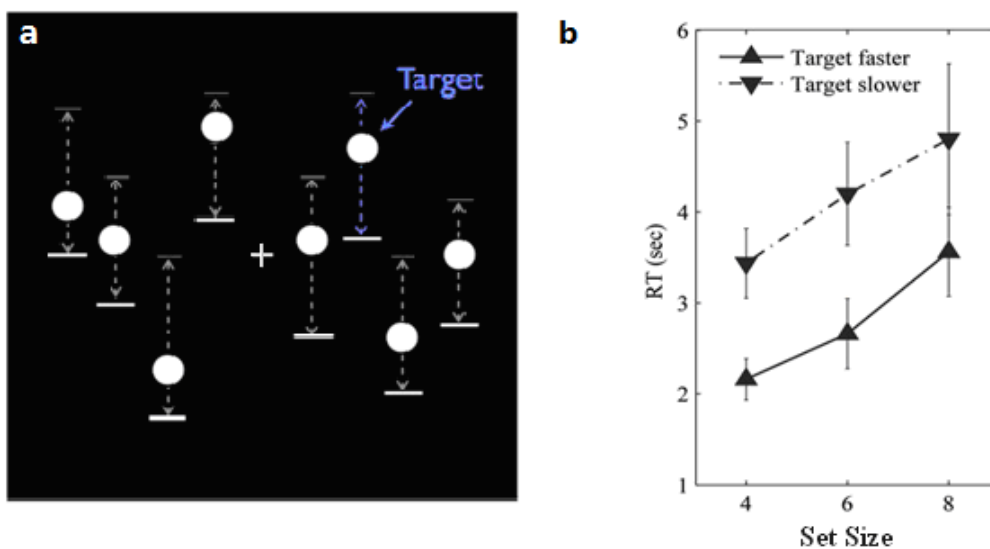
The first experiment has been published on Cognitive Processing (Li, Bao, Poeppel & Su, 2013). The second and the third experiment are in preparation to be published as a separate paper, and so is the fourth experiment. In the following parts, brief introductions to the four experiments are given respectively.

### 1.3.1.1 Experiment 1: Does visual rhythm perception of periodically moving stimuli need attention?

This experiment aimed to find out whether a unique visual rhythm of moving stimuli might capture our attention automatically, thus showing efficient or parallel search. Intuitively, one might expect that a moving object with a unique visual rhythm should "pop out", just like an "odd ball" with a different visual rhythm in a synchronous activity is immediately and effortlessly spotted. However, as visual rhythm is a temporal phenomenon that requires time to develop, it may require a considerable amount of attentional resources to be processed. Thus, a serial search seems also possible. Such uncertainty requires clarification on an experimental level.

In order to test the hypothesis, I used a visual search paradigm in which vertically "bouncing dots" with regular rhythms were presented simultaneously

on a display (Fig.1a). The search target dot was defined by a unique visual rhythm among rhythmic distractor dots. All the distractor dots bounced with the same period (550 ms), while the target bounced either faster (period = 400 ms) or slower (700 ms) than the distractors. Each trial started with the fixation cross on the center of the screen. After 1000 ms, either 4, 6, or 8 dots appeared on the screen and they were bouncing periodically and simultaneously. All the dots started bouncing at randomized phases, and their amplitudes were randomized within the predefined range, thus leaving the bouncing period being the only distinctive feature for target search. In each trial participants were required to search for the target and indicate whether the target was faster or slower than the distractors by pressing two predefined keys as fast and accurately as possible.



*Fig.1 Demonstration of search display and results.*

Eight participants (3 male, ages 23-30 years, Mean = 26.38, SD = 2.39) participated in this experiment. The experiment followed a design of 3 (set size: 4, 6 or 8) × 2 (target type: faster or slower tempo) as within-subjects variables. The median RTs in each experimental condition for each participant were

analyzed using a 3 (set size) × 2 (target type) repeated-measures ANOVA. The main effect of target type was found significant, $F(1, 7) = 7.29$, $P < 0.05$, $\eta_p^2 = 0.51$, revealing a longer RT for slower target (4.14 s) than for faster target (2.80 s). Moreover, there was a significant main effect of set size, $F(2, 14) = 15.49$, $P < 0.001$, $\eta_p^2 = 0.69$, indicating that RTs increased with the number of distractors, and the average search slope measured 346 ms/item (Fig.1b). No significant interaction was found between set size and target type $F(2, 14) = 0.23$, $P = 0.80$, $\eta_p^2 = 0.03$. The same ANOVA as for RT was also carried out for error rate, but no significant effect was found. The overall mean error rate was 12.08%.

The results of this experiment showed a shorter search RT for faster tempo targets compared to slower tempo targets. This rhythm search asymmetry may reflect a higher saliency for moving objects with faster tempi which is possibly shaped by evolutionary preference with high-pass speed filters, since movements with higher frequency may have survival advantages as they may signal a predator, a prey, of danger in the environment.

With respect to the main question of interest, the current experiment showed for the first time that the reaction time needed for perceiving a unique visual movement rhythm increased as the number of distractors increased. This RT pattern indicates that the visual rhythm is not processed in a parallel manner, but it is processed in a serial mode. Contrary to other features such as color, size, orientation and motion, the visual rhythm of a moving object turns out to be a non-guiding feature since a unique visual rhythm does not 'pop-out' or capture attention automatically. The large search slope (346 ms/item) of the linear function suggests that considerable attentional resource is required during the search of a movement rhythm. One explanation to the result is that visual rhythm perception might be related to attention entrainment, similar to the mechanism in the processing of rhythmic auditory input (Large & Jones, 1999; Jones et al., 2002). In order to perceive the visual rhythm of a moving

object, attention might need to be entrained to the period of the rhythmic movement, which is possibly triggered by the hitting the bottom line of the bouncing balls in space as critical phase transitions. Thus, it is impossible to process all the rhythmic moving objects at the same time, and a serial search process has to be initiated.


**1.3.1.2 Experiment 2: Whether a cross-modal congruency effect would occur between auditory and visual motion rhythms?**
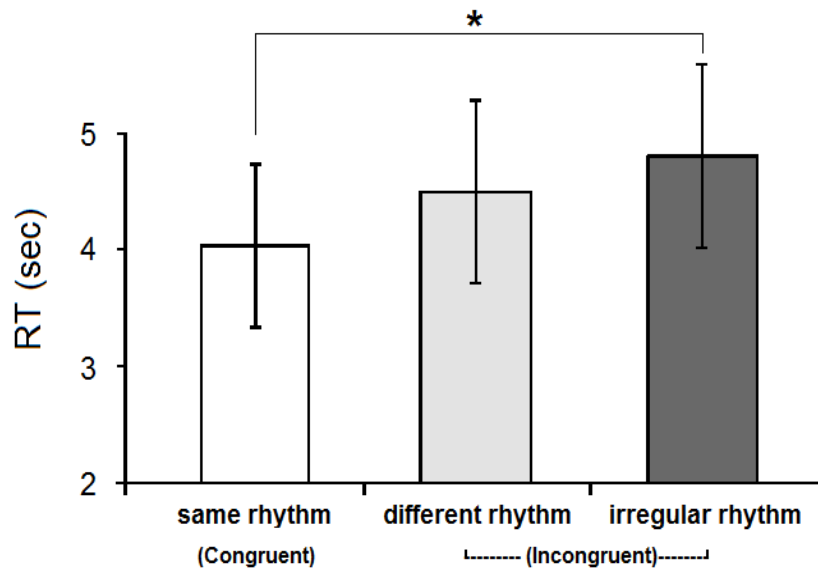
Studies comparing cognitive processing of visual rhythm and auditory rhythm often show different perceptual performances. However, these results could be due to the preferred use of stationary flashing stimuli by many researchers. In order to match a more natural manner of rhythm in the visual world, visual motion rhythm was used in the present study. It is hypothesized that, the perception of visual motion rhythm and auditory rhythm may have a common or overlapping processing mechanism. Therefore, visual motion rhythm and auditory rhythm with the same temporal characteristic should show a cross-modal congruency effect in perceptual tasks. Experiment 2 examined this hypothesis in a visual search task.

Three kinds of auditory conditions were employed in this experiment: 1) isochronous tone sequences with the same tempo as the visual target movement ('same rhythm'); 2) isochronous tone sequences with a different tempo as the visual target movement ('different rhythm'); 3) an irregular tone sequence ('irregular rhythm'). The 'same rhythm' tone sequence was always presented in anti-phase to the visual bounce. The search efficiency was represented either by RT or the slope of the function of RT against set size. A shorter RT or a smaller slope means a higher efficiency. If there were cross-modal correspondence between auditory rhythm and visual motion

rhythm, then a cross-modal congruency effect might be expected, resulting in higher efficiency for visual searches accompanied by 'same rhythm' sequences (congruent) when compared to visual search accompanied by 'different rhythm' or 'irregular rhythm' (incongruent).

The same dynamic search display as in the first experiment was used, with the set-size been reduced to 4 and 8. In each trial, a tone sequence was presented concurrently with the visual search display, which could be one of the three aforementioned conditions. Participants were instructed to ignore the tones and to perform the same visual search task as described in the first experiment.

Thirteen participants (6 male, age 21 - 32, Mean = 25.31, SD = 2.95) participated in this experiment. The experiment followed a design of 2 (set size) × 2 (target type) × 3 (auditory condition) as within-subject factors. The mean median RTs in each experimental condition were analyzed using a 2 (set size) × 2 (target type) × 3 (auditory condition) repeated-measures ANOVA. A main effect of target type was found, $F(1, 12) = 11.49$, $P < 0.01$, $\eta_p^2 = 0.49$, revealing a longer mean median RT for slower target (5.15 s) than for faster target (3.71 s). A significant main effect of set size was also found, $F(1, 12) = 13.33$, $P < 0.01$, $\eta_p^2 = 0.53$, showing that RTs increased with the number of distractors. Most importantly, there was a significant main effect of auditory condition, $F(2, 24) = 6.44$, $P < 0.01$, $\eta_p^2 = 0.35$. Post-hoc paired t-tests showed that the RT in the 'same rhythm' condition (4.02 s) was significantly shorter than RT in the 'irregular' condition (4.80 s). However, the RTs between the 'same rhythm' and the 'different rhythm' condition, as well as between the 'different rhythm' and the 'irregular' condition, were not significantly different, $Ps > 0.5$. No significant interaction was found among the three experimental factors, $P > 0.6$. The same ANOVA as for RT was conducted for error rate, but no significant effect was found. The overall mean error rate was 9.42%.

