Correlation between Binocular Fixation Disparity and Emotional Faces in Video Stimuli: Binocular Coordination Aftereffect

Authors: Chen Feihu, Zhao Guangping

Date: 2019-02-25T00:00:00+00:00

Abstract

To maintain the holistic nature and singularity of subjective visual perception, precise coordination and integration between the two eyes is required during information gathering, i.e., binocular coordination. However, existing literature predominantly conducts "monocular vision" research based on Hering' s law—which treats the two eyes as a single organ with completely identical patterns of action. When children with autism and control group children viewed videos of emotional faces, a Tobii eye tracker recorded oculomotor trajectories with high precision. After targeted filtering of different noise components in the high-precision data, temporal variations in binocular fixation disparity were clearly revealed. The analysis results indicated: (1) Binocular fixation disparity varied with both stimulus and time. This refutes Hering' s law and indirectly supports the Helmholtz hypothesis: the two eyes are relatively independent and cooperate with each other; (2) Under different emotional faces, the binocular fixation point separation in typically developing children exhibited consistent patterns of variation, indicating that emotional faces exert a modulatory effect on binocular fixation point separation; (3) In contrast, autistic participants exhibited significant specificity in binocular fixation disparity across all emotional faces, further revealing the specificity of emotional face perception in children with autism.

Full Text

Emotional Face Correlation of Binocular Fixation Disparity in Videos: Binocular Coordination

CHEN Feihu, ZHAO Guangping*

Minnan Normal University, Fujian 36300, China
*Corresponding author: ZHAO Guangping, 278153233@qq.com

Abstract

To maintain the integrity and uniformity of subjective visual perception, precise coordination and integration between both eyes is required during information collection—a process known as binocular coordination. However, existing literature predominantly studies "monocular vision" based on Hering' s law, which treats both eyes as a single organ with identical patterns of action. This study recorded eye movements with high-precision Tobii eye trackers while children with autism spectrum disorder (ASD) and typically developing (TD) children watched videos of emotional faces. After targeted filtering of various noise components in the high-precision data, the temporal dynamics of binocular fixation disparity became clearly visible. Analysis results demonstrated that: (1) Binocular fixation disparity changes with stimuli and time, contradicting Hering's law and indirectly supporting Helmholtz's hypothesis that the two eyes are relatively independent yet cooperative; (2) TD children exhibited consistent patterns of change in binocular fixation point distance across different emotional faces, indicating that emotional faces regulate binocular fixation spacing; and (3) ASD subjects showed significant specificity in binocular fixation disparity across all emotional faces, further revealing the unique characteristics of emotional face perception in autistic children.

Keywords: Hering' s Law; Helmholtz Hypothesis; Filtering; Emotional Face; Autism

1. Introduction

Humans possess two eyes yet perceive a single, unified world. To maintain this perceptual unity, precise and systematic control and coordination between the eyes is necessary, allowing both foveae to align approximately on the same target. This process is called binocular coordination, which emphasizes the mutual cooperation and harmonization between the eyes prior to the formation of subjective "single vision" (Blythe et al., 2006; Kirkby et al., 2008; King, 2011; Yang & Kapoula, 2003). Some studies have used the distance between fixation points of the two eyes—binocular fixation disparity—as an index to measure the magnitude of differences in visual images received by both eyes (陈飞虎 et al., 2016; 高世欢 et al., 2017; 李龙珠, 2017).

However, existing eye movement and visual perception research predominantly builds upon the concept of "monocular vision," which treats both eyes as a unified organ with completely synchronized and uniform movements, focusing both lines of sight on a single point. This research approach is also constrained by the "monocular vision"concept, as visual perception studies frequently employ the "Cyclopean Eye" (the average of left and right eye coordinates) or even data from a single eye. Meanwhile, widely used eye trackers such as EyeLink, Tobii, and IView, along with their accompanying software, only consider and provide "monocular" metrics and data processing procedures (see relevant manuals for details).

2.1 Domestic and International Research

The emphasis on "monocular vision" in research content manifests in the concentration of visual psychological and cognitive functions on higher-level processing after binocular visual information fusion, such as depth perception and binocular rivalry. Some researchers even implicitly consider binocular disparity as an inherent, unchanging physiological attribute of individuals. For instance, Wang Ling and Wang Huaiqing (2007) emphasized that the interocular distance in the head plays a decisive role in binocular coordination.

