Journal of Visualized Experiments

A comparison of eyetracking data of social videos between children with high functioning ASD, comorbid ADHD and control

--Manuscript Draft--

Article Type:	Methods Article - JoVE Produced Video
Manuscript Number:	JoVE58694R1
Full Title:	A comparison of eyetracking data of social videos between children with high functioning ASD, comorbid ADHD and control
Keywords:	ASD, ADHD, eyetracking, area of interests, scanpath, fixation counts, fixation duration
Corresponding Author:	Kwan Lan Vicky TSANG, PhD EdUHK Hong Kong, SAR CHINA
Corresponding Author's Institution:	EdUHK
Corresponding Author E-Mail:	vtsang@eduhk.hk
Order of Authors:	Kwan Lan Vicky TSANG, PhD
	Patrick Chun Kau Chu
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	hong kong

TITLE:

Comparing Eye-tracking Data of Children with High-functioning ASD, Comorbid ADHD, and of a Control Watching Social Videos

AUTHORS AND AFFILIATIONS:

Vicky Tsang[1], Patrick Chun Kau Chu[1]

[1]Department of Special Education and Counselling, Education University of Hong Kong, Hong Kong

Corresponding Author:

Vicky Tsang (vtsang@eduhk.hk)

Email Addresses of the Co-author:

Patrick Chun Kau Chu (chunkau@eduhk.hk)

KEYWORDS:

ASD, ADHD, eye tracking, area of interests, scanpath, fixation counts, fixation duration

SUMMARY:

This is a qualitative comparative case study analysis of eye-tracking data on the first moments of social video scenes as viewed by three participants: one with autism spectrum disorder, one with comorbid attention deficit-hyperactive disorder, and one neurotypical control.

ABSTRACT:

Children with autism spectrum disorders (ASD) are known to have sensory-perceptual processing deficits that weaken their abilities to attend and perceive social stimuli in daily living contexts. Since daily social episodes consist of subtle dynamic changes in social information, any failure to attend to or process subtle human nonverbal cues, such as facial expression, postures, and gestures, might lead to inappropriate social interaction. Traditional behavioral rating scales or assessment tools based on static social scenes have limitations in capturing the moment-to-moment changes in social scenarios. An eye-tracking assessment, which can be administered in a video-based mode, is therefore preferred, to augment clinical observation. In this study, using the single-case comparison design, the eye-tracking data of three participants, a child with autism spectrum disorder (ASD), another with comorbid attention deficit-hyperactive disorder (ADHD), and a neurotypical control, are captured while they view a video of social scenarios. The eye-tracking experiment has helped answer the research question: How does social attention differ between the three participants? By predefining areas of interest (AOIs), their visual attention on relevant or irrelevant social stimuli, how fast each participant attends to the first social stimuli appearing in the videos, for how long each participant continues to attend to those stimuli within the AOIs, and the gaze shifts between multiple social stimuli appearing concurrently in the same social scene are captured, compared, and analyzed in a video-based eye-tracking experiment.

INTRODUCTION:

Persons with ASD are known to be characterized by behavioral deficits in social communication, based on conventional behavioral evidence from structured observational assessments and parent interviews. In addition, sensory processing abnormalities have been recently incorporated into the DSM-5 diagnostic criteria of ASD[1]. Social information processing involves the lower level sensory-perceptual processing and higher level social cognitive processing of social information. Sensory-perceptual processing refers to the ability to attend to social stimuli and encode them in a short-term memory bank for instant retrieval and response-planning, while social cognitive processing refers to the interpretation of social information by social reasoning and problem-solving[2-3]. As such, social information-processing deficits often lead to other psychobehavioral characteristics, such as social anxiety and inattentiveness. This can be illustrated by the high comorbid prevalence rate of ASD with attention deficit-hyperactive disorder (ADHD). The range of comorbidity for ADHD in ASD has been estimated at 30% to 80%, whereas the presence of comorbid ASD in ADHD has been estimated at 20% to 50%[4].

Two major hypotheses have been put forward to account for the deficits in social information processing—namely, enhanced perceptual functioning (EPF) and weak central coherence (WCC). EPF refers to the overattentiveness to or preoccupation with specific parts by individuals with ASD, whereas WCC refers to their weakness to derive the essence of wholes by pulling together the interelement relationships of the parts[5]. Both theoretical frameworks attest to their failure to globally configure or process the multiple stimuli concurrently presented in a confined social context[6-7]. In an earlier face emotion recognition study using static face expression photos[8], it was found that the ASD group tended to show localized processing of facial features (such as the shape of the mouth) using EPF, but seem to be weaker in configural processing, which demands pulling together the more abstract perceptual concepts as postulated by WCC, such as the spatial relationships between multiple facial components (*e.g.*, the distance between the eyebrows and the intensity of the eye gaze)[9-10].