*Fig.2 Results of experiment 2.*

Besides confirming the serial search pattern found in the first experiment, the results of this experiment showed that, visual search was faster in the congruent condition where the tone sequence share the tempo that matched the tempo of the visual rhythm, as compared to the incongruent condition where the tone sequence did not have a perceivable underlying rhythm. However, although the result showed a tendency of a longer RT for the 'different rhythm' condition compared to the 'same rhythm' condition, the difference between them was not significant. To confirm this non-significance was not due to the lack of sensitivity of the experimental settings and task, Experiment 3 with less task difficulty and less parameters in the display was designed to further address the question.

### 1.3.1.3 Experiment 3: Further test to the cross-modal congruency effect.

The third experiment was designed to further test the hypothesis that the cross-modal correspondence between auditory and visual rhythm would have
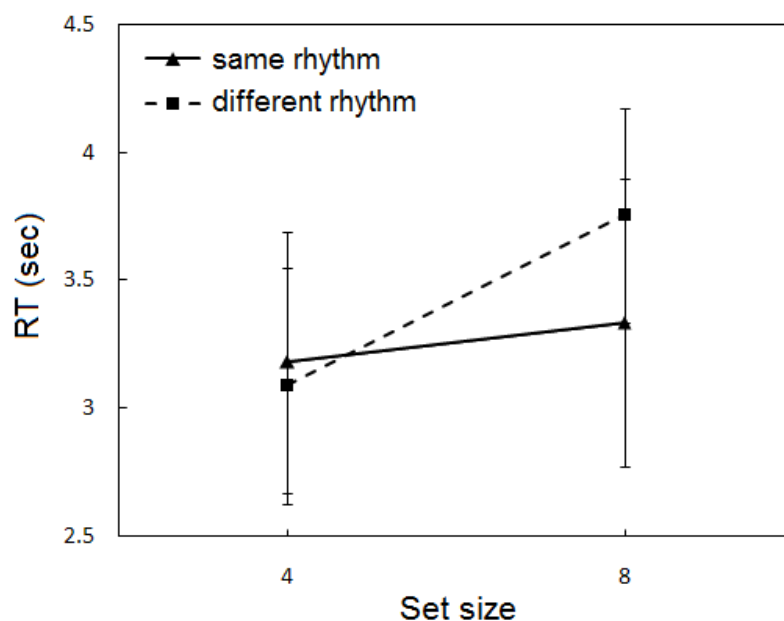
an impact on the search efficiency for visual rhythms. Two kinds of auditory conditions were included in this experiment: 1) isochronous tone sequences with the same tempo as the visual targets movement ('same rhythm') and 2) isochronous tone sequences with the same tempo as the distractors movement ('different rhythm'). The 'same rhythm' tone sequence was always presented in anti-phase to the visual bounce. The search efficiency was represented either by RT or the slope of the function of RT against set size.

Instead of a target discrimination task as employed in Experiment 2, a target detection task was used in this experiment to reduce task difficulty. Besides, faster and slower targets were presented in separate blocks, and participants were informed of the type of targets before they started each block. In each trial, participants were instructed to ignore the tones while detecting whether a visual target was present or not. They were required to press one of the two keys that corresponded to 'target present' and 'target absent', respectively.

Eight participants (3 male, ages 20 - 28, Mean = 24.25, SD = 3.54) participated in this experiment. The experiment followed a design of 2 (set size) × 2 (target type) × 2 (auditory condition) × 2 (target presence) as within-subject variables. The mean median RTs were analyzed using a 2 (set size) × 2 (target type) × 2 (auditory condition) × 2 (target presence) repeated-measures ANOVA. A main effect of target type was found, $F(1, 7) = 4.95$, $P < 0.05$, $\eta_p^2 = 0.41$, revealing a longer mean median RT for slower target (4.46 s) than for faster target (2.88 s). There was a significant main effect of set size, $F(1, 7) = 13.29$, $P < 0.01$, $\eta_p^2 = .66$, showing that RT for set size 4 (3.19 s) was shorter than that for set size 8 (4.14 s). More importantly, a significant three-way interaction was found amongst set size, auditory condition, and target presence, $F(1, 7) = 5.74$, $P < 0.05$, $\eta_p^2 = .45$. Simple effect analysis showed that, when the target was present under the auditory 'same rhythm' condition, there was no significant difference between RTs in set size 4 (3.2 s) and in set

size 8 (3.3 s), $P > 0.5$, showing a flat search function with a average slope measured 25 ms/item. On the other hand, when the target was present under the auditory 'different rhythm' condition, RTs in set size 4 (3.1 s) differed significantly from RTs in set size 8 (3.8 s), $P < 0.05$, showing a search function with a average slope measured 175 ms/item. No such interaction was found when the target was absent. The overall mean error rate was 8.41%. There was neither significant main effect nor interaction between any factors on error rate following the same four-way ANOVA.



*Fig.3 Results of experiment 3.*

The results of Experiment 3 further demonstrated the cross-modal congruency effect between auditory rhythm and visual motion rhythm, showing that the search efficiency under the congruent condition was significantly higher than the incongruent condition. Together, the results of experiment 2 & 3 support the hypothesis that auditory rhythm and visual motion rhythm may be processed with same or overlapping mechanisms.
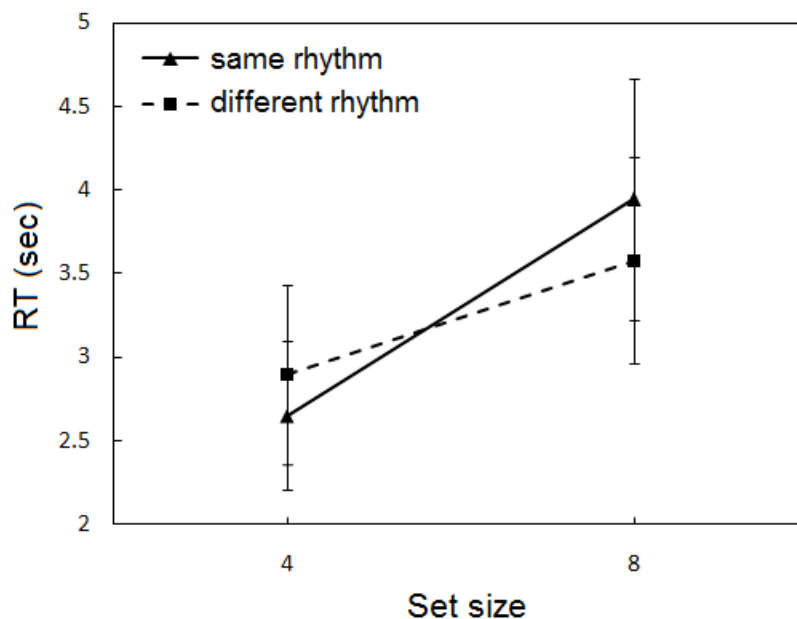
**1.3.1.4 Experiment 4**

One perceptual phenomenon that can be used to infer how information is presented in the brain is provided by analyzing aftereffect. Such effects are considered to be caused by the adaptation of neurons to sensory stimuli. It refers to the decrease of neural activities in response to a stimulus as a result of constant presentations of that stimulus. For instance, if a neuron initially fires vigorously in response to a vertical line, continued stimulation of the neuron will eventually cause a decrease in firing rate. This effect can be seen by measuring perceptual consequences. Cross-modal adaptation occurs when the adapted process in the brain induced by one modality has consequences for processes in another modality. For example, motion aftereffect can transfer between vision, audition and touch (Konkle, Wang, Hayward & Moore, 2009; Kitagawa & Ichihara, 2002; Jain, Sally & Papathomas, 2008). The cross-modal adaptation implies that in the central processing, the modalities share underlying neural representations.

In the present experiment, it is reasoned that, if there were shared neural representations in the processing of auditory and visual motion rhythm, then a long time exposure to a auditory rhythm will get the underlying neurons be adapted, thus leading to a cross-modal adaptation in the processing of visual motion rhythm. Once the rhythm aftereffect occurs, the cross-modal congruency effect observed in former experiments would disappear. Thus, an inefficient serial search would be expected for searches under 'same rhythm' auditory condition.

The procedure of the fourth experiment was similar to the third experiment except the following changes: first, i.e., to answer the core question in a concise way, only faster rhythm targets were used; second, during each block, there was a concurrent auditory rhythm which could be a 'same rhythm' tone sequence or a 'different rhythm' tone sequence, presented incessantly. The

same dynamic search displays as in the Experiment 3 were used.

Eight participants (4 male, aged 21- 29, Mean = 23, SD = 3.02) were tested in this experiment. The experiment followed a design of 2 (set size) × 2 (auditory condition) × 2 (target presence) as within-subject variables. The mean median RTs were analyzed using a 2 (set size) × 2 (auditory condition) × 2 (target presence) repeated-measures ANOVA. There was a significant main effect of set size, $F(1, 7) = 12.96$, $P < 0.01$, $\eta_p^2 = .65$, showing that RT for set size 4 (2.73 s) was shorter than that for set size 8 (3.91 s). More importantly, a significant three-way interaction was found amongst set size, auditory condition, and target presence, $F(1, 7) = 7.36$, $P < 0.05$, $\eta_p^2 = .51$. Simple effect analysis showed that, when the target was present under the auditory 'same rhythm" condition, RTs in set size 4 (2.6 s) differed significantly from RTs in set size 8 (3.9 s), $P < 0.05$. No such interaction was found when the target was absent. The overall mean error rate was 9.11%. There was neither a significant main effect nor an interaction between any factors on error rate following the same three-way ANOVA.



Fig.4 Results of experiment 4.

