A Pupil-Dilation Technique to Test Developmental Differences in Visual Synchrony During Free Viewing

Tommaso Ghilardi[1,2], Ori Ossmy[1,2]

[1] Centre for Brain and Cognitive Development, Birkbeck, University of London
[2] Department of Psychological Sciences, Birkbeck, University of London

Corresponding Author: Ori Ossmy, 32 Torrington Square, London, UK, WC1E 7JL.
ori.ossmy@bbk.ac.uk

Keywords: Child development; Pupil synchrony; free viewing; Visual synchrony; Pupil dilation; Eye tracking;

Abstract

Visual synchrony, a form of coordinated behavior wherein observers look at displays in a similar manner, is important for understanding how coordinated visual attention influences cognitive, emotional, and social development. Traditional developmental research tested visual synchrony using gaze location metrics—assessing the convergence of children's visual focus at any given moment. However, gaze location is not the only looking measure linked to attention and cognitive states. Pupil dilation—the process of the pupils increasing in size as a physiological response to visual stimuli—offers a window into the autonomic nervous system, providing insights into cognitive load, emotional arousal, and attention shifts. The aim of the current study was to validate a new technique to test developmental changes in visual synchrony based on pupil dilation. We demonstrate our approach in previously collected data from preschoolers and adults during free viewing of cartoon videos. We found stable and significant time windows where the two age groups differed in synchrony, suggesting different physiological responses to the videos within each age group. All analyses and tutorials are shared. Findings highlight the potential for using pupil dilation to explore how individuals, from children to adults, synchronize their attention and emotions. Such a technique offers a richer picture of what and how children share visual information during observation.

Introduction

Behavioral synchrony—the coordination of actions between individuals—plays a critical role in fostering social cohesion and understanding, significantly influencing group dynamics, cooperation, and the establishment of social connections (Hoch et al., 2021; Marsh et al., 2009; Valdesolo et al., 2010). Within the domain of child development, there is much focus on visual synchrony, and specifically how coordinated visual attention between caregivers and children is instrumental for cognitive and social development (Feldman, 2007; Kuhl et al., 2003). The mechanisms underlying visual synchrony, including joint attention and shared gaze, are foundational for learning (Suarez-Rivera et al., 2019; Yu & Smith, 2012, 2017), facilitating the acquisition of language (Kuhl, 2004; Righi et al., 2018; Schroer & Yu, 2023), the development of theory of mind (Baimel et al., 2015; Brandone & Stout, 2023), and the understanding of social cues (Brooks & Meltzoff, 2005; Mundy & Newell, 2007). These processes are not only important for acquiring specific skills but also for the child's overall developmental trajectory, influencing emotional regulation, social interaction patterns, and the ability to engage in complex social environments (Carpenter et al., 1998; Sebanz et al., 2006).

Traditional research of visual synchrony focused on how individuals coordinate their gaze through shared everyday interactions such as conversations even in the absence of direct visual interaction (Richardson et al., 2007; Risko et al., 2016). The variability in visual synchrony across different contexts, influenced by factors like task complexity, emotional content, and interpersonal relations, further illustrates the adaptive nature of human social cognition (Henderson et al., 2007; Hollingworth et al., 2001). Those findings mostly relied on eye-tracking technologies that record where and how long an individual looks at different areas of a visual field, thereby enabling researchers to identify patterns of how people align their gaze when exposed to the same stimuli, thereby linking perceptual performance and cognitive states (Brennan et al., 2008; Hasson et al., 2012; Louwerse et al., 2012; Richardson & Dale, 2005).

The development of visual synchrony is a crucial aspect of their social and cognitive development, marked by developmental changes in synchrony (Gredebäck et al., 2010; Kirkorian et al., 2009) that reflect a developmental shift from responsive joint attention in infancy (L. B. Smith et al., 2011; Tomasello, 1995; Yu & Smith, 2012) where infants follow the gaze or pointing gesture of caregivers to share attention towards the same object or event to more sophisticated forms of interaction in childhood and adolescence. Those include initiative joint attention (the child takes the lead in directing another's attention towards an object or event of interest; Bruner, 1985), shared attention mechanisms (children not only share a focus but are also aware that they are sharing this focus; (Moore & Dunham, 1995), social referencing (the child looks to a caregiver or another individual for cues about how to react to a novel or uncertain situation), and understanding intentions (recognizing that others have different intentions and desires;(Wellman et al., 2001). Visual synchrony is also linked to developmental changes in neural networks that underpin attention, social cognition, and emotional regulation (Triesch et al., 2007).