Since daily social episodes consist of dynamic moment-to-moment subtle changes in social information, any failure to attend or engage in the sensory-perceptual processing of subtle human nonverbal cues, such as facial expression, postures, and gestures, and to make sense of the relationships of the different social stimuli might lead to inappropriate social cognitive processing. Eye-tracking experiments have been increasingly used to supplement clinical observation in social information processing studies. Eye-tracking data, in the form of scanpath patterns, visual fixation counts, and visual duration, have been major biomarkers to investigate social information processing in ASD[11-15].

In this study, we illustrate the use of the eye-tracking technique to investigate whether the two participants with ASD and with ASD-ADHD process the first moments of social video scenes differently than the neurotypical child. The eye tracker equipment captures four major indices during viewing: the number of visual fixations, the first fixation duration, the total fixation duration, and the scanpath patterns in the form of spatial arrangement and sequence of fixation points. In this way, how fast each participant attends to the audio-visual stimuli predefined by AOIs as they first appear into the social scenes, for how long they continue to

look at those AOIs, and their gaze shifts between multiple AOIs appearing concurrently in the same social scene can be captured. Any delay to fixate AOIs during the first moments (*i.e.*, 500 ms) and the trajectory of the scanpaths provide important evidence for data analysis. Representative findings from the qualitative analysis of this single-case comparative study using this paradigm are reported.

PROTOCOL:

Parental and participant consent was obtained during the recruitment process in a primary school and a children service center for ASD in Hong Kong and the study was approved by the university ethical review committee of the Education University of Hong Kong.

1. Use of a Video-based Assessment

1.1. Produce several social videos, about one minute long, that consist of daily life scenarios involving several people in a social context (**Figure 1**).

1.1.1. For **Video 1**, demonstrate the following social scenario. In a crowded cafeteria, a student spots an unoccupied seat that is simultaneously occupied by a lady who is talking on the phone and places her bag on the seat with no awareness of his request.

1.1.2. For **Video 2**, demonstrate the following social scenario. Students are playing a chess game while an unfamiliar student comes too close to watch them playing the game.

1.1.3. For **Video 3**, demonstrate the following social scenario. A boy's painting is ruined when his friend accidentally spills water from a cup on the table.

1.2. Conduct expert reviews of all the videos. Select those social scenarios that are agreed on the most by the experts as containing the actors' intention, emotions, and thoughts through their expressions and gestures.

2. Recruitment of the Participants

2.1. From the pool of participants who satisfied the research inclusion criteria, select and match participants with ASD, with ASD-ADHD, and neurotypical controls using their medical diagnostic reports and the percentile scores of Raven's Standard Progressive Matrices[16].

2.2. Convert their Raven percentile scores to five percentile ranks. Select participants who perform at ranks II or III (average) and exclude those who scored above rank I (above average) or at rank IV (below average).

3. Eye-tracking Experiment

3.1. Experimental set-up

3.1.1. On one side of the eye-tracking room, display the videos on a 23-inch color LCD monitor with a screen resolution of 1920 x 1080 pixels, using an eye tracker at a distance of approximately 60 cm from the participant. Have a research investigator operate the eye tracker from the other side of the eye-tracking room (**Figure 2**).

3.1.2. Have another research investigator sit next to the participant and instruct the participant to look at the screen of the monitor. Place the monitor in front of the child on the other side of the partition and connect to the eye tracker. The choice of eye-tracking equipment, testing environment, and the set-up procedures are previously discussed[17].

3.2. Calibration process

3.2.1. Instruct the participants to watch the calibration dots that set the viewing boundaries across the screen by capturing the eye movements using infrared corneal reflectance technology (**Video 4**). The calibration is properly done if all the green dots or lines fall within the grey circle dots.

3.2.2. Repeat the calibration if some of the green dots or lines do not fall within the grey circle dots.

3.3. Viewing of the videos

3.3.1. Instruct the participant to view the social videos one after another, and capture their eye movement data during viewing using the eye tracker.

4. Data Analysis

4.1. Define and set up the first-moment fixation within AOIs (**Video 5**).

4.1.1. Choose context-relevant targets (face, hands, targeted objects, *etc.*) in their initial 500 ms of appearance in each scene of the video as AOIs (**Figure 3**) and label the AOIs in the information box on the left panel.

4.1.2. Upon the completion of the addition and selection of the AOIs in the current frame, move the cursors in the timeline bar at the bottom panel to the next frame.