The results of Experiment 4 demonstrated a cross-modal adaptation between auditory rhythm and visual motion rhythm, showing that the cross-modal congruency effect observed in Experiment 3 disappeared when auditory rhythms were incessantly presented during blocks. Moreover, the search efficiency under the incongruent condition was significantly higher than the congruent condition. This may be the case because the auditory ''different rhythm'' presented in the incongruent condition used the same rhythm as visual distractors, therefore the adaptation to the rhythm of distractors increased the sensitivity to the target rhythm as a novel feature. This cross-modal adaptation further supports the hypothesis that processing for auditory rhythm and visual motion rhythm may have shared neural representations.

## 1.3.1.5 Summary of the experiments

The four experiments presented here investigated the perception of visual motion rhythm by using a visual search paradigm. In Experiment 1, participants searched for a target dot bouncing with an either 'faster' or 'slower' tempo among distracting dots bouncing with a fixed tempo. Search time was increased as the number of distractors increased, and was shorter for the 'faster' tempo target than for the 'slower' tempo target. Experiment 2 and 3 revealed a cross-modal congruency effect between auditory rhythm and visual motion rhythm. The search efficiency was higher when the trial was accompanied by an auditory tone sequence with 'same rhythm' than by a tone sequence with 'different rhythm'. Furthermore, in Experiment 4, it was tested whether a cross-modal adaptation effect would occur between auditory rhythm and visual motion rhythm. As expected, the cross-modal congruency effect disappeared when the trial was companied by long-exposed auditory rhythms.

Taken together, these experiments provide a better understanding to the attention deployment in dynamic visual scenes characterized by rhythmic movements, and they offer an empirical demonstration of cross-modal congruency effects between auditory and visual motion rhythms. The finding of a cross-modal adaptation effect further suggests shared neural substrates for processing auditory and visual motion rhythms.

## 1.3.2 Part II: Temporal window of IOR in different visual fields

In view of early evidences for a functional subdivision of the human visual field, Bao and Pöppel checked the effect of stimulus eccentricity on IOR. (Bao and Pöppel, 2007). By systematically manipulating the eccentricities of cues and targets, they found that IOR was much stronger in the far periphery relative to the perifoveal visual field, indicating a functional dissociation of attention control in the visual field. In order to understand the temporal dynamics underlying this spatial dissociation, the present study was conceived to further examine the temporal window of IOR at the two functional regions of the visual field (see the second publication). In this study, I took the responsibility of experiment design, programming, data collecting and analysis.

Two stimulus eccentricities (7 & 21 degree) were used in the experiment. The cue-target SOAs were manipulated systematically from 500 ms to 4500 ms with 500 ms increments to capture the offsets of IOR in both regions of the visual field. Twenty-five students (13 males) participated in the experiment (aged from 18 to 24, mean age = 21, SD = 1.77). The mean RTs for correct test trials were analyzed using a 2 (stimulus eccentricity) × 5 (SOAs) × 2 (target location) repeated-measures ANOVA. The results again confirmed the eccentricity effect of IOR as reported in early study. Moreover, as to the temporal dynamics of this effect, analysis on ratio IOR data and a further trend

analysis by linear regression suggested a common temporal window for the perifoveal and peripheral IOR with different attenuation characteristics.

Overall, the study confirmed a stronger IOR for more peripheral stumuli, and indicated a same temporal window for IOR in both regions of visual field.

*In sum*, the present thesis explored the perception of visual motion rhythm and the temporal dynamics of IOR in different regions of the visual field. The experiments on rhythm yield new insights into the underlying mechanism of rhythm perception, which may have implications for clinical rehabilitation for Parkinson's patients (e.g. Azulay et al., 1999; Azulay, Mesure & Blin, 2006; Jiang & Norman, 2006; McAuley, Daly & Curtis, 2009). The study on IOR revealed a spatial subdivision and a temporal unifying mechanism of attentional control in perifoveal and peripheral visual fields.

## 2. Published Scientific Works

## 2.1 A unique visual rhythm does not pop out.

Hui Li, Yan Bao, Ernst Pöppel, Yi-Huang Su

Author contributions:

Li: Conception and design of the study, data collection and analysis, writing.

Bao: Contribution to discussion and writing.

Pöppel: Discussion and final acceptance of the paper.

Su: Programming, contribution to design, discussion and writing.

SHORT REPORT

# A unique visual rhythm does not pop out

Hui Li · Yan Bao · Ernst Pöppel · Yi-Huang Su

**Abstract**  We investigated attentional demands in visual rhythm perception of periodically moving stimuli using a visual search paradigm. A dynamic search display consisted of vertically "bouncing dots" with regular rhythms. The search target was defined by a unique visual rhythm (i.e., a shorter or longer period) among rhythmic distractors with identical periods. We found that search efficiency for a faster or a slower periodically moving target *decreased* as the number of distractors *increased*, although searching for a faster target was about one second faster than searching for a slower target. We conclude that perception of a visual rhythm defined by a unique period is not a "pop-out" process, but a serial one that demands considerable attention.

**Keywords**  Attention · Visual search · Rhythm perception · Temporal processing

H. Li · E. Pöppel · Y.-H. Su
Institute of Medical Psychology and Human Science Center, Ludwig Maximilian University, Munich, Germany

Y. Bao (✉)
Department of Psychology and Key Laboratory of Machine Perception (MoE), Peking University, 5 Yiheyuan Road, Beijing 100871, China
e-mail: baoyan@pku.edu.cn

Y. Bao · E. Pöppel
Parmenides Center for Art and Science, Pullach, Germany

E. Pöppel
Institute of Psychology, Chinese Academy of Sciences, Beijing, China

Y.-H. Su
Department of Movement Science, Technical University Munich, Munich, Germany

## Introduction

Rhythm is typically associated with auditory modality as experienced in music or poetry (Pöppel 1989; Turner and Pöppel 1988), but it also refers to regular temporal patterns exhibited in vision, touch, proprioception, or movement (Feldmann 1955; Fröhlich 1920; Pöppel 2009; Su and Pöppel 2012), in sensorimotor synchronization (Franek et al. 1991; Mates et al. 1994) and even in counting (Rubia et al. 1997). Although vision seems dominant in processing information in humans, as exemplified for instance in attentional control in the visual field (Bao and Pöppel 2007; Bao et al. 2013), this is not the case in rhythm perception where auditory processing plays a crucial role. Infants and even newborns are capable to perceive auditory rhythm or "feel the beat", while the ability to perceive visual rhythms is developed later in life (Bahrick and Lickliter 2004; Brandon and Saffran 2011; Winkler et al. 2009); compared to auditory rhythms, visual rhythms are perceived and discriminated more poorly (Grahn 2012; Kosonen and Raisamo 2006). Multimodal studies investigating synchronized tapping with auditory and visual stimuli also reveal that visual rhythms are more difficult to be synchronized as compared to auditory rhythms (Kolers and Brewster 1985; Patel et al. 2005; Repp and Penel 2002). Thus, the rhythm perception is assumed to be modality-specific, i.e., showing an auditory dominance. As a consequence, research on rhythm perception has mainly focused on the auditory modality.

Previous studies investigating visual rhythms typically use stationary rhythmic stimuli like light flashes to produce rhythmic visual percepts. This stationary nature of stimuli does not match on a phenomenal level the manner in which natural rhythms occur in the visual world where motion cues play an important role as seen in regular periodic

movements of an object. A recent study has revealed that although infants of 7-month old are not able to discriminate stationary visual rhythms, they are capable to discriminate visual rhythms that occur with an apparent motion component (Brandon and Saffran 2011). This observation suggests the importance of using motion-related dynamic scenes when studying visual rhythms.

As a matter of fact, in the real visual world, we are confronted with many instances of rhythmic information that is patterned in time. Periodic movements such as synchronous walking of a group, violinists playing together in an orchestra, rhythmic movement of branches in a tree when the wind is blowing, the swinging of a pendulum, or the bouncing of a ball, all induce the perception of visual rhythms. Given the ubiquity of rhythms in the visual world, it is important to know how rhythms are extracted from the visual environment and how they are perceived. Although visual rhythm is a temporal phenomenon that involves time to develop, will a unique visual rhythm in a dynamic scene pop out, just like a red letter among green ones capturing our attention automatically and effortlessly? Does visual rhythm operate as a guiding feature similar to other visual features such as color or orientation in the visual search domain? To the best of our knowledge, no study thus far has addressed this question.

The visual search paradigm has been extensively adopted to study the deployment of attention (e.g., Treisman and Gelade 1980; Pashler 1987; van Zoest et al. 2006), and it provides a useful tool to answer the question raised above. In a typical visual search experiment, subjects are asked to search for a target that differs from the distractors by a unique visual feature such as a red letter among green ones or a horizontal bar among vertical ones. Such targets defined by a unique visual feature tend to "pop out" immediately, showing a rather flat (a slope <10 ms/item) reaction time (RT) function against set size (total number of items including both distractors and target in the display), indicating an efficient search or parallel processing of all items simultaneously. Visual search that leads to a linear RT function with a slope greater than 20 ms/item is interpreted as inefficient or serial search, suggesting an effortful or sequential processing of items in a search display (Wolfe 1998). The difference in search performance between efficient and inefficient search is believed to reflect different requirements for attentional resource involved in feature processing: When a search is efficient, the target feature requires relatively little attentional resource, whereas in an inefficient search, the target feature requires greater attentional resource to be processed (Treisman and Gelade 1980; Wolfe et al. 1989). Using visual search paradigms, two types of features have been distinguished, i.e., one is guiding feature as observed in parallel or efficient search, the other is non-guiding feature as

revealed in a serial or inefficient search (Wolfe and Horowitz 2004; Wolfe et al. 2011). Features such as color, size, orientation, and short-range motion have been shown to be guiding features eliciting a "pop out" or parallel search (Cavanagh et al. 1990; Dick et al. 1987; McLeod et al. 1988; Treisman and Gormican 1988; Treisman and Souther 1985; Wolfe et al. 1999), whereas features like optic flow, intersections, and faces require a serial search (Braddick and Holliday 1991; Nothdurft 1993; Tong and Nakayama 1999; Wolfe and DiMase 2003).