Despite the importance of visual synchrony, its development has been studied only using gaze location metrics—assessing the convergence of children's visual focus at any given moment. This metric is beneficial to identify developmental milestones in how children engage with and perceive visual stimuli (Dorr et al., 2010; Hasson et al., 2008; Mital et al., 2011; Shepherd et al., 2010; T. J. Smith & Mital, 2013; ’t Hart et al., 2009; Wang et al., 2012). However, gaze location is not the only looking metric that is linked to attention and cognitive states and can be recorded using eye-tracking technology. Pupil dilation—the process of the pupils increasing in size as a physiological response to visual stimuli—offers a window into the autonomic nervous system, providing insights into the cognitive load, emotional arousal, and shifts in attention (Beatty & Lucero-Wagoner, 2000; Laeng et al., 2012). Presumably, synchronized changes in pupil size across individuals (Kang & Wheatley, 2017; Mathôt et al., 2018; Sirois & Brisson, 2014) introduce an additional, unexplored measure to developmental changes in coordinating their attention.

In the adult literature, researchers showed that when individuals engage in tasks requiring attention or are exposed to emotional stimuli, their pupils tend to dilate in a synchronized manner, suggesting a linkage between their perceptual and emotional processing (Harrison et al., 2006; Murphy et al., 2014). Further, work by Kret and De Dreu (2017) has demonstrated that pupil synchrony can enhance social cohesion and trust within groups, indicating its role in facilitating nonverbal communication and social bonding. Pupil synchrony has also been explored as a marker for shared cognitive workload and mutual understanding during cooperative tasks, highlighting its potential as a tool for measuring the dynamics of social interaction and collaboration (Dalmaso et al., 2012; Elliott et al., 2007). These findings highlight the significance of pupil synchrony as a window into the complex interplay of perception, cognition, and emotion.

Building on these insights, the current study is the first to explore the potential of using pupil synchrony to test developmental changes in visual synchrony. To that end, we performed a secondary data analysis on data collected from the same preschoolers (3- to 5-year-olds) and adults reported in Ossmy et al. (2021). In between blocks of watching hammering tasks, children and a subset of adults were watching the same short video clips. We analyzed changes in their pupil size during this free viewing. We aimed to form an analytic procedure that significantly identifies differences in pupil synchrony between the two age groups. We expected more pupil synchrony within the adult group compared to the child group, and we predicted that these differences would be distributed equally across time and video content.

Methods

With participants' permission, videos and demographic data are shared in the Databrary web-based library: https://nyu.databrary.org/volume/321. The volume also includes all the films used as stimuli and the eye-tracking data that we used for analysis. All analysis scripts and a tutorials are shared in https://github.com/Physical-Cognition-Lab/A-Pupil-Dilation-Technique-to-Test-Developmental-Differences-in-Visual-Synchrony

Participants

The original study tested 22 children from 3.09 to 5.49 years of age (M = 4.06 years; 11 girls) and 22 adults from 19.37 to 26.40 years of age (M = 22.02 years; 14 women). Thirteen adults watched different animated videos and, therefore, were excluded from this secondary analysis. Due to technical issues (n = 1) or not enough data (n = 5; watched less than 75% of each video), our secondary analysis focused on data from 16 children (M = 3.94 years, 14 girls). Children were recruited from families in the NYC area who were interested in participating in psychology research, and the adults were recruited through word of mouth. Participants received a robot toy, photograph magnet, and tote bag for their participation. The experiment conformed to the guidelines approved by the ethics committee at New York University. All participants were healthy with normal vision.

Procedure

Children and adults were seated in either a child or adult-sized chair facing a 60-cm widescreen LCD monitor with a resolution of 1920 × 1200 (Figure 1A). The height and orientation of the monitor were adjusted to align with the participants' eye level. We used remote eye tracking (desk-mounted SMI eye tracker; SensoMotoric Instruments, RED, 120 Hz) to record participants' pupil size while they observed videos of actors using tools and four family-friendly animated short videos in between. We used a 4-point routine to calibrate the tracker and validated the calibration using a second 4-point routine. To ensure the calibration remained accurate, a third 4-point routine was conducted at the session's conclusion.

The current study only focuses only on the animated videos. Data on tool-use observation was previously reported in a different article (Ossmy et al., 2021) and was excluded here. Participants also wore an EEG cap for purposes unrelated to the current study, and their neural activity has been reported in a separate study (Ossmy et al., 2021). Participants were not instructed to focus on anything specific during the observation.