4.1.3. Adjust the location and boundary of the AOIs in each frame of the video in the presentation video software of the eye tracker manually as the target areas change in each time frame of the video due to the movement of the people or objects as the story of the social video develops.

4.1.4. Click the **Select** button on the top panel and add new AOIs to the new scene if necessary. If some existing AOIs are present for 500 ms in the current scene (the timestamp of the video

can be checked in the bottom left panel) or if they are not relevant in the new frame in the video, right-click on these AOIs to deactivate them in the new frame.

4.2. Run a statistical analysis of the eye-tracking indices. Follow the steps of statistical data processing on the eye tracker as described below (**Video 6**).

4.2.1. Choose the recordings of the children.

4.2.2. Select the Media file for analysis.

4.2.3. Select from the available videos.

4.2.4. Click **Analyze selected media**.

4.2.5. Choose the **Descriptive statistics** (*e.g.*, Sum).

4.2.6. Choose the dependent measures in **Metrics** (*e.g.*, First fixation duration, visit count).

4.2.7. Choose **Recordings** in **Rows**.

4.2.8. Select **AOI Media Summary** in **Columns**.

4.2.9. Click **Update** to analyze the eye-tracking patterns. The results of the eye-tracking pattern metrics are shown on the screen.

4.3. Create the scanpath of a scene from the eye-tracking data (**Video 7**).

4.3.1. Choose **Visualization** and **GazePlot** in the software.

4.3.2. Select the **Media** and **Recordings** in the left panel for visualization.

4.3.3. In the bottom timeline, move the lower cursor to the beginning of the target scene and move the upper cursor to the end of the target scene.

4.3.4. Make sure **Accumulate** is chosen for the **Data** field to show the accumulative scanpath.

4.3.5. Click **Export** and **Visualization image** to save the scanpath as a separate image file.

REPRESENTATIVE RESULTS:

The eye-tracking data of the three Cantonese-speaking children (with ASD, with ASD-ADHD, and a control) aged between the ages of 7 and 9 viewing three social videos using the aforementioned paradigm is presented here (**Table 1**).

The first fixation duration (per 500 ms target AOI) was longer for the neurotypical child (150 ms) than for the ASD and ASD-ADHD children (both 110 ms). The total fixation duration (per 500 ms target AOI) was shorter for the ASD-ADHD child (120 ms) than for both the neurotypical child (170 ms) and the ASD child (180 ms). The total number of fixation counts (per 500 ms target AOI) was the largest for the ASD child (4.62), second for the neurotypical child (4.09), and the shortest for the ASD-ADHD child (3.19).

A scanpath plot captures the visual scanning of multiple AOIs in a social scene. An example of the scanpaths of the three children for one 10 s episode in **Video 1** is shown in **Figure 4** and **Videos 8** - **10**.

FIGURE AND TABLE LEGENDS:

Video 1: Social scenario one. In a crowded cafeteria, a student spots an unoccupied seat which is simultaneously occupied by a lady who is talking on the phone and places her bag on the seat with no awareness of his request.

Video 2: Social scenario two. Students are playing a chess game while an unfamiliar student comes too close to watch them playing the game.

Video 3: Social scenario three. A boy's painting is ruined when his friend accidentally spills water from a cup on the table.

Video 4: Eye tracker calibration process.

Video 5: The process of selecting the target AOIs in the social videos.

Video 6: The process of conducting statistical analysis.

Video 7: The process of creating the scan path of a scene.

Video 8: Scanpaths of the control.

Video 9: Scanpath of the child with ASD.

Video 10: Scanpath of the child with ASD-ADHD.

Figure 1: An example of essential social scenes in Video 1. In the first scene, the boy is waiting to get his meal from the cafeteria staff. In the second scene, he is looking for a seat near the lady who is talking on the phone. In the third scene, he asks the lady whether he can sit on the empty chair next to her. In the last scene, the lady does not notice his request and puts a bag on the unoccupied chair. The boy is disappointed because he could not find a place to sit.

Figure 2: Eye-tracking experimental set-up. A research investigator gave instructions to the child about viewing the videos in front of the monitor on one side of the eye-tracking experiment room. The display of the videos was controlled by another investigator using another computer on the other side of the same room separated by a partition.

Figure 3: An example of the target AOIs in Video 1. The colored ovals are the AOIs (*i.e.*, face, eyes, mouth, hands, mobile phone, and the bag of the lady) that show the first moments in one of the scenes in **Video 1**.