Although a great deal of information has been accumulated, no study thus far has focused on temporal features like visual rhythm in the visual search domain. Therefore, it is still unknown whether a visual rhythm is extracted with a parallel or serial process. Considering the dynamic nature of our environment and the ubiquity of rhythms in our visual world, we aimed to find out whether a unique visual rhythm of moving stimuli might capture our attention automatically, thus showing an efficient or parallel search. Intuitively, one might expect that a moving object with a unique visual rhythm should "pop out", just like an "odd ball" with a different visual rhythm in a synchronous activity is immediately and effortlessly spotted. However, as visual rhythm is a temporal phenomenon that requires time to develop, it may require a considerable amount of attentional resources to be processed. Thus, a serial search seems also possible. Such uncertainty requires clarification on an experimental level.
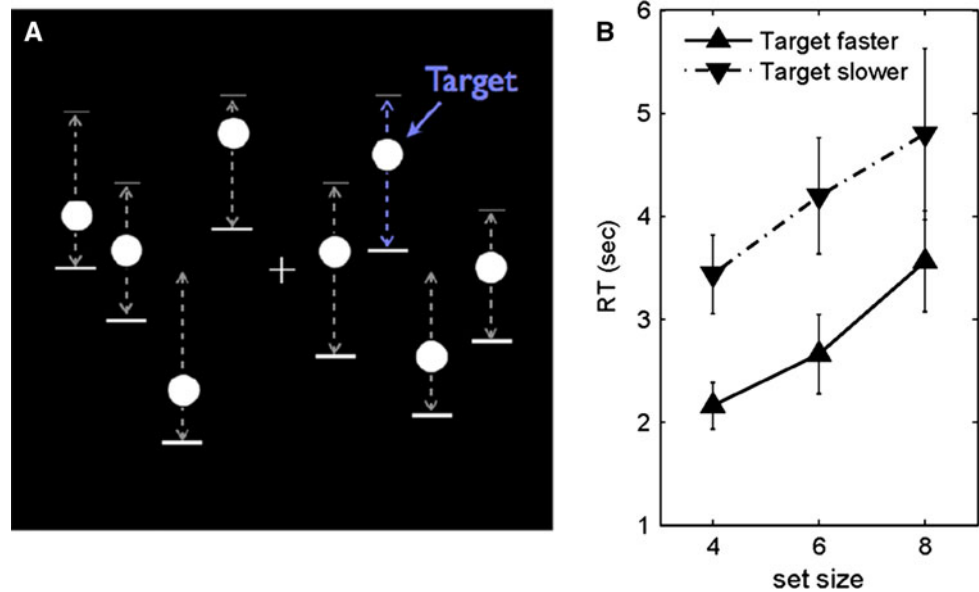
## Method

### Participants

Eight participants (3 male, age range = 23–30 years, mean age = 26 years, SD = 2.39 years) participated in the present study. All reported having normal or corrected-to-normal vision. Informed consent was given before the experiment. Each participant received a small honorarium for participating in the experiment.

### Materials and procedure

The visual search stimuli were generated by MATLAB (2009a) (Mathworks) using Psychophysics Toolbox extensions version 3 (Brainard 1997) and were presented on a 21 inch CRT monitor with a display resolution of $1{,}280 \times 800$ pixels. The search display was a square field subtending $10° \times 10°$ around the fixation (a white cross centering at the screen) and consisted of 4, 6, or 8 white dots against a black background. Each dot subtended 0.8° in diameter. The search display was equally divided into 4, 6, or 8 imaginary columns. The initial

**Fig. 1** Demonstration of search display and results. **a** A sample search display consisting of both the distractors and the target, which is defined by a unique rhythm. **b** Mean median RT as a function of set size for faster and slower target conditions



position and the movement trajectory of each individual dot were constrained within such predefined columns. Each dot moved independent of the others in a vertical manner with constant speed, resembling a bouncing ball, neglecting the influence of gravitation in the real world in this experimental design. For each dot, the bouncing amplitude was chosen between 6° and 8° vertical visual angle.

Each trial started with the fixation cross on the center of the screen. After 1,000 ms, either 4, 6, or 8 dots appeared on the screen and they were bouncing periodically and simultaneously (Fig. 1a). All the dots started bouncing at randomized phases (i.e., different initial positions within their trajectories), and their amplitudes were randomized within the predefined range, thus leaving the bouncing period being the only distinctive feature for target search. All the distractor dots bounced with the same period (550 ms), while the target bounced either faster (period = 400 ms) or slower (700 ms) than the distractors. Participants were asked to fix their eyes at the central cross throughout the trial; meanwhile, search for the target and indicate whether the target was faster or slower than the distractors by pressing two predefined keys as fast and accurately as possible. The experiment followed a 3 (set size 4, 6, or 8 dots) × 2 (target type: faster or slower period target) within-subjects design and consisted of two blocks of 90 trials each. All the 6 experimental conditions were randomized across blocks. A practice of 20 trials was added before the main test to ensure that subjects understood and performed the task correctly according to the instruction. A short break of 5 min between the two blocks was required to avoid fatigue. The entire experimental session lasted approximately 45 min.

## Results

Median RTs for correct trials in each experimental condition were analyzed with a 3 (set size) × 2 (target type) repeated-measure ANOVA. The main effect of target type was significant, $F(1,7) = 7.29$, $p < 0.05$, $\eta_p^2 = 0.51$, revealing a longer RT for the slower target (4.14 s) than for the faster target (2.80 s). Importantly, the main effect of set size was also significant, $F(2,14) = 15.49$, $p < 0.001$, $\eta_p^2 = 0.69$, showing an increased RT with the number of distractors. The average RTs for set size 4, 6, and 8 were 2.8 s, 3.43 s, and 4.18 s, respectively, with a search slope of 346 ms/item. No significant interaction between set size and target type was observed, $F(2,14) = 0.23$, $p = 0.80$, $\eta_p^2 = 0.03$. These results indicated serial search for both faster and slower targets (Fig. 1b). The same ANOVA as for RT was carried out for the error rates of the six experimental conditions, but no significant effects were observed. The overall error rate was 12.08 %.

## Discussion

The present study demonstrated a shorter search RT for faster rhythmic targets compared to slower rhythmic targets, showing an important search asymmetry. This result seems consistent with a previous observation on the detection of motion velocity, which shows searching for a faster target among slower distractors is more efficient than searching for a slower target among faster distractors (Ivry and Cohen 1992). However, this motion search asymmetry can be explained by a simple mathematical model (Rosenholtz 1999), since the Mahalanobis distances between

the target velocity and the mean of distractor distribution are different in the two situations, whereas our result cannot be explained with this model since the periods of all distractor bouncing balls were the same, and the periods of the faster and the slower targets equally deviated from the distractor period. The rhythm search asymmetry observed in the present study rather reflects a higher saliency for moving objects with faster tempi. This higher saliency is possibly shaped by evolutionary preference with high-pass speed filters, since movements with higher frequency may have survival advantages as they may signal a predator, a prey, or danger in the environment. Such high saliency objects with faster tempi might be noticed even if we are not explicitly looking for them.

With respect to the main question of interest, the current study showed for the first time that reaction time needed for perceiving a unique visual movement rhythm increased as the number of distractors increased. This RT pattern indicates that visual rhythm is surprisingly not processed in a parallel manner as one might expect from the everyday experience, but it is processed in a serial mode. Contrary to other features such as color, size, orientation, and motion, the visual rhythm of a moving object turns out to be a non-guiding feature since a unique visual rhythm does not "pop-out" or captures attention automatically. The large search slope (346 ms/item) of the linear function suggests that considerable attentional resource is required during the search of a movement rhythm. The reason of this serial search function might be related to the temporal nature of visual rhythm. In a natural environment, visual rhythm is almost always demonstrated in the context of motion, e.g., walking, dancing, music playing, or repetitive actions such as tool use. Since movement in space needs "time" to demonstrate, visual rhythm also requires time to develop, e.g., in the bouncing ball case, at least one period is needed for observers to "sense" the rhythm. Correspondingly, more attentional resources are engaged in this time-consuming process. Thus, it is impossible to process all the rhythmic moving objects at the same time, and a serial search process has to be initiated.

Another aspect of visual rhythm perception might be related to attention entrainment, similar to the mechanism in the processing of rhythmic auditory input (Large and Jones 1999; Jones et al. 2002). In order to perceive the visual rhythm of a moving object, attention might need to be entrained to the period of the rhythmic movement, which is possibly triggered by the maxima and minima of the bouncing balls in space as critical phase transitions. However, in case of phase coupling of several moving objects with the same amplitude like the synchronous walking of a group as observed in the natural visual world, an effortless "pop-out" of an "odd ball" still may happen. This implies that the conjunction of different parameters, i.e., period,

phase, and amplitude, presumably decreases the attentional demand involved in rhythm perception of visual movement, thus leading to an effortless processing of visual rhythm.

# References

Bahrick LE, Lickliter R (2004) Infant's perception of rhythm and tempo in unimodal and multimodal stimulation: a developmental test of the intersensory redundancy hypothesis. Cogn Affect Behav Neurosci 4:137–147

Bao Y, Pöppel E (2007) Two spatially separated attention systems in the visual field: evidence from inhibition of return. Cogn Process 8:37–44

Bao Y, Wang Z, Liang W, Wang Y, Pöppel E, Li H (2013) Inhibition of return at different eccentricities in the visual field share the same temporal window. Neurosci Lett 534:7–11

Braddick OJ, Holliday IE (1991) Serial search for targets defined by divergence or deformation of optic flow. Perception 20:345–354

Brainard DH (1997) The Psychophysics Toolbox. Spatial Vis 10:433–436

Brandon M, Saffran JR (2011) Apparent motion enhances visual rhythm discrimination in infancy. Atten Percept Psychophys 73:1016–1020

Cavanagh P, Arguin M, Treisman A (1990) Effect of surface medium on visual search for orientation and size features. J Exp Psychol Hum Percept Perform 16:479–491

Dick M, Ullman S, Sagi D (1987) Parallel and serial processes in motion detection. Science 237:400–402

Feldmann H (1955) Das Wesen des Rhythmus im Experiment an Gehörlosen und Normalsinnigen. Archiv für Psychiatrie und Zeitschrift für Neurologie 194:36–61

Franek M, Mates J, Radil T, Beck K, Pöppel E (1991) Sensorimotor synchronization: motor responses to regular auditory patterns. Percept Psychophys 49:509–516