Nevertheless, some studies have long noted and clearly demonstrated that during visual processing, both eyes do not fixate on the exact same point, even for single characters or dots (Kirkby et al., 2008; Kirkby et al., 2010; Paterson et al., 2009). Furthermore, these studies discovered patterns of interocular coordination: when a fixation point is about to shift (i.e., at saccade onset), one eye acts as the "abducting eye," using its fovea to process the original fixation point before jumping ahead to search for the next potential fixation point, while the other eye, the "adducting eye," leaves the original fixation point to follow after a small time interval (Collewijn et al., 1988; Vernet & Kapoula, 2009). Consequently, binocular disparity undergoes a process of first increasing and then decreasing during the period from saccade initiation to reaching the next fixation point (Collewijn et al., 1988; Vernet & Kapoula, 2009; Yang & Kapoula, 2003). After collecting sufficient information, both eyes enter the next cycle of behavioral patterns.

2.2 The Root of "Monocular Vision" and Corresponding Research Methods: Hering' s Law and Helmholtz' s Hypothesis

Two major hypotheses address the relationship between the eyes: Hering' s hypothesis and Helmholtz' s hypothesis. This debate began in the 19th century when von Helmholtz proposed that the connection between the two eyes is not a mandatory anatomical mechanism but can be influenced and modified by will alone, making it learnable and trainable (Helmholtz, 1962). This hypothesis was later challenged by Hering' s hypothesis, which stated: "When considering the influence of eye movements on vision, both eyes can be treated as a unified organ (operated by the nervous system)," "like a person pulling both reins of a horse simultaneously" (Hering, 1977). However, because early research instruments were not sufficiently precise, most studies supported Hering's hypothesis, leading to its elevation from hypothesis to law—Hering' s Law of Equal Innervation (Howard & Rogers, 1995). Consequently, research often assumes both eyes act as a single organ with synchronized movements, i.e., "monocular vision," though this approach is limited by laboratory constraints and lacks ecological validity.

2.3 Continuous Analysis in (Semi-)Naturalistic Contexts

The prevalence of "monocular vision" research paradigms has naturally influenced and constrained binocular vision research. First, some studies artificially control binocular fixation disparity by manipulating differences between two images presented separately to each eye. This approach suffers from artificiality, passivity, and low ecological validity. Second, binocular coordination research has been confined to laboratories, focusing on micro-saccadic studies that extract individual eye movement features on small time scales (typically in milliseconds, with total duration not exceeding 1000 ms). This paradigm cannot adapt to naturalistic or semi-naturalistic contexts with large time scales and sequential influences.

In response to these challenges, the naturalness, bidirectionality, dynamic nature, and large-scale, continuous, and coherent characteristics of eye trackers become particularly prominent.

2.4 Binocular Fixation Disparity in Emotional Face Processing in Autistic Children

In the search for specific characteristics of autistic individuals, domestic research has found that children with Autism Spectrum Disorder (ASD) exhibit significant differences in binocular fixation point distance compared to TD children when viewing animated social videos (not involving depth perception) (陈飞虎 et al., 2016), a conclusion repeatedly confirmed by subsequent studies (高世欢 et al., 2017; 李龙珠, 2017). However, existing research paradigms cannot advance further investigation because they cannot answer critical questions: How does binocular fixation disparity manifest when individuals read, browse, or watch planar materials such as books, pictures, or films in natural states (without involving depth perception)? If binocular fixation disparity exists uniformly throughout continuous visual perception processing, then it is relatively invariant, representing merely a fixed physiological characteristic. Otherwise, binocular fixation disparity is variable. If variable, is the change completely random or does it occur systematically in response to specific factors? If the former, binocular fixation disparity still possesses physiological significance but is only relatively stable, not absolutely invariant. If binocular fixation disparity changes systematically in response to specific factors, then clarifying this relationship would establish its psychological significance.

Theoretically, the root cause of researchers' focus on "monocular vision" while neglecting binocular vision and coordination needs clarification. Practically, the specificity of binocular fixation disparity in ASD children requires further elucidation. Simultaneously, to address the low ecological validity of laboratory research, we need to develop long-term, continuous, and visual analysis models based on eye movement data.

Methodology

Participants: The data for this study were derived from 陈飞虎 et al. (2016). Participants included 28 children with ASD selected from a special education school in Zhangzhou, Fujian Province, and 28 typically developing (TD) children selected from a kindergarten.

Materials: Three video clips containing complete storylines were extracted from the animated cartoon "Transport Cars," which is used to promote emotion recognition in children with ASD. Each clip was approximately one minute long, featuring happiness, fear, and sadness as the primary emotions, with "neutral" face segments in each video serving as baseline comparisons.

Design: A 2 (participant group: ASD children, TD children) × 3 (task type: different emotional faces "happy, sad, and fear") experimental design was employed. Video clips were presented to participants in random order to balance positional differences.

Data Generation: Tobii eye trackers recorded the coordinates of both eyes while ASD and TD children watched the animated clips. The pixel distance between the fixation points of the two eyes was calculated as binocular fixation disparity.