Video Clips

Four animated short films were presented in the same order to all participants (Figure 1A): (1) 'Soar' (270 seconds)—a short film that explores the heartwarming story of a young girl who befriends a tiny boy pilot who crashes his miniature flying machine. Together, they embark on a journey to help him repair his plane, navigate through challenges, and ultimately enable him to soar once again, highlighting themes of friendship, perseverance, and the power of collaboration; (2) 'Dustin' (382.8 seconds)—a CGI animated film about a pug who, much to his chagrin, has to arrange with an automatic cleaning robot as his new roommate; (3) 'Lifted' (258 seconds)—a movie about a young alien that struggles to abduct a sleeping human under the watchful eye of his instructor, facing numerous comedic mishaps as it attempts to control the spaceship's complicated machinery. Despite his failures, the story concludes with an unexpected twist that highlights the value of perseverance and learning from mistakes; and (4) 'Boundin' (248 seconds)—a short musical film on a once-proud lamb loses its confidence after being sheared, feeling exposed and ridiculed by the other animals. A wise jackalope comes along and teaches the lamb the importance of resilience and self-acceptance, inspiring it to "bound" joyfully again regardless of its appearance).

The four video clips differ significantly in visualisation, characters and narrative structures. 'Soar' emphasises creativity and collaboration, engaging participants in problem-solving and the joys of friendship. 'Dustin' introduces humour and the challenges of adaptation through the interaction between a dog and a robot, prompting reflections on technology and companionship. 'Lifted' explores the themes of learning and failure, invoking empathy and amusement as viewers witness the alien's comedic attempts at abduction. Lastly, 'Boundin' deals with resilience and self-acceptance, eliciting emotional responses related to empathy and encouragement as the lamb regains its confidence.

This spectrum of thematic differences allowed us to explore whether differences in pupil synchrony of adults and children are related to perceptual, cognitive, or emotional experiences. Perceptually, the different visual styles may capture attention differently, affecting how observers synchronise their focus on key visual elements. Cognitively, the distinct stories and lessons challenge observers to engage with diverse problem-solving strategies, moral judgments, and understanding of character motivations, potentially aligning or diverging their interpretations and reflections. Emotionally, the range from humour to empathy across the films can synchronize observers' emotional responses.

Each film was split into three time chunks that were interspersed with the tool-use clips. Because we were interested in the effect of time on pupil synchrony, we analysed the different time intervals separately. Further information about the visual stimuli across the entire session can be found in Ossmy et al. 2021.

Pupil Dilation Pre-processing

For each participant, a median filter with a 90-ms window was applied to attenuate noise and smooth the signal (de Winter et al., 2021). After filtering, data were down-sampled to 20Hz, and linear interpolation was used to fill missing time points up to a maximum gap of 1.5 seconds. Any gaps larger than this threshold were not interpolated.

Subsequently, visual inspection was carried out to detect and label sections or entire video chunks that displayed artefacts. Video chunks or portions thereof identified as containing artefacts were filled with NaNs. Following artefact identification, video chunks containing less than 25% of the data were discarded (0 chunks). Participants with less than 25% of the total amount of video chunks were excluded from subsequent analysis (2 children).

Pupil Synchrony Calculation

We calculated a pupil synchrony index PS for the children group (PS_c) and adults group (PS_a). For each time interval, we calculated Pearson correlation between each pair of participants using a centered rolling window of 2 seconds (see Figure 1B). This resulted in a correlation matrix for each time point, representing the similarity in pupil dilation across participants. Averaging the correlation matrix in each group yielded an average PS_c and PS_a for each timepoint. To normalize the distribution for inferential statistics, we applied the Fisher Z-Transformation to the values within the matrix before averaging the correlation coefficients. The transformation facilitated the mean calculation, ensuring equal variance across values, resulting in a more accurate measure of central tendency. The average was calculated by first applying Fisher Z-Transformation to the values within the matrix, and then calculating the mean of the matrix. We used the inverse of the Fisher Z-Transformation to transform the extracted means back. Windows that contained less than 25% of the data were excluded from further analyses.

Analytic Procedure

Figure 1B shows a schematic illustration of our analytic procedure. Given the unequal distribution of participants across the two groups (comprising 9 adults and 16 children), we employed a bootstrapped t-test to ensure unbiased comparisons. We performed 1,000 iterations in the bootstrap sampling procedure, and randomly selected 9 children in each iteration to match the adult group size. Pupil synchrony indices PS_c and PS_a were calculated for each bootstrap sample and each video time point (see Pupil Synchrony Calculation). Subsequently, we calculated the mean pupil synchrony over 10-second bins for each video chunks. Then, a t-test was calculated for each iteration to compare the mean PS_c and PS_a. The *p*-values extracted from these tests were used to fit a density estimation curve, enabling us to identify the most frequently occurring value.