Fröhlich FW (1920) Über die Physiologie des Zeitsinns. Zeitschrift für Sinnesphysiologie 51:153–158

Grahn JA (2012) See what I hear? Beat perception in auditory and visual rhythms. Exp Brain Res 220:51–61

Ivry R, Cohen A (1992) Asymmetry in visual search for targets defined by differences in movement speed. J Exp Psychol Hum Percept Perform 18:1045–1057

Jones MR, Moynihan H, MacKenzie N, Puente J (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. Psychol Sci 13:313–319

Kolers PA, Brewster JM (1985) Rhythms and responses. J Exp Psychol Hum Percept Perform 11:150–167

Kosonen K, Raisamo R (2006) Rhythm perception through different modalities. Proc Eurohaptics 2006:365–369

Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. Psychol Rev 106:119–159

Mates J, Müller U, Radil T, Pöppel E (1994) Temporal integration in sensorimotor synchronization. J Cogn Neurosci 6:332–340

MATLAB (2009a) The mathworks, Inc., Natick, Massachusetts, US

McLeod P, Driver J, Crisp J (1988) Visual search for a conjunction of movement and form is parallel. Nature 332:154–155

Nothdurft HC (1993) Faces and facial expressions do not pop out. Perception 22:1287–1298

Pashler H (1987) Detecting conjunctions of color and form: reassessing the serial search hypothesis. Percept Psychophys 41:191–201

Patel AD, Iversen JR, Chen Y, Repp BH (2005) The influence of metricality and modality on synchronization with a beat. Exp Brain Res 163:226–238

Pöppel E (1989) The measurement of music and the cerebral clock: a new theory. Leonardo 22:83–89

Pöppel E (2009) Pre-semantically defined temporal windows for cognitive processing. Philos Trans R Soc B 363:1887–1896

Repp BH, Penel A (2002) Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. J Exp Psychol Hum 28:1085

Rosenholtz R (1999) A simple saliency model predicts a number of motion popout phenomena. Vis Res 39:3157–3163

Rubia K, Schuri U, von Cramon DY, Pöppel E (1997) Time estimation as a neuronal network property: a lesion study. NeuroReport 8:1273–1276

Su Y, Pöppel E (2012) Body movement enhances the extraction of temporal structures in auditory sequences. Psychol Res 76:373–382

Tong F, Nakayama K (1999) Robust representations for faces: evidence from visual search. J Exp Psychol Hum Percept Perform 25:1016–1035

Treisman A, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97–136

Treisman A, Gormican S (1988) Feature analysis in early vision: evidence from search asymmetries. Psychol Rev 95:15–48

Treisman A, Souther J (1985) Search asymmetry: a diagnostic for preattentive processing of separable features. J Exp Psychol Gen 114:285–310

Turner F, Pöppel E (1988) Metered poetry, the brain, and time. In: Rentschler I, Herzberger B, Epstein D (eds) Beauty and the brain: biological aspects of aesthetics. Birkhäuser Verlag, Basel, pp 71–90

van Zoest W, Giesbrecht B, Enns JT, Kingstone A (2006) New reflections on visual search: interitem symmetry matters! Psychol Sci 17:535–542

Winkler I, Háden GP, Ladinig O, Sziller I, Honing H (2009) Newborn infants detect the beat in music. PNAS 106:2468–2471

Wolfe JM (1998) Visual search. In: Pashler H (ed) Attention. University College London Press, London, UK, pp 13–73

Wolfe JM, DiMase JS (2003) Do intersections serve as basic features in visual search? Perception 32:645–656

Wolfe JM, Horowitz TS (2004) What attributes guide the deployment of visual attention and how do they do it? Nat Rev Neurosci 5:495–501

Wolfe JM, Cave KR, Franzel SL (1989) Guided search: an alternative to the feature integration model for visual search. J Exp Psychol Hum Percept Perform 15:419–433

Wolfe JM, Klempen NL, Shulman EP (1999) Which end is up? Two representations of orientation in visual search. Vis Res 39:2075–2086

Wolfe JM, Vo ML, Evans KK, Greene MR (2011) Visual search in scenes involves selective and nonselective pathways. Trends Cogn Sci 15:77–84

## 2.2 Inhibition of return at different eccentricities in the visual field share the same temporal window.

Yan Bao, Zhiyuan Wang, Wei Liang, Yi Wang, Ernst Pöppel, Hui Li:

Author contributions:

Bao: Conception and writing.

Wang: Part of data analysis.

Liang: Part of data analysis.

Wang: Contribution to graphics.

Poeppel: Discussion.

Li: Design, programming, data collection and analysis.

# Inhibition of return at different eccentricities in the visual field share the same temporal window

Yan Bao [a,b,e,*], Zhiyuan Wang [a,c], Wei Liang [a], Yi Wang [a], Ernst Pöppel [a,d,e], Hui Li [a,d]

[a] Department of Psychology, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China
[b] Key Laboratory of Machine Perception (Ministry of Education), Peking University, 5 Yiheyuan Road, Beijing 100871, PR China
[c] School of Physics, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China
[d] Institute of Medical Psychology, Ludwig Maximilian University Munich, Goethestr. 31, 80336 Munich, Germany
[e] Human Science Center, Ludwig Maximilian University Munich, Goethestr. 31, 80336 Munich, Germany

## HIGHLIGHTS

► IOR magnitude was larger at 21° relative to 7° eccentricity in the visual field.
► IOR at both 7° and 21° eccentricities is characterized by a passive decay over time.
► IOR at both 7° and 21° eccentricities disappears at approximately the same time of 3 s.
► IOR in the visual field is controlled by a common temporal mechanism.

## ARTICLE INFO

## ABSTRACT

Neurobiological and psychophysical evidence indicates a functional subdivision of the human visual field with a border at approximately 10–15° eccentricity. Recent support for this inhomogeneity comes from an attention study on inhibition of return (IOR), which shows a much stronger IOR effect in the periphery relative to the perifoveal visual field (Bao & Pöppel [1]). Is this inhomogeneity of the visual field also reflected in the temporal dynamics of IOR? To answer this question, we examined when IOR effects disappear at the two functional regions of the visual field. Consistent with previous observations, IOR is much stronger in the periphery relative to the perifoveal visual field, but the two decay functions reach threshold at approximately the same time. This observation suggests a common temporal control window for IOR in both perifoveal and peripheral visual fields.

© 2012 Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

Visual attention can be captured by the sudden onset of a peripheral cue, leading to a biphasic processing of a subsequent target appearing at the same location. Facilitation of responding is usually observed when the target immediately follows the cue. However, when the cue-target interval becomes longer, a delayed responding to targets appearing at the cued location relative to the uncued locations will be observed. This latter effect is called "Inhibition of Return" (IOR) and has been generally interpreted as an attentional bias toward novel locations against the previously inspected ones [9,20,21]. This phenomenon has been intensely investigated with respect to its various characteristics and potentially underlying mechanisms [5–8,22,23]. However, one aspect related to the spatial distribution of IOR, i.e. whether IOR is homogeneously distributed throughout the visual field, has not been addressed until an eccentricity effect of IOR was recently described [1].

Neurobiological and psychophysical evidence indicates a functional subdivision of the human visual field with a border at approximately 10–15° eccentricity along the horizontal and vertical meridian. The spatial distribution of light-difference thresholds shows a higher sensitivity for the central and perifoveal visual field which is surrounded by a plateau of constant sensitivity in the peripheral areas [19]. In a study with patients who had suffered injuries of the central visual pathways but leaving some perifoveal and peripheral vision intact, it was found that critical flicker fusion was reduced in the perifoveal region but not beyond [18]. A functional dissociation was also observed in studies of residual vision

* Corresponding author at: Department of Psychology, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China. Tel.: +86 10 62753200; fax: +86 10 62761081.

E-mail address: baoyan@pku.edu.cn (Y. Bao).

or "blindsight" [26]. All these studies suggest that the neuronal processing modes for stimuli appearing at perifoveal and peripheral regions of the visual field are qualitatively different.

Motivated by these observations, Bao and Pöppel [1] further asked whether attentional control in the visual field might also underlie a functional dissociation, i.e., showing different processing mechanisms for stimuli appearing in the perifoveal and peripheral regions of the visual field. Since IOR can be seen as an attentional bias in sampling locations in the visual field, it provides a useful and valid measurement for evaluating attentional control in the visual field. By systematically manipulating the stimulus eccentricities of the cues and targets from 5° to 30°, an eccentricity effect of IOR, i.e., a stronger IOR in the periphery relative to the perifoveal visual field, was also demonstrated. This observation suggests that attentional control in the visual field cannot be considered as a homogenous phenomenon, but is characterized by a spatial dissociation. Is this eccentricity effect, however, really a robust phenomenon that can be consistently observed when different stimulus eccentricities from the two functional regions are compared? Will this effect possibly disappear after subjects receive extensive practice during the task? A further study addressed these questions and demonstrated that the eccentricity effect of IOR is a stable phenomenon, i.e., it can be observed when different stimulus eccentricities are compared, and the effect is resistant to subjects' practice [2]. Being convinced of the robustness of the eccentricity effect, we further asked whether the apparent spatial inhomogeneity of the visual field is also reflected in the temporal dynamics of IOR at different eccentricities, or whether attentional control in the time domain is independent of these spatial factors.

## 2. Methods

In order to examine the temporal dynamics of IOR in the two functional regions of the visual field, we selected two stimulus eccentricities (7° and 21°), and manipulated the cue-target SOAs (the time interval between the onset of the cue and the onset of the target) in a systematic way. To capture when IOR effects start to disappear, we tested a relatively longer SOA range from 500 ms to 4500 ms. We expect that such a long SOA range is sufficient to capture the offsets of IOR in both regions of the visual field.

Twenty-five students (13 males) aged from 18 to 24 years (mean age = 21.36 years, SD = 1.77 years) from Peking University participated in the experiment for payment. All of them reported normal or corrected-to-normal vision and were naïve to the purpose of the study. The stimuli were white figures on a black background, consisting of a fixation cross at the center, five outline boxes (subtending 1.5°) serving as cues, and a solid dot (0.8°) serving as target. The outline boxes were only presented during the cueing process and did not appear as place holders at other times. The target was preceded by a peripheral cue, which appeared either left or right to the fixation at the same stimulus eccentricity of the target. Participants were asked to respond to the target by pressing the space bar of the keyboard with their dominant hand.