Data Processing: Data were analyzed using Matlab R2014a and SPSS 20.0. For outliers in the data trend, amplitude limiting filtering was applied for removal, followed by replacement of missing values with the median of adjacent data points. As shown in [Figure 1: see original paper], the overall trends for both participant groups before and after amplitude-limiting median filtering reveal that after noise removal, data trends become clearer and more pronounced, with the data range significantly reduced from approximately 0-200 to 10-60. Due to extremely dense detailed fluctuations, the trends remain somewhat unclear.

3.2 Filtering Analysis of Raw Eye Movement Data

Before fully realizing the advantages of eye trackers in binocular fixation disparity research, one issue must be addressed: noise in eye movement data. Due to the high precision of eye movements, they are easily affected by environmental noise, instrument conditions, blinking, and head movements. The aforementioned dense detailed fluctuations originate from eye movements themselves: microtremor, drift, and jitter. These eye movement features, which serve only physiological adaptation purposes, are not essential for cognitive processing (Møller et al., 2006) and are characterized by randomness, small amplitude, and high frequency. As shown in [Figure 2: see original paper] and [Figure 3: see original paper], after applying mean filtering (a linear filtering method), data trends become more prominent and clear while preserving overall patterns, making specific information about temporal sequence differences highly visible.

Results

4.1 Analysis of Filtering Effects

陈飞虎 et al. (2016) found significant differences in binocular fixation disparity between ASD and TD children in Multidimensional Scaling (MDS) space. To verify the filtering effects in this study, we compared the twice-filtered data with raw data in MDS space. To clearly demonstrate filtering effects on a two-dimensional plane while meeting all MDS criteria, we randomly selected one of the three video segments for comparison. As shown in [Figure 4: see original paper], the two-dimensional MDS distributions for both groups at each stage are reasonable (all three MDS metrics are appropriate: Stress$<$0.25, DAF$>$0.9). The boundary between the two groups becomes increasingly neat and clear from a to c. Additionally, we extracted MDS spatial coordinates for each participant under all three conditions, calculated the multidimensional Euclidean distance of binocular fixation disparity for each participant, and performed ROC curve analysis on the discriminative ability of binocular fixation disparity Euclidean distances under each condition. Results showed AUCa=0.69, AUCb=0.75, AUCc=0.78, indicating that as filtering progressed, discrimination between the two groups became increasingly pronounced, suggesting excellent discriminatory power between ASD and TD.

4.2 Temporal Sequence Analysis of Both Groups and Three Emotional Faces

After twice filtering, the overall trends and differential details of binocular fixation disparity emerged clearly. As shown in [Figure 3: see original paper] and [Figure 5: see original paper], binocular fixation disparity continuously changes with time and materials. First, binocular fixation disparity is variable; second, this change is not uniform. Specific findings are as follows: (1) When screen content showed scenes, the binocular fixation disparity trends of ASD and TD children were comparable, with ASD children’ s disparity even smaller than TD children’ s in some segments; (2) On different emotional faces, ASD children’ s binocular fixation disparity was significantly larger than TD children’ s. During continuous presentation of the same emotional face, ASD children’ s binocular fixation disparity showed greater fluctuation; (3) As emotions shifted from positive to negative (happy-sad-fear), TD children’ s binocular fixation disparity increased, while ASD children’ s first increased then decreased during fear.

However, temporal analysis can only reveal visual differences but cannot explain whether these differences are significant. To further verify the descriptive temporal analysis conclusions, we extracted binocular fixation disparity data for both groups (TD, ASD) during corresponding emotional face periods (neutral, happy, sad, fear) from the three video segments for quantitative analysis.

Results indicated significant main effects of participant type, $F(1, 6661) = 47388.12$, $p<0.001$, $p^2=0.49$, and emotional face type, $F(3, 6661) = 367.65$, $p<0.001$, $p^2=0.17$. The interaction was also significant, $F(3, 6661) = 738.87$, $p<0.001$, $p^2=0.33$. Simple effects analysis revealed that ASD children’ s binocular fixation disparity was significantly larger than TD children’ s on neutral faces and all three emotional faces (happy, sad, fear) ($p<0.01$) (as shown in [Figure 7: see original paper]). Significant differences existed among ASD children’ s binocular fixation disparity across the three emotional faces ($p<0.01$). Additionally, as emotions shifted from positive to negative (happy-sad-fear), TD children' s binocular fixation disparity increased significantly, while ASD children' s first increased significantly then decreased significantly on fear faces. ANOVA results validated the temporal trend analysis conclusions for binocular fixation disparity under three emotional faces in both groups.