Figure 4: Scanpaths of the control (top), the child with ASD (middle), and the child with ASD-ADHD (bottom). Taking a social scene in **Video 1** as an example, the blue dots trace the scanpaths for the neurotypical control child, the green dots for the ASD child, and the red dots for the ASD-ADHD child. The dots in the figure indicate the locations of the visual fixations. The bigger the dots are, the longer the child attend to that particular spot on the visual stimulus. The numbers in the dots represent the sequence of visual fixations within 500 ms of the video scene.

Table 1: Descriptive statistics of the eye-tracker measurements of the three children.

DISCUSSION:

The first-moment fixation duration was shorter for the ASD-ADHD and ASD children than for the neurotypical child. The total fixation duration was shorter for the ASD-ADHD child than for the neurotypical child, demonstrating a general reduction in visual attention to social stimuli. This showed that the ASD-ADHD child showed a delay in attending to the entry of social stimuli in a social scene. This delay might cause the child to skip registering important momentary social information, which may lead to the misinterpretation of social information and subsequent social cognitive processing.

The total number of fixation counts was lower for the ASD-ADHD child than for the neurotypical child, while the total number of fixation counts within localized AOIs was the highest for the ASD child. This seems to support past ASD findings under the framework of enhanced perceptual functioning (EPF), which suggests that children with ASD employ featural processing; hence, they visually attend to more details of the AOIs then neurotypical controls do.

When the results of the three children are compared, it shows that the ASD child performed the fewest scans across multiple AOIs of social stimuli. This might be explained by the difficulty experienced by the ASD child in pulling together the relationship between relevant social stimuli. This can be accounted for by the weak central coherence theory (CWW), which states that ASD shows deficits in sensory perceptual processing which demands simultaneous attending to and scanning between multiple AOIs.

For scanpath analysis, several limitations are noted. Even though the same scanpath picture is used, it actually contains different scenes within a temporal period (in this study, it was predefined as a video length of 10 seconds). Therefore, there might be spatial errors of gaze

spots on the scanpath plot that do not necessarily represent the actual locations of what the participant is focusing on the plot (**Video 7**). Investigators need to be cautious of these potential eyeballing errors during data analysis and interpretation.

Since the AOIs have to be marked manually on the eye tracker, there might be a latency of visual fixation from the markers themselves. Since the AOIs were manually plotted against the moving social stimuli, there might be slight errors in the duration of how long each AOI lasts across all AOIs. For example, for a predefined 500 ms, an AOI may have been marked for 498 ms or 510 ms. This may make the comparison of performances across different videos, in contrast to that in the same video, difficult as the performance baselines differ from one video to another. Nonetheless, this artifact will have the same impact on all three participants, and therefore, this may not create a bias for a particular type of participant.

ACKNOWLEDGMENTS:
The authors acknowledge that the wider study from which this paper is generated is financially supported by the General Research Fund under the University Grants Council of Hong Kong Special Administration Region, China (grant number: GRF 844813); and by the Research Support Scheme 2017/18 of the Department of Special Education and Counselling at the Education University of Hong Kong.

DISCLOSURES:
The authors have nothing to disclose.

REFERENCES:
1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders – 5th Edition (DSM-V)*. Washington, DC: American Psychiatric Association (2013).

2. Williams, D.L., Goldstein, G., Minshew, N.J. Neuropsychologic functioning in children with autism: further evidence of disordered complex information processing. *Child Neuropsychology*. **12**, 279-298 (2006).

3. Bons, D. *et al*. Motor, emotional, and cognitive empathy in children and adolescents with autism spectrum disorder and conduct disorder. *Journal of Abnormal Child Psychology.* **41**, 425-443 (2013).

4. Van der Meer, J.M. *et al*. Are autism spectrum disorder and attention-deficit/hyperactivity disorder different manifestations of one overarching disorder? Cognitive and symptom evidence from a clinical and population-based sample. *Journal of the American Academy of Child and Adolescent Psychiatry.* **51** (11), 1160-1172 (2012).

5. Brosnan, M.J., Scott, F.J., Fox, S.,Pye, J. Gestalt processing in autism: Failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry*. **45**(3), 459-469 (2004).

6. Behrmann, M., Thomas, C., Humphreys, K. Seeing it differently: Visual processing in autism. *Trends in Cognitive Sciences*. **10** (6), 258-264 (2006).

7. Happé, F.G.E., Frith, U. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders.* **36**, 5-25 (2006).

8. Tsang, V. Eye-tracking study on facial emotion recognition tasks in individuals with high-functioning autism spectrum disorders. *Autism*. **22** (2), 161-170 (2018).