The experiment took place in a dimly illuminated room. Subjects were seated 45 cm from the computer with their heads rested on a chin rest. The center of the screen was set at the subjects' eye level. A detection task with a typical double-cue IOR paradigm (see Fig. 1) was presented on the computer screen. Each trial started with a fixation cross at the center and remained visible throughout the trial. Following the onset of the fixation cross for 1000 ms, one of the boxes appeared randomly at either 7° or 21° eccentricities to the left or right of the fixation for 100 ms. After an interval of 70 ms, a central cue (same box) appeared at the fixation location for 100 ms. Following a varied interval of 230/1230/2230/3230/4230 ms, a target appeared randomly at either the peripherally cued location or
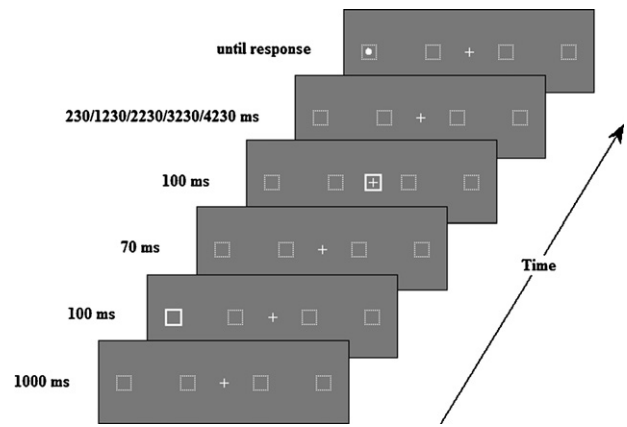


**Fig. 1.** Sample trial sequence of a typical double-cue IOR paradigm (for details see text).

the uncued opposite symmetric location with the same stimulus eccentricity as the cue. The target remained on the screen until the spacebar was pressed. An inter-trial interval of 1000 ms blank screen was inserted before the next trial started. Participants were informed that the peripheral cues did not predict where the target would occur, and they were required to keep their fixation at the cross throughout each trial and detect the targets as quickly and as accurately as possible. On catch trials where there was no target following the cues, participants were asked to withhold their responses, and catch trial ended after 2500 ms of the offset of the central cue. If participants pressed the space bar during catch trials, an error signal (500 Hz tone) was presented for 100 ms. Eye movements of the participants were not monitored in this study since previous studies have shown that subjects only make very few fixation errors [22] and that the pattern of results does not change when eye movements are monitored [e.g., 23]. However, the experimenter did check during the practice session whether the participants can fix their gaze appropriately on the central cross during each trial, and all of them seemed to be able to follow our fixation requirement very well.
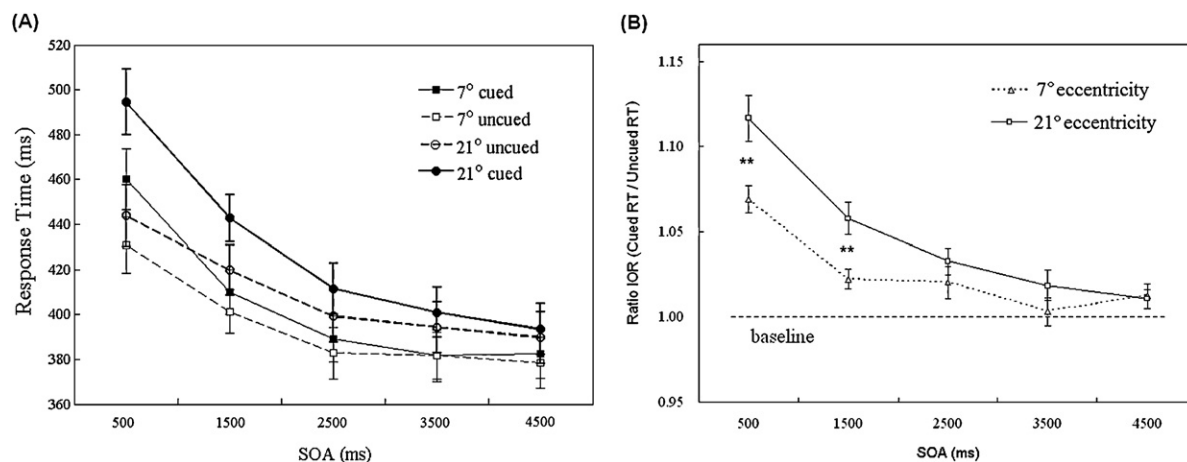
After a practice block of 48 trials, all participants started the main test, which included 600 target trials and 120 catch trials. All trials were randomized completely and arranged into 15 blocks with 48 trials in each. Trial types were balanced among two stimulus eccentricities (7°/21°), five cue-target SOAs (500/1500/2500/3500/4500 ms) and two target locations (cued location/uncued location).

## 3. Results

Only response times (RTs) for correct test trials were analyzed. The response time data for each subject were first submitted to a descriptive statistics and RTs beyond 3 standard deviations were excluded. A lower RT limit of 120 ms was further employed to exclude those RTs that are physiologically impossible. Mean RTs as a function of cue-target SOA and target location are shown for each stimulus eccentricity in Fig. 2A.

An analysis of variance (ANOVA) with stimulus *eccentricity* (7° and 21°), *SOA* (500 ms, 1500 ms, 2500 ms, 3500 ms, 4500 ms) and target *location* (cued location, uncued location) as within-subjects factors was conducted on the mean RT data. The main effect of *eccentricity* was significant [$F(1,24) = 44.009$, $P < 0.001$, $\eta_p^2 = 0.647$], revealing a slower RT for more peripheral locations which was consistent with previous observations [27]. Furthermore, the *eccentricity* effect interacted with target *location* [$F(1,24) = 23.437$, $P < 0.001$, $\eta_p^2 = 0.494$], showing a significantly larger IOR effect (mean RT for cued location trials minus mean RT for uncued

**Fig. 2.** The time course of IOR for different eccentricities as indicated by mean response times and ratio IOR values. (A) Mean response times at both cued and uncued locations as a function of SOA for both 7° and 21° eccentricities. A gradual decay of IOR was demonstrated for both 7° and 21° eccentricities (see text). (B) Ratio IOR values as a function of SOA for both 7° and 21° eccentricities. Compared to the baseline value of "1" which indicating no effect of IOR, significantly larger values were observed at the SOAs of 500 ms, 1500 ms and 2500 ms for stimuli presented at both 7° and 21° eccentricities. Significant difference of the ratio IOR values between the two stimulus eccentricities are indicated by asterisks (**$P$ < .01). Error bars indicate standard errors of mean.

*Note*: Ratio IOR = $RT_{cued}/RT_{uncued}$. ($RT_{cued}$ stands for individual RT for cued location trials; $RT_{uncued}$ stands for individual RT for uncued location trials).

location trials) at the 21° relative to the 7° eccentricity (20 ms vs. 10 ms on average). This result again demonstrated an eccentricity effect of IOR as observed previously [1,2]. The three-way interaction between *eccentricity*, *SOA* and *location* was also significant [$F(4,96) = 3.609$, $P = 0.009$, $\eta_p^2 = 0.131$]. Further analysis of this interaction revealed different temporal dynamics of IOR at 7° and 21° eccentricity conditions. For the 7° eccentricity, a significant interaction between *location* and *SOA* [$F(4,96) = 14.889$, $P < 0.001$, $\eta_p^2 = 0.383$] was observed, and further $t$-tests showed that IOR was presented at the first two SOAs (500 ms and 1500 ms, $Ps < 0.001$). For the 21° eccentricity, the interaction between *location* and *SOA* was also significant [$F(4,96) = 27.184$, $P < 0.001$, $\eta_p^2 = 0.531$], and further $t$-tests showed that IOR occurred at the first three SOAs (500 ms, 1500 ms and 2500 ms SOAs, $Ps < .001$). These results seemed to suggest a longer temporal window of the peripheral IOR (21° eccentricity) relative to the perifoveal IOR (7° eccentricity).
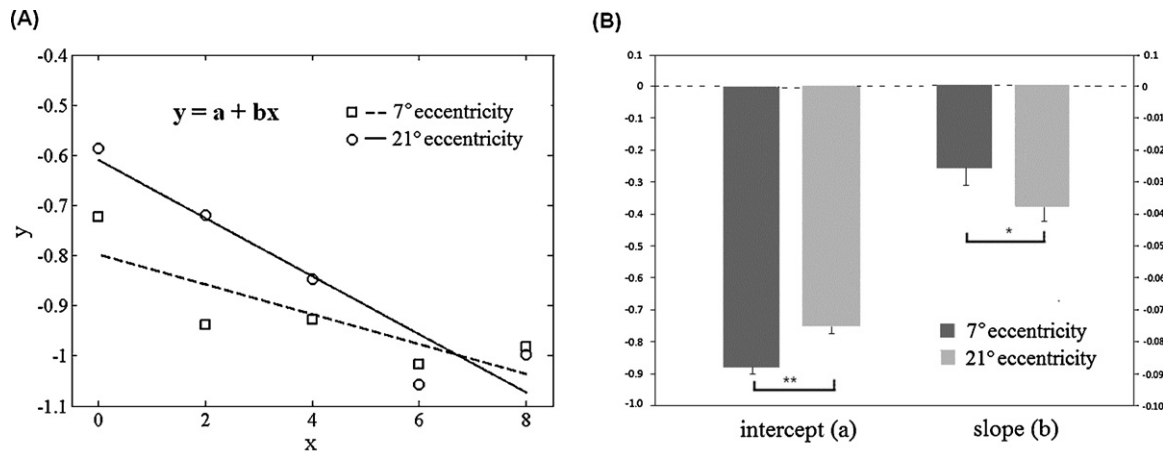
However, since the absolute RT at 21° eccentricity was generally slower than that at 7° eccentricity, which made the direct comparison using original RT data not appropriate, we further analyzed the ratio IOR for each stimulus eccentricity (Fig. 2B) with the individual RT for cued location trials divided by the RT for uncued location trials ($RT_{cued}/RT_{uncued}$) as an index of the relative IOR magnitude. A ratio value of "1" indicating no effect of IOR (i.e., same RTs were measured for both the cued and the uncued location trials) which provided a baseline for measuring IOR. This new ANOVA with *eccentricity* and *SOA* as two within-subject factors yielded significant main effects of both *eccentricity* [$F(1,24) = 19.815$, $P < 0.001$, $\eta_p^2 = 0.452$] and *SOA* [$F(4,96) = 31.991$, $P < 0.001$, $\eta_p^2 = 0.571$]. While the main effect of *SOA* indicated a gradual decay of IOR, the main effect of *eccentricity* revealed again a larger IOR effect for 21° vs. 7° eccentricity (1.05 vs. 1.03). Of our interest, a significant two-way interaction [$F(4,96) = 3.184$, $P = 0.017$, $\eta_p^2 = 0.117$] was again observed. Further analyses showed that IOR (significantly larger ratio value than baseline value "1") was observed at the first three SOAs (500 ms, 1500 ms and 2500 ms) for both 7° and 21° eccentricity stimuli ($Ps < .05$), but only at the first two SOAs (500 ms and 1500 ms) larger ratio IOR values for 21° vs. 7° eccentricity were observed ($Ps < .01$). These new results not only indicated that the eccentricity effect of IOR (larger IOR for 21° vs. 7° stimuli) finally disappeared at the SOA of 2500 ms, but also revealed a common offset of IOR at ca. 3 s (after 2500 ms and before 3500 ms) for both 7° and 21° eccentricity stimuli.