A rolling t-test was used to determine the specific time points in the video segments where pupil synchrony significantly differed between the two groups. This test was conducted using a centered rolling window of 4 seconds, during which we compared the PS_c and PS_a values. The window size was dynamically adjusted at the edges of the data to include the maximum available data. This allowed to calculate a p-value for each time point of the video chunks. We corrected multiple comparisons from the rolling t-tests using False Discovery Rate (Benjamini & Hochberg, 1995). Sections of the videos where there was a significant difference in pupil synchrony between the two groups were identified using the extracted p-values.

We assessed the significance by performing an identical analytic procedure as described above, using the same data but with shuffled participant labels. To obtain a distribution of shuffle-labeled differences in synchrony, this was repeated 100 times. In each iteration, participants were randomly assigned to either the adult or child group. We defined a time window as significant if the synchrony differences based on real labels exceeded the shuffle-labelled distribution with a significance level of .05 (Ossmy & Mukamel, 2018). A tolerance of ±0.25 seconds was allowed in identifying repeated segments to account for minor variations in timing. Our analytic procedure and full tutorial are shared in https://github.com/Physical-Cognition-Lab/A-Pupil-Dilation-Technique-to-Test-Developmental-Differences-in-Visual-Synchrony

Results

Pupil-dilation synchrony between adults differed from pupil-dilation synchrony between children. The analysis of *p*-value distributions from multiple *t*-tests on bootstrapped equal-sized groups revealed peak value of 0.002, indicating a statistically significant difference in pupil-dilation synchrony between adults and children.

Following the identification of a general significant difference between the two age groups, we focused our analysis on the specific time windows in which synchronies differed between adults and children. Figure 2A shows 128 time-windows with significant differences between the two groups (*max* *p*-value = 0.049, *min* *t*-value = -27.81). The significant windows are distributed equally across all video chunks (M = 16.8, SD = 6.95 significant time windows per chunk). We found that the average significant time window duration when comparing adults to children is M = 5.4, SD = 0.8 seconds (see Figure 2B for the duration distribution). No difference was observed across video chunks in the duration of the significant time windows ($F(11, 116) = 0.46$, $p = 0.91$). This supports our prediction that differences do not depend on specific time of video content.

We then examined how many of the time windows we found to differ between age groups are still different when participants are randomly assigned to two groups (see Methods). Table 1 shows that only 20 (15.6%) of the time windows were found in the randomized groups.

Finally, we tested how many significant time windows are found in the randomized groups compared to the age groups and whether there is a significant difference in their duration. Figure 2C shows the distribution of the number of significant time windows in the randomized groups, which is lower than the number of windows in which age groups differ. Moreover, the average time window duration in the randomized group (M = 3.7s, SD = 1.56) was significantly lower than those of the age groups (M = 5.3, SD = 3.1). A Kolmogorov-Smirnov test comparing the distribution of durations in the age-group data compared to the distribution of durations in the randomized-group data was significant ($W = 123678$, $p < 0.001$).

Discussion

Understanding visual synchrony between humans is crucial for understanding their shared experiences and social interactions. This secondary data analysis aimed to explore the possibility of using pupil dilation as a novel metric for assessing visual synchrony over development. We found significant differences in pupil-dilation synchrony between preschool children and adults during free viewing of animated video clips. The findings underscore the developmental changes in how visual stimuli are processed at both cognitive and physiological levels. In essence, the variation in pupil synchrony observed suggests that adults and children do not merely see things differently; presumably, their levels of shared attention and emotional engagement with peers in the same age group vary significantly. Adults typically displayed greater synchrony, implying a more uniform response across individuals in this age group compared to children, who exhibited a more variable response.

Our analysis focused on free viewing because it offers a unique lens into spontaneous and not directed responses. Traditional developmental studies of visual synchrony focus on directed screen-based tasks where participants' gaze is guided (Franchak, 2020). Free viewing, by contrast, allows subjects to look wherever they wish, providing insights into natural viewing habits and unscripted cognitive processing (Franchak et al., 2016). This method can reveal fundamental aspects of developmental cognitive and emotional processes as they occur in real-time, outside the structured settings of many psychological experiments (Madsen & Parra, 2022; Nuthmann & Henderson, 2010).

The differences we found in pupil dilation synchrony provide a window into the distinct perceptual worlds inhabited by children and adults. Pupil dilation, a link to autonomic nervous activity, offers objective cues to underlying shifts in cognitive load and emotional states. For children, whose cognitive and emotional systems are still developing, these cues can be particularly varied. The fact that adults generally showed higher synchrony might reflect a more developed, standardized processing of emotional and cognitive stimuli, which stabilizes with development.