9. Renzi, C. *et al*. Featural and Configural processing of faces are dissociated in the dorsolateral prefrontal cortex: A TMS study. *NeuroImage*. **74**, 45-51 (2013).

10. Samson, F., Mottron, L., Soulieres, I., Zeffiro, T.A. Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping*. **33** (7), 1553-1581 (2012).

11. Sasson, N.J., Turner-Brown, L.M., Holtzclaw, T.N., Lam, K.S.L., Bodfish, J.W. Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. *Autism Research.* **1**, 31-42 (2008).

12. Klin, A., Jones, W., Schultz, R., Volkmar, F., Cohen, D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*. **59**, 809-816 (2002).

13. Rutherford, M.D., Towns, A.M. Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*. **38** (7), 1371-1381 (2008).

14. Wallace, S., Coleman, M., Bailey, A. An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition and Emotion*. **22** (7), 1353-1380 (2008).

15. Byrge, L., Dubois, J., Tyszka, J.M., Adolphs, R., Kennedy, D.P. Idiosyncratic Brain Activation Patterns Are Associated with Poor Social Comprehension in Autism. *The Journal of Neuroscience*. **35**, 5837-5850 (2015).

16. Raven, J. *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No.1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales.* San Antonio, Texas: Harcourt Assessment (1981).

17. Sasson, N.J., Elison, J.T. Eye tracking young children with autism. *Journal of Visualized Experiments.* (61), e3675 (2012).

Figure 1. An example of essential social scenes in video one

Figure 2. Eyetracking experimental set up

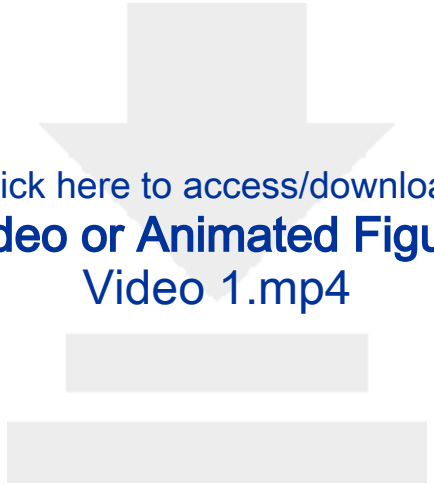
Click here to access/download;Figure;Figure 2.png

Figure 3. An example of the target AOIs in video one

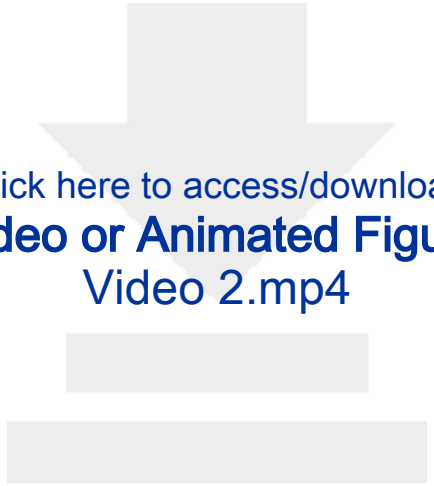
Click here to access/download;Figure;Figure 3.png