In order to quantitatively compare the different attenuation functions between the IOR effects at 7° and 21° eccentricities, we further conducted a trend analysis (for details see Note section) for ratio IOR data by linear regression ($y = a + bx$) using transformed ratio IOR values ($y$) as a function of transformed SOA ($x$). The results (Fig. 3) showed that IOR at 21° eccentricity was not only stronger (the intercept "$a$" for 7° eccentricity is −0.8755, and for 21° eccentricity is −0.7508, $t(24) = −4.668$, $P < .001$), but also attenuated faster (the slope "$b$" for 7° eccentricity is −0.0259, and for 21° eccentricity is −0.0381, $t(24) = 2.205$, $P = 0.037$). This further trend analysis complemented our original ratio IOR data and suggests a common temporal window for the perifoveal and peripheral IOR with different attenuation speed.

## 4. Discussion

The eccentricity effect of IOR as reported previously [1,2] was again confirmed in the present study, supporting the concept of a functional dissociation of attention control in the visual field [1,3,17]. The main question studied here was whether the different regions of the visual field share the same temporal window for IOR, or whether the temporal dynamics are different. As indicated in Figs. 2 and 3, although both the perifoveal (7°) and peripheral IOR (21°) followed a passive decay across time, their attenuation functions were quite different. The more peripheral IOR showed a much steeper decline of the decay function. Combining the absolute ratio IOR data which showed a non-significant difference between the two eccentricity conditions at the SOA of 2500 ms and the complete offsets of IOR for both conditions afterwards, one can conclude that IOR in the visual field share a common temporal window of approximately 3 s, matching nicely the time constant as indicated by previous research [13–16,25].

It has been shown that IOR can be observed up to a temporal interval of a few seconds [4,10,12,24]. Previous research on temporal processing has indicated that many cognitive functions appear to be embedded within a temporal window of a few seconds only [15,16]. Evidence comes for instance from experiments on temporal reproduction, sensorimotor synchronization or motor control. Sensory stimuli can be reproduced veridically and with small variance up to 2–3 s approximately, but not beyond; a regular sequence of acoustic stimuli can be synchronized under attentional control with motor taps precisely up to this temporal limit, but not beyond

**Fig. 3.** Trend analysis for ratio IOR data by linear regression using transformed ratio IOR values ($y$) as a function of transformed SOAs ($x$). (A) Linear regression line of a typical subject (No. 3) using the regression function of $y = a + bx$. Here, $y = \log_{10}(\text{Ratio} - 0.91)$ and $x = (\text{SOA} - 500)/500$, where Ratio represents ratio IOR value and SOA represents the cue–target onset asynchrony in milliseconds (see text). It is obvious that the regression line for the condition of 21° eccentricity is steeper than that for the condition of 7° eccentricity. (B) Mean values of the intercept ($a$) and the slope ($b$) of the linear regression function ($y = a + bx$) across all subjects. Significant differences between the 7° and the 21° eccentricities were observed for both the intercept and the slope (*$P < .05$, **$P < .01$). Error bars indicate standard errors of mean.

[13,14]; intentional behavioral acts show a preferential duration of 2–3 s in different cultures indicating an anthropological universal [25]. All these observations indicating a time constant of ca. 3 s was further supported by the results of the present study. The fact that prifoveal and peripheral IOR disappear at the same time but attenuate with different time constants (attenuation rates) also suggests a common temporal control system, i.e., independent of eccentricity a common baseline of reactivity is programmed to be reached after approximately the same temporal interval. Such a common mechanism would secure homogeneity of temporal processing within the visual field independent of stimulus eccentricity, and would have important behavioral consequences, as the two attention systems of the visual field would share the same time horizon.

An equivalent mechanism with respect to homogeneity of the visual field has been demonstrated for brightness: although sensitivity is falling off toward the periphery under photopic adaptation conditions as measured with light-difference threshold, apparent brightness throughout the visual field is constant, thus, creating the homogeneous visual field [19]. Thus, although we can observe (at least) two attention systems in the visual field as indicated by the eccentricity effect of IOR, and although threshold sensitivity within the visual field changes as a function of eccentricity, both spatial and temporal homogeneity are created to overcome challenges resulting from neuronal constraints, providing a unitary operating platform for visual cognition.

In sum, by comparing the temporal dynamics of IOR in both 7° and 21° eccentricities, our study demonstrated on the one hand a stronger IOR magnitude for more peripheral stimuli confirming previous observations, but on the other hand indicating that the temporal window within which IOR can be observed is approximately the same. This observation not only confirms that IOR in the visual field is not homogenous, but shows a robust eccentricity effect with respect to inhibitory magnitude, but also reveals that perifoveal and peripheral regions of the visual field might be controlled by two qualitatively different inhibitory systems [11] which, however, share the same temporal control window.

**Note**

We further conducted a *trend analysis* based on the ratio IOR data to compare directly the time course of IOR for 7° and 21° eccentricities. As can be seen from the ratio IOR data (Fig. 2B), *linear function* does not fit the raw data since ratio IOR value will not decrease any

more at very long SOAs but approaches a baseline. Therefore, an *exponential decay* can best describe the time course of IOR:

$$\text{Ratio} = C \cdot 10^{-k \cdot \text{SOA}} + 1$$

where, $C$ and $k$ are coefficients that define the strength and declining rate of ratio IOR quantity. The constant 1 corresponds to the general assumption that ratio IOR ultimately attenuates to the baseline 1 ($\text{RT}_{\text{cued}}/\text{RT}_{\text{uncued}} = 1$). Then, we derive a linear dependence on SOA by applying logarithm of 10:

$$\log_{10}(\text{Ratio} - 1) = -k \cdot \text{SOA} + \log_{10} C$$

Due to the fact that individual data of ratio IOR have a minimum of 0.92, we have changed the constant baseline 1 to 0.91 so as to make the logarithm applicable. Thus, the above equation is modified to:

$$\log_{10}(\text{Ratio} - 0.91) = -k \cdot \text{SOA} + \log_{10} C$$

Next, two new variables are used as substitutes for the Ratio IOR value and SOA:

$$y = \log_{10}(\text{Ratio} - 0.91)$$

$$x = \frac{\text{SOA} - 500}{500}$$

Using these two new variables, a mathematical model for our linear regression can be derived:

$$y = a + b \cdot x$$

where the intercept

$$a = \log_{10} C - 500\,k$$

and the slope

$$b = -500\,k$$

With the transformed SOA (new variable $x$), the intercept "$a$" (when $x = 0$, $y = a$) of the linear regression is now meaningful since $x$ can be 0, and the original SOA cannot be 0 (since no IOR effect occurs at a SOA of 0 ms). Using the transformed two new variables ($x$ and $y$) and the above mathematical model of linear regression ($y = a + b \cdot x$), we first calculated the regression function for each subject, and then compared the means of both the intercept ($a$) and the slope

(*b*) between 7° and 21° eccentricity conditions by using paired-sample *t*-tests since repeated measurement was involved in the present study. The results are illustrated in Fig. 3.