Our technique and tutorials opens new avenues for research into how perceptual processes evolve from childhood to adulthood. By integrating pupillometry with traditional measures of visual synchrony, researchers can gain a more holistic view of the cognitive and emotional resources that are harnessed during visual engagement. Such a view could unveil the dynamics of how children with varying developmental profiles, such as those with attention deficit hyperactivity disorder (ADHD), process visual stimuli not only at the cognitive level but also through the lens of physiological responses (Ansarinasab et al., 2022; Braithwaite et al., 2020; Fitzpatrick et al., 2018; Tansey et al., 2022).

Based on this secondary analysis, we propose that pupil-dilation synchrony can serve as a powerful tool in the toolkit of developmental psychology, particularly in educational settings (Bühler et al., 2024; Schneider et al., 2022). Recognizing the patterns of cognitive regulation through physiological measures like pupil dilation can help tailor educational and developmental interventions more effectively. For instance, understanding the variability in how children engage with visual stimuli can lead to more personalized learning experiences that cater to children's unique developmental needs. Future studies should expand on this foundational research by exploring a broader range of ages and developmental stages, incorporating diverse types of visual stimuli, and examining the potential longitudinal changes in pupil synchrony. Additionally, integrating more granular neurophysiological data, such as EEG measures, could enrich our understanding of the neural correlates of these observed behaviors (Ossmy et al., 2021, 2022). Such integrated research is crucial in harnessing the full potential of pupillometry in developmental science, providing deeper insights into the maturation of visual and cognitive faculties across the human lifespan.

Acknowledgements

We thank Ruben Zamora for helping with the online data collection. This work was supported by the ESRC New Investigator grant ES/W009242/1, BA Talent Award TDA21\210038, Waterloo Foundation grant 917-4975, Leverhulme Trust research grant RPG-2022-327, and the Birkbeck/Wellcome Trust Institutional Strategic Support Fund to OO.

Declaration of interest statement

The authors report there are no competing interests to declare.

References

Ansarinasab, S., Panahi, S., Ghassemi, F., Ghosh, D., & Jafari, S. (2022). Synchronization stability analysis of functional brain networks in boys with ADHD during facial emotions processing. *Physica A*, *603*(127848), 127848.

Baimel, A., Severson, R. L., Baron, A. S., & Birch, S. A. J. (2015). Enhancing "theory of mind" through behavioral synchrony. *Frontiers in Psychology*, *6*, 870.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–162). Cambridge University Press.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *57*(1), 289–300.

Braithwaite, E. K., Gui, A., & Jones, E. J. H. (2020). Social attention: What is it, how can we measure it, and what can it tell us about autism and ADHD? *Progress in Brain Research*, *254*, 271–303.

Brandone, A. C., & Stout, W. (2023). The origins of theory of mind in infant social cognition: Investigating longitudinal pathways from intention understanding and joint attention to preschool theory of mind. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, *24*(3), 375–396.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465–1477.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111–114.

Bühler, B., Bozkir, E., Deininger, H., Gerjets, P., Trautwein, U., & Kasneci, E. (2024). On task and in sync: Examining the relationship between gaze synchrony and self-reported attention during video lecture learning. *Proceedings of the ACM on Human-Computer Interaction*, *8*(ETRA), 1–18.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1–143.

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology Letters*, *8*(3), 450–452.

de Winter, J. C. F., Petermeijer, S. M., Kooijman, L., & Dodou, D. (2021). Replicating five pupillometry studies of Eckhard Hess. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *165*, 145–205.

Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, *10*(10), 28.

Elliott, J. C., Vallines, I., & Peltier, C. (2007). The influence of synchrony on pupillary responses to dynamic emotional expressions. *Cognition & Emotion*, *21*(1), 1–12.

Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *48*(3–4), 329–354.

Fitzpatrick, P., Frazier, J. A., Cochran, D., Mitchell, T., Coleman, C., & Schmidt, R. C. (2018). Relationship between theory of mind, emotion recognition, and social synchrony in adolescents with and without autism. *Frontiers in Psychology*, *9*, 1337.

Franchak, J. M. (2020). Visual exploratory behavior and its development. In *Psychology of Learning and Motivation* (pp. 59–94). Elsevier.

Franchak, J. M., Heeger, D. J., Hasson, U., & Adolph, K. E. (2016). Free viewing gaze behavior in infants and adults. *Infancy: The Official Journal of the International Society on Infant Studies*, *21*(3), 262–287.

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848.

Harrison, N. A., Singer, T., Rotshtein, P., Dolan, R. J., & Critchley, H. D. (2006). Pupillary contagion: central mechanisms engaged in sadness processing. *Social Cognitive and Affective Neuroscience*, *1*(1), 5–17.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(10), 2539–2550.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In *Eye Movements* (pp. 537–III). Elsevier.