Figure 4. Scanpaths of Control, ASD and ASD-ADHD comorbid child

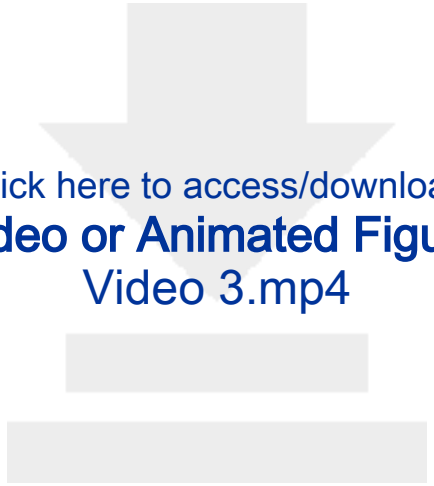
Video 1. Social scenario one



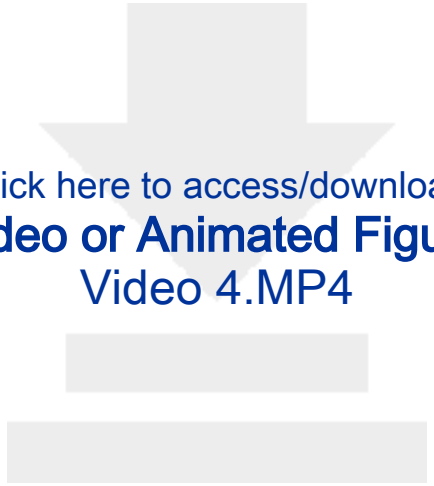
Video 2. Social scenario two



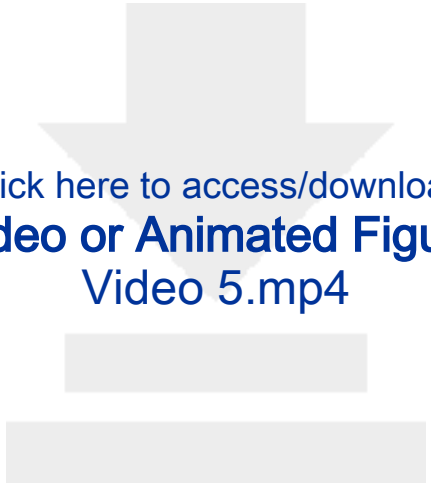
Video 3. Social scenario three



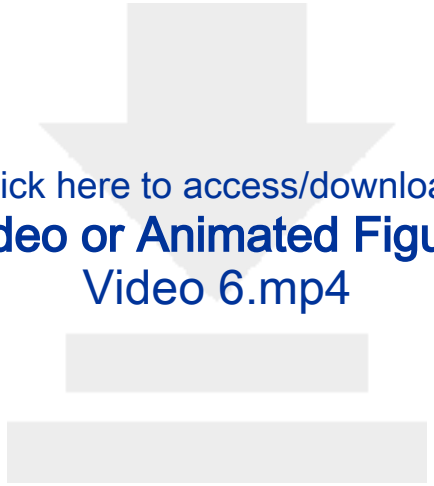
Video 4. Eyetracker calibration process



Video 5. Process of selecting the target AOIs in the social videos



Video 6. Process of conducting statistical analysis

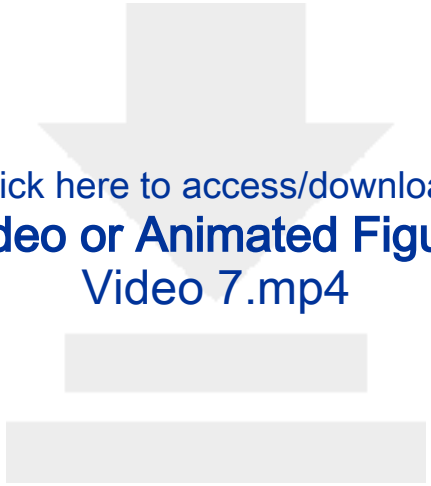


Click here to access/download

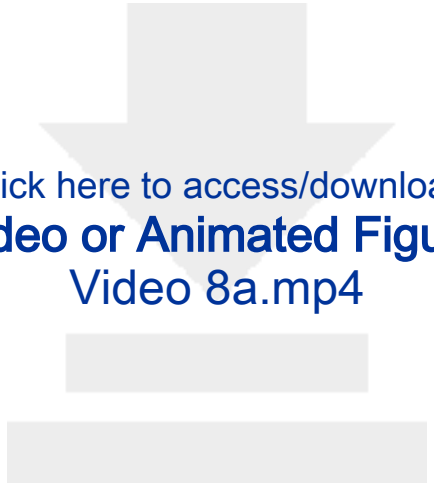
Video or Animated Figure

Video 6.mp4

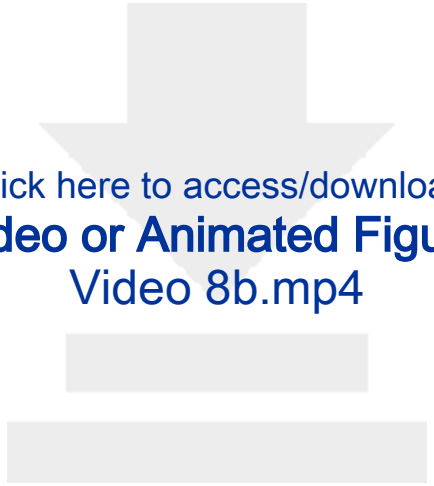
Video 7. Process of creating the scanpath of a scene



Video 8. Scanpath of the neurotypical control viewing a scene in Video 1



Video 9. Scanpath of the ASD child viewing a scene in Video 1



Video10. Scanpath of the ASD-ADHD comorbid child viewing a scene in Video 1

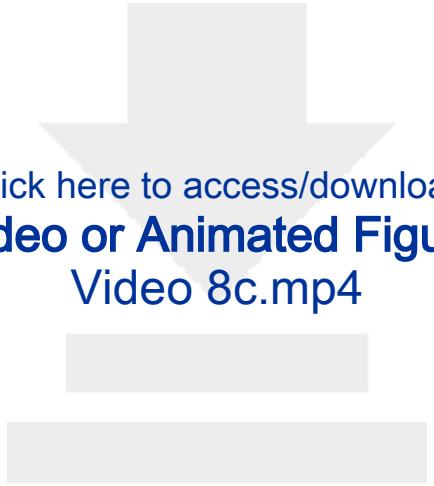


Table 1. Descriptive statistics of the eye-tracker measurements of the three childre

Participant groups	Raven Score	Grade	First fixation duration (ms)
Control	120	3	150
ASD	129	1	110
ASD-ADHD	115	3	110

n.