## Acknowledgements

## References

[1] Y. Bao, E. Pöppel, Two spatially separated attention systems in the visual field: evidence from inhibition of return, Cognitive Processing 8 (2007) 37–44.
[2] Y. Bao, T. Sander, L. Trahms, E. Pöppel, Q. Lei, B. Zhou, The eccentricity effect of inhibition of return is resistant to practice, Neuroscience Letters 500 (2011) 47–51.
[3] Y. Bao, Y. Wang, E. Pöppel, Spatial orienting in the visual field: a unified perceptual space? Cognitive Processing 13 (Suppl. 1) (2012) 93–96.
[4] Y. Bao, J. Zhou, L. Fu, Aging and the time course of inhibition of return in a static environment, Acta Neurobiologiae Experimentalis 64 (2004) 403–414.
[5] P.J. Bennett, J. Pratt, The spatial distribution of inhibition of return, Psychological Science 12 (2001) 76–80.
[6] G. Berlucchi, Inhibition of return: a phenomenon in search of a mechanism and a better name, Cognitive Neuropsychology 23 (2006) 1065–1074.
[7] L.J. Fuentes, M. Boucart, R. Alvarez, A.B. Vivas, M.A. Zimmerman, Inhibitory processing in visuospatial attention in healthy adults and schizophrenic patients, Schizophrenia Research 40 (1999) 75–80.
[8] A. Kingstone, J. Pratt, Inhibition of return is composed of attentional and oculomotor processes, Perception & Psychophysics 61 (1999) 1046–1054.
[9] R. Klein, Inhibition of return, Trends in Cognitive Sciences 4 (2000) 138–146.
[10] L.K. Langley, L.J. Fuentes, A.B. Vivas, A.L. Saville, Aging and temporal patterns of inhibition of return, Journal of Gerontology: Psychological Science 62B (2007) 71–77.
[11] Q. Lei, Y. Bao, B. Wang, E. Gutyrchik, fMRI correlates of inhibition of return in perifoveal and peripheral visual field, Cognitive Processing 13 (Suppl. 1) (2012) 223–227.
[12] J. Lupianez, E.G. Milan, F.J. Tornay, E. Madrid, P. Tudela, Does IOR occur in discrimination tasks? Yes, it does, but later, Perception & Psychophysics 59 (1997) 1241–1254.
[13] J. Mates, U. Müller, T. Radil, E. Pöppel, Temporal integration in sensorimotor synchronization, Journal of Cognitive Neuroscience 6 (1994) 332–340.
[14] Y. Miyake, Y. Onishi, E. Pöppel, Two types of anticipation in synchronization tapping, Acta Neurobiologiae Experimentalis 64 (2004) 415–426.
[15] E. Pöppel, A hierarchical model of temporal perception, Trends in Cognitive Sciences 1 (1997) 56–61.
[16] E. Pöppel, Pre-semantically defined temporal windows for cognitive processing, Philosophical Transactions of the Royal Society B 363 (2009) 1887–1896.
[17] E. Pöppel, Y. Bao, The visual field paradox: a theoretical account on the reafference principle providing a common frame for the homogeneity and inhomogeneity of visual representation, Cognitive Processing 13 (Suppl. 1) (2012) 285–287.
[18] E. Pöppel, D. von Cramon, H. Backmund, Eccentricity-specific dissociation of visual functions in patients with lesions of the central visual pathways, Nature 256 (1975) 489–490.
[19] E. Pöppel, L.O. Harvey Jr., Light-difference threshold and subjective brightness in the periphery of the visual field, Psychological Forschung 34 (1973) 145–161.
[20] M.I. Posner, Y. Cohen, Components of visual orienting, in: H. Bouma, D.G. Bouwhuis (Eds.), Attention and Performance. X. Control of Language Processes, Erlbaum, Hove, UK, 1984, pp. 531–556.
[21] M.I. Posner, R.D. Rafal, L.S. Choate, J. Vaughan, Inhibition of return: neural basis and function, Cognitive Neuropsychology 2 (1985) 211–228.
[22] J. Pratt, R.A. Abrams, Inhibition of return in discrimination tasks, Journal of Experimental Psychology: Human Perception and Performance 25 (1999) 229–242.
[23] J. Pratt, T. Spalek, E. Bradshaw, The time to detect targets at inhibited and non-inhibited locations: preliminary evidence for attentional momentum, Journal of Experimental Psychology: Human Perception and Performance 25 (1999) 730–746.
[24] A.G. Samuel, D. Kat, Inhibition of return: a graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties, Psychonomic Bulletin & Review 10 (2003) 897–906.
[25] M. Schleidt, I. Eibl-Eibesfeldt, E. Pöppel, A universal constant in temporal segmentation of human short-term behavior, Naturwissenschaften 74 (1987) 289–290.
[26] P. Stoerig, E. Pöppel, Eccentricity-dependent residual target-detection in visual field defects, Experimental Brain Research 64 (1986) 469–475.
[27] B. Zhou, Y. Bao, T. Sander, T. Trahms, E. Pöppel, Dissociation of summation and peak latencies in visual processing: an MEG study on stimulus eccentricity, Neuroscience Letters 483 (2010) 101–104.

# 3. References

Azulay, J. P., Mesure, S., Amblard, B., Blin, O., Sangla, I., Pouget, J. (1999). Visual control of locomotion in Parkinson's disease. *Brain*, 122(Pt 1): 111-20.

Azulay, J. P., Mesure, S., & Blin, O. (2006). Influence of visual cues on gait in Parkinson's disease: contribution to attention or sensory dependence? *Journal of the Neurological Sciences*, 248(1-2): 192-5.

Bao, Y., & Pöppel, E. (2007). Two spatially separated attention systems in the visual field: Evidence from inhibition of return. *Cognitive Processing*, 8: 37-44.

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, 41, 254-311.

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, 87, 241–247.

Broadbent, D. E. (1958). Perception and communication. London: Pergamon Press.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision* 10:433–436.

Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S. J., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: the role of attending. *Frontiers in Psychology*, 1, 224.

Coull, J.T., Frith, C.D., Büchel, C., & Nobre, A.C. (2000). Orienting attention in time: behavioural and neuroanatomical distinction between exogenous and endogenous shifts. *Neuropsychologia* 38: 808-819.

Dick, M., Ullman, S., & Sagi, D. (1987). Parallel and serial processes in motion detection. *Science*, 237, 400-402.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, Referent Period, Focal

Attending.  *Cognition,* 77(3), pp. 251-288.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1), 6:1–12.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine audio-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, 57, 802–816.

Frost, D., & Pöppel, E. (1976). Different programming modes of human saccadic eye movements as a function of stimulus eccentricity: indications of a functional subdivision of the visual field. *Biological Cybernetics* 23, 39–48.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, 68, 1191–1203.

Geiser, E., Ziegler, E., Jäncke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, 45, 93-102.

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *The European Journal of Neuroscience*, 32,1979-85.

Goldsmith, M. (1998). What's in a location? Comparing object-based and space-based models of feature integration in visual search. *Journal of Experimental Psychology: General*, 127(2): 189-219.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, 220:51–61.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: implications for musical structure learning. *Cognitive Psychology*, 50, 354-77.

Horowitz, T. S., Wolfe, J. M., DiMase, J. S., & Klieger, S. B. (2007). Visual search for type of motion is based on simple motion primitives. *Perception*,

36, 1624-1634.

Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American Journal of Psychology*, 109, 219–238.

Jain, A., Sally, S. L., & Papathomas, T. V. (2008). Audiovisual short term influences and aftereffects in motion: examination across three sets of directional pairings. *Journal of Vision* 8:1-13.

Jiang, Y. & Norman, K. E. (2006). Effects of visual and auditory cues on gait initiation in people with Parkinson's disease. *Clinical Rehabilitation*, 20(1): 36-45.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*,13, 313-319.

Jones, J. A., & Jarick, M. (2006). Multisensory integration of speech signals: The relationship between space and time. *Experimental Brain Research*, 174, 588–594.

Kitagawa, N., & Ichihara, W. (2002). Hearing visual motion in depth. *Nature* 416:172-4.

Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance*, 11:150–167.

Konkle, T., Wang, Q., Hayward, V., & Moore, C. I. (2009). Motion aftereffects transfer between touch and vision. *Current Biology*. 19: 745-750.

Koopman, B. O. (1956a). The theory of search I: Kinematic bases. *Operations Research* 4: 324-346.

Koopman, B. O. (1956b). The theory of search II: Target detection. *Operations Research* 4: 503-531.

Koopman, B. O. (1957). The theory of search III: The optimum distribution of searching effort. *Operations Research* 5: 613-626.

Kosonen, K., Raisamo, R. (2006). Rhythm perception through different

modalities. Proc Eurohaptics 2006:365–369.

Ladinig, O., Honing, H., Háden, G. P., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception*, 26, 377-386.

Lakatos, P., et al. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology* 94(3): 1904-1911.

Large, E. W. & Jones M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119-159.

Li, H., Bao, Y., Pöppel, E., & Su, Y. H. (2014). A unique rhythm does not pop out. *Cognitive Processing*, 15(1): 93-97.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 384–394.

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903–923.

McAuley, J. H., Daly, P. M., & Curtis, C. R. (2009). A Preliminary investigation of a novel design of visual cue glasses that aid gait in Parkinson's disease. *Clinical Rehabilitation*, 23(8): 687-95.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature*, 332, 154–155.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, 163:226–238.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society B*, 363,1887-1896.

Pöppel, E., & Harvey, L. O. Jr. (1973) Light-difference threshold and subjective

brightness in the periphery of the visual field. *Psychologische Forschung* 34, 45–161.

Pöppel, E., Held, R., & Frost, D. (1973). Residual visual function after brain wounds involving the central visual pathways in man. *Nature*, 243,295–296.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. InH. Bouma & D.G. Bouwhuis (Eds.), *Attention and Performance X* (pp.531-556). Hillsdale,NJ: Erlbaum.

Poulton, E.B. (1890). The Colours of Animals. New York: Appleton.

Repp, B. H., Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 28:1085.

Schroeder, C. E., Lakatos, P., Chen, C. M., Radman, T., & Barczak, A. M. (2009, October). Aligning the brain in a rhythmic world. In N. Kopell (Organizer), Rhythms of the neocortex: Where do they come from and what are they good for? Short course conducted at *the annual meeting of the Society for Neuroscience*, Chicago, IL.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuro Report*, 12, 7–10.

Spalek, T. M., Kawahara, J., & Di Lollo, V. (2009). Flicker is a primitive visual attribute in visual search. *Canadian Journal of Experimental Psychology,* 63, 319-322.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention Perception & Psychophysic,* 73 (4): 971-995.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, 51, 599-606.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory–visual speech perception. *Neuropsychologia*, 45, 598–607.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 2468-71.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419-433.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495-501.

Wolfe, J. M., Vo, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, 15, 77-84.

# Acknowledgements

# Curriculum Vitae

## Hui Li

### Personal information

date of birth     23th of April, 1982

place of birth   Anhui, China

### Education

**2009 - 2013**   PhD studies in the Institut für Medizinische Psychologie, Ludwig Maximillians University, Munich.

**2009**          M.S. in Cognitive Neuroscience, Peking University, Beijing, China.

**2005**          B.S. in Nursing, Anhui Medical University, Anhui, China.

### Publications

**Li, H.**, Bao, Y., Pöppel, E., & Su, YH. (2013). A unique rhythm does not pop out. *Cognitive Processing*, 2013, volume15: 93-97.

Bao, Y.,Wang, Z. Y., Liang, W., Wang, Y., Poeppel, E.,& **Li, H.** (2013). Inhibition of return at different eccentricities in the visual field share the same temporal window. Neuroscience Letters 534: 7-11.