Hoch, J. E., Ossmy, O., Cole, W. G., Hasan, S., & Adolph, K. E. (2021). “dancing” together: Infant-mother locomotor synchrony. *Child Development*, *92*(4), 1337–1353.

Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, *8*(4), 761–768.

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology. General*, *146*(4), 569–576.

Kirkorian, H. L., Pempek, T. A., Murphy, L. A., Schmidt, M. E., & Anderson, D. R. (2009). The impact of background television on parent-child interaction. *Child Development*, *80*(5), 1350–1359.

Kret, M. E., & De Dreu, C. K. W. (2017). Pupil-mimicry conditions trust in partners: moderation by oxytocin and group membership. *Proceedings. Biological Sciences*, *284*(1850). https://doi.org/10.1098/rspb.2016.2554

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience*, *5*(11), 831–843.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(1), 18–27.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*(8), 1404–1426.

Madsen, J., & Parra, L. C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus*, *1*(1), gac020.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320–339.

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106.

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, *3*(1), 5–24.

Moore, C., & Dunham, P. J. (1995). *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum Associates.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274.

Murphy, P. R., Vandekerckhove, J., & Nieuwenhuis, S. (2014). Pupil-linked arousal determines variability in perceptual decision making. *PLoS Computational Biology*, *10*(9), e1003854.

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision*, *10*(8), 20.

Ossmy, O., Han, D., Kaplan, B. E., Xu, M., Bianco, C., Mukamel, R., & Adolph, K. E. (2021). Children do not distinguish efficient from inefficient actions during observation. *Scientific Reports*, *11*(1), 18106.

Ossmy, O., Kaplan, B. E., Han, D., Xu, M., Bianco, C., Mukamel, R., & Adolph, K. E. (2022). Real-time processes in the development of action planning. *Current Biology: CB*, *32*(1), 190-199.e3.

Ossmy, O., & Mukamel, R. (2018). Behavioral and neural effects of congruency of visual feedback during short-term motor learning. *NeuroImage*, *172*, 864–873.

Richardson, D. C., & Dale, R. (2005). Looking to understand: the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407–413.

Righi, G., Tenenbaum, E. J., McCormick, C., Blossom, M., Amso, D., & Sheinkopf, S. J. (2018). Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. *Autism Research: Official Journal of the International Society for Autism Research*, *11*(4), 645–653.

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science. *Current Directions in Psychological Science*, *25*(1), 70–74.

Schneider, B., Sung, G., Hassan, J., & Bhinder, H. (2022). What Matters Most for Learning From Online Videos, Seeing the Instructor's Face or Gaze? Impact on Instructor-Student Synchrony. In *Proceedings of the 15th International Conference on Computer-Supported Collaborative Learning-CSCL 2022*.

Schroer, S. E., & Yu, C. (2023). Embodied attention resolves visual ambiguity to support infants' real-time word learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

Shepherd, S. V., Steckenfinger, S. A., Hasson, U., & Ghazanfar, A. A. (2010). Human-monkey gaze correlations reveal convergent and divergent patterns of movie viewing. *Current Biology: CB*, *20*(7), 649–656.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews. Cognitive Science*, *5*(6), 679–692.

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: the dynamics of toddler visual experience. *Developmental Science*, *14*(1), 9–17.

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision*, *13*(8). https://doi.org/10.1167/13.8.16

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*(1), 96–109.

't Hart, B. M., Schmidt, K. L., & Roth, E. A. (2009). Saccadic eye movements are accurate and precise in the central visual field. *Experimental Brain Research*, *192*(1), 1–20.

Tansey, R., Graff, K., Rohr, C. S., Dimond, D., Ip, A., Dewey, D., & Bray, S. (2022). Inattentive and hyperactive traits differentially associate with interindividual functional synchrony during video viewing in young children without ADHD. *Cerebral Cortex Communications*, *3*(1), tgac011.

Tomasello, M. (1995). *Joint Attention: Its Origins and Role in Development*.

Triesch, J., Jasso, H., & Deák, G. O. (2007). Emergence of mirror neurons in a model of gaze following. *Adaptive Behavior*, *15*(2), 149–165.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, *46*(4), 693–695.

Wang, H. X., Freeman, J., Merriam, E. P., Hasson, U., & Heeger, D. J. (2012). Temporal eye movement strategies during naturalistic viewing. *Journal of Vision*, *12*(1), 16.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655–684.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262.

Yu, C., & Smith, L. B. (2017). Hand-eye coordination predicts joint attention. *Child Development*, *88*(6), 2060–2078.