Total fixation duration (ms)	Fixation counts
170	4.09
180	4.62
120	3.19

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
Tobii Pro TX300	Tobii	N/A	Screen based eye-tracker (300Hz refreshing rate)
Tobii Pro Studio	Tobii	N/A	Software for analyzing eyetracking data

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: A comparison on scanpaths eyetracking data of social videos between individuals with high functioning ASD, comorbid ADHD and control

Author(s): Vicky Tsang, Patrick Chun Kau Chu

Item 1 (check one box): The Author elects to have the Materials be made available (as described at http://www.jove.com/author) via: ☑ Standard Access ☐ Open Access

Item 2 (check one box):

☑ The Author is NOT a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. Defined Terms. As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of *The Journal of Visualized Experiments;* "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. Background. The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. Grant of Rights in Article. In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. Retention of Rights in Article. Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. Grant of Rights in Video – Standard Access. This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. Grant of Rights in Video – Open Access. This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this Section 6 is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. Government Employees. If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. Likeness, Privacy, Personality. The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

9. Author Warranties. The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

10. JoVE Discretion. If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have

full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

11. Indemnification. The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in *JoVE* or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

12. Fees. To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

13. Transfer, Governing Law. This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement required per submission.

CORRESPONDING AUTHOR:

Name:	Vicky Tsang
Department:	Special Education and Counselling
Institution:	The Education University of Hong Kong
Article Title:	A comparison on scanpaths eyetracking data of social videos between individuals with high functioning ASD, comorbid ADHD and control
Signature:	Vicky Tsang
Date:	June 11 2018

Please submit a signed and dated copy of this license by one of the following three methods:

1) Upload a scanned copy of the document as a pfd on the JoVE submission site;
2) Fax the document to +1.866.381.2236;
3) Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02139

For questions, please email submissions@jove.com or call +1.617.945.9051

REBUTTAL - JoVE manuscript submission by Vicky Tsang and Patrick Chun Kau Chu

No.	**Editorial Comments**	**Authors' Responses & Actions taken**
1	Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.	The manuscript has been proofread and checked for spelling and grammatical errors.
2	Protocol Language: The JoVE protocol should be almost entirely composed of numbered short steps (2-3 related actions each) written in the imperative voice/tense (as if you are telling someone how to do the technique, i.e. "Do this", "Measure that" etc.). Any text that cannot be written in the imperative tense may be added as a brief "Note" at the end of the step (please limit notes). Please re-write your ENTIRE protocol section accordingly. Descriptive sections of the protocol can be moved to Representative Results or Discussion.	The whole section has been rewritten using the imperative tense and the descriptive sections of the protocol has been moved to Representative results.
3	Protocol Detail: Please add all specific details (e.g. button clicks for software actions, numerical values for settings, etc) to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.	The steps in 3.2 and 4.1 regarding the eyetracker calibration and the selection of AOI in the video have been expanded.
4	Protocol Numbering: There must be a one-line space between each protocol step.	A one-line space has been added between each protocol step.
5	Protocol Highlight: Please highlight ~2.5 pages or less of text in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps.	The steps that should be visualized has been highlighted in yellow.
6	Discussion: 1) Please ensure that the discussion covers the following in detail: modifications and troubleshooting, limitations of the technique, significance with respect to existing methods, future applications and critical steps within the protocol.	We have modified the content in Discussion further, please see the red parts.
7	Please remove the numbered list from the discussion.	The numbered lists have been removed from the figure.
8	Figures: All panels of a figure (e.g. a,b,c should be on one page).	Figures 4a-c has been combined into one single figure in one page.
9	Figure/Table Legends: Please expand the legends to adequately describe the figures/tables.	The legends have been expanded for Figures 1, 2, 4 and 6.
10	Please define all abbreviations at first use.	The full form has been defined for ASD, ADHD and AOI at first use in the summary, abstract and the main text.
No.	**Reviewer #1**	
11	Number of participants - should discuss this limitation	The manuscript is meant to be a case comparison report. So no group data is reported.
12	In Introduction, the authors may frame the present work as studying "social attention".	Thanks for the suggestions and information. Social attention has been used and related references added to the manuscript.
13	There are several studies along the same line of the present work using videos and the authors should be aware of these research	Thanks for the comments and information. The references such as Klin et al. (2002) and Byrge et al.