Discussion

5.1 Helmholtz' s Hypothesis and Binocular Coordination

The temporal trends in binocular fixation disparity reflect the coordination performance of both eyes during visual processing, highlighting the "coordination" aspect of binocular coordination. This temporal analysis model accommodates diverse experimental materials, greatly improving ecological validity and expanding future research directions. Helmholtz' s hypothesis has been indirectly supported by developmental and rehabilitation training studies on binocular coordination, suggesting that binocular coordination appears to be driven by visual experience and based on neuroplasticity (Coubard, 2015; King, 2011). This study clearly demonstrates the incomplete consistency between the two eyes through temporal presentation of binocular fixation disparity, showing systematic changes over time and across emotional faces. These findings oppose Hering' s law of the eyes as a "unified organ" and extreme "anatomical mechanism" views. Conversely, Helmholtz' s hypothesis (Helmholtz, 1962) posits that interocular connection is not a mandatory anatomical mechanism but can be influenced and modified by will, making it learnable and trainable. Therefore, our results indirectly support Helmholtz' s hypothesis. Additionally, the coherent analysis of animated video materials exemplifies the advantages of this approach. Finally, temporal analysis can guide researchers to discover changes, locate time periods, identify triggering events, and potentially uncover psychological meanings embedded in these changes. Kirkby et al. (2008) emphasized in their review that the psychological significance of binocular coordination requires confirmation and specific investigation.

5.2 Emotional Face Regulation of Binocular Fixation Disparity

Faces are considered the most important means of information transmission in social interaction, with emotional faces being particularly complex and subtle. Meanwhile, individuals with autism exhibit special characteristics in emotional face processing. Therefore, this study selected emotional faces as experimental materials to explore interocular relationships. The results indicate that this material satisfied the research objectives. Emotional faces regulate binocular fixation disparity, and different emotional faces have different regulatory effects on different participant types. First, there are differences in regulation magnitude—for example, TD children's binocular fixation disparity differs significantly across emotions (as shown in [Figure 5: see original paper] and [Figure 6: see original paper]). Second, there are differences in regulation direction: positive emotions cause slight decreases in TD children' s binocular fixation disparity, while negative emotions cause significant increases. Research found that as emotions shifted from positive to negative (happy-sad-fear), TD children' s binocular fixation disparity increased significantly, echoing previous studies. Gross et al. (1991) reviewed literature indicating that 4-5-year-old children can identify happy, sad, and fearful facial expressions, but with decreasing accuracy across these three emotions. Guo et al. (2012) showed that compared to happy and sad faces, primary school students exhibited the largest pupil distance and longest fixation duration when viewing fearful faces.

5.3 Specificity of Binocular Coordination in Emotional Face Processing in ASD

Wu and Xu (2012) introduced perceptual specificity in ASD. Building on previous research (陈飞虎 et al., 2016; 高世欢 et al., 2017; 李龙珠, 2017), this study further explored additional characteristics of visual perceptual specificity in ASD children—namely, more pronounced specificity in binocular coordination during emotional face processing. First, ASD children' s binocular fixation disparity was significantly larger than TD children' s across all three emotional faces, with greater variability in ASD. This aligns with the view that autistic children have difficulties processing faces and emotions. Second, as emotions shifted from positive to negative (happy-sad-fear), both TD and ASD children' s binocular fixation disparity increased significantly; however, ASD children showed a turning point on fear faces with a significant decrease. This corroborates findings from other studies that ASD children have abnormal processing difficulties with negative emotions such as fear (Pelphrey et al., 2002; 严淑琼, 2008).

Conclusion

This study demonstrates that: (1) Amplitude-limiting median filtering and mean filtering can effectively eliminate noise in binocular fixation disparity derived from raw eye movement data, making temporal trends more intuitive. (2) During visual processing, binocular fixation disparity is influenced by emotional faces and time in the temporal sequence. First, this directly opposes Hering' s law while indirectly supporting Helmholtz' s hypothesis. Second, it shows that emotional faces regulate binocular fixation disparity. (3) During emotional face processing, ASD children' s binocular fixation disparity shows specificity: it is significantly higher than TD children' s across all emotional faces. Additionally, while TD children' s binocular fixation disparity increases significantly as emotions shift from positive to negative, ASD children' s disparity first increases significantly then decreases significantly on negative emotional faces.

Tobii eye trackers collect data under natural conditions without requiring equipment, leading some studies (徐静俭, 2009; 徐娟, 2012) to conclude they have higher ecological validity and are suitable for children and special populations. However, this increases uncertainty in the distance between participants and the screen, preventing conversion of binocular fixation disparity units from pixels to degrees and limiting comparison with other studies. We hope this provides a reference for future research with adult or typically developing participants.

Acknowledgments

When receiving the acceptance notification for this paper, it coincided with the 60th anniversary of my alma mater. Though I cannot be there in person, I dedicate this work to the celebration! At the same time, this commemorates the fulfilling scientific research conducted in Laboratory 218 of the Entrepreneurship Building.

Author Statement: The authors declare no conflicts of interest.
Ethical Statement: All participants provided informed consent.

Note: Figure translations are in progress. See original paper for figures.