Figures

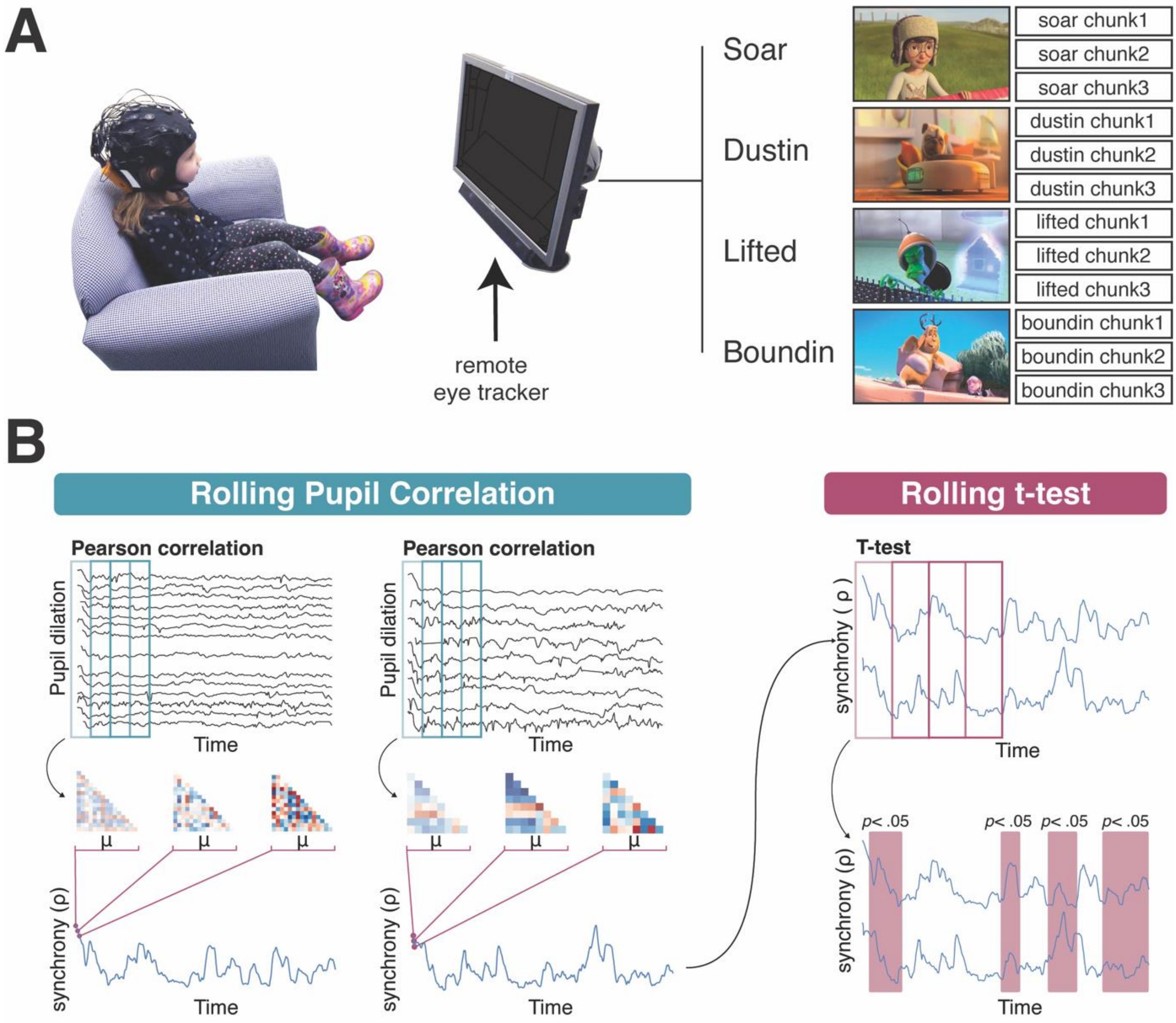

Figure 1. (A) Experimental setup: A child participant is seated in front of a monitor with a remote eye tracker. The screen displays images from four different animated short films (Soar, Dustin, Lifted, and Boundin), each divided into three chunks. **(B)** Schematic illustration of the analytic procedure. The left panel shows the Rolling Pupil Correlation analysis, depicting Pearson correlations of pupil dilation over time for two separate groups. The correlations are visualized as time series and corresponding heatmaps. From each correlation matrix, a single value was extracted as the average Pearson correlation. The right panel illustrates the Rolling t-test analysis, showing the comparison of synchrony between groups over time, with highlighted areas indicating statistically significant differences ($p < .05$).