	before claiming "There have not been much research done on dynamic social videos".	(2015) have been added in the introduction and the sentence removed.
14	The authors need to carefully check the abstract — there are many grammatical errors, especially there is a mixture of tenses. The present version of abstract does not read smooth to the reviewer.	Thanks for the suggestions and comments. The manuscript has been proofread and checked for spelling and grammatical errors.
	Reviewer #2	
15	While the title is about scan paths, there seems to be more time spent discussing other types of measures … and little is said about how to conduct systematic analyses on scan path data… It is also unclear how this stands out from the article by Sasson & Elison (2012) and what methodology it is adding (perhaps detailed approaches to analyzing scan paths would have added a novel part to the current work).	Thanks for the suggestions and comments. Some details on the creation of gazepath have been added in 4.3 and the title have been changed to "A comparison of eyetracking data of social videos between children with high functioning ASD, comorbid ADHD and control" to include scanpath analysis as one of the components of eyetraking data. One major difference between the current article and Sasson & Elison (2012) is the introduction of the methodology in using social videos and the detailed delineation of its relevant target AOIs for analysis.
16	… if the authors want to go in the direction of a case study/qualitative analysis on scanpaths, the article would need to be completely re-conceived.	Most of the article has been rewritten and restructured. Please see the changes in red.
17	This article feels a bit all over the place and does not have a strong and clear story thread that the reader is following to understand what this work is contributing above and beyond past work (e.g., Sasson & Elison, who are cited by the authors), especially based on the Introduction.	The major contribution of this article is the introduction of target AOI analysis on dynamic social videos when compared with the viewing of photos which are described in most of the previous literature. This is emphasized in the revised abstract.
18	This entire manuscript needs to be revised for English language, as at least half of the sentences are difficult to interpret because of language-related issues (word choice, grammar, etc).	The manuscript has been proofread and checked for spelling and grammatical errors.
19	In the Title and Abstract, there is no mention of children, but this study is with 1st - 3rd graders… but it is confusing, as this study is clearly about children.	In the title and abstract, the word "individuals" is changed to "children" following Reviewer's suggestions.
20	In the Abstract, there is no mention of ADHD +ASD…The Introduction says nothing about ADHD (e.g., why are you including a subject with comorbidity, what is expected to differ, what has past work shown). It is also unclear the point of getting into detail about the three theories if you don't re-visit any of that or use it to make any predictions about your data.	Relevant details about the comorbid ASD-ADHD group have been added in the abstract and the introductions; and, the theories have been revisited in the introduction and discussion sections.
21	Final paragraph of Intro discusses 'comparative case study approach' but then the article takes a more traditional statistical analysis approach, despite having only 3 subjects. Again, speaks to the confusion about the authors' intention behind this article.	This article follows the single-case study approach and the regression section has been removed to make our study aim clearer to readers.

22	I think there is something about scan paths and using more qualitative approaches (combined with quantitative) that could be interesting, but there isn't much at all said to the reader about how to do scan path analyses, whether qualitatively or quantitatively, only about more traditional metrics (# fixations, first fixation duration, total fixation duration).	Details on how to do scanpath analysis have been added in 4.3.
23	In fact, in the Discussion, it mentions a set of reasons why those kinds of analyses are extremely difficult, especially when using video stimuli with complex social scenes, again causing the reader to wonder what the authors are trying to teach the reader, etc, and why they would use a method full of error, as they say.	We purposefully added the limitations in eyetracking procedures as this is specifically requested in JoVE manuscript guidelines. Therefore we reported this to let future researchers aware of its limitations should they desire to conduct similar procedures.
24	Results/Table 2 - No statistical analyses of other more quantitative aspects of eye-tracking can be properly looked at with three children across three groups. If this paper is about 3 subjects, it has to be a case study/qualitative approach, and it makes no sense to do any sort of inferential statistics when there is an n=1 for each group.	We have removed the regression table from the results and focused on qualitative analysis in our revised report.
25	Discussion 2.2 is first moment mentioning anything about the parts of the videos involving people vs. not, and if this is important or worth explaining, it should come up earlier in choices about the videos and the AOIs used, etc.	Thanks for the reviewer's suggestions. We have added a detailed section on defining and setting up the first moment fixation within AOIs in section 4.1 with a video illustration.