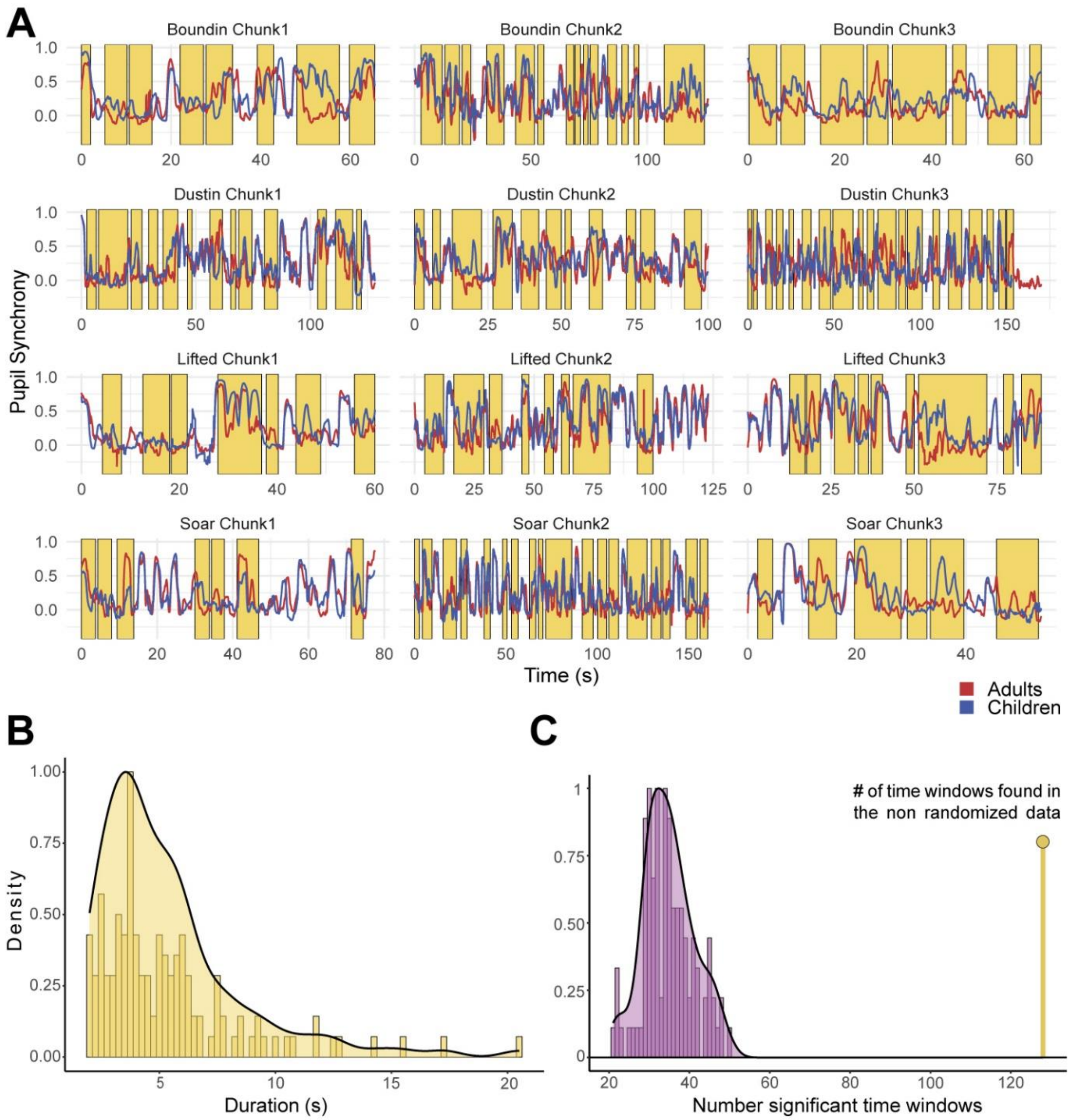

Figure 2. Analysis of pupil synchrony across different video chunks and groups. **(A)** Pupil synchrony over time for adults (blue) and children (red) across all chunks of four animated short films: Boundin, Dustin, Lifted, and Soar. Yellow highlighted areas indicate time windows with significant differences in synchrony between adults and children. **(B)** Density distribution of the duration of time windows where significant pupil synchrony differences were observed. The x-axis shows the duration in seconds and the y-axis shows the density. **(C)** Density distribution of the number of significant time windows which are longer than 2 seconds and found in randomized data. The purple histogram shows the distribution from randomized data, while the yellow dot represents the number of significant time windows found in the non-randomized (actual) data.

	START	END	How many times the time window were found in the randomized data
Boundin Chunk1	0.00	2.05	1
Boundin Chunk1	39.20	42.85	1
Boundin Chunk2	20.45	24.25	1
Boundin Chunk2	83.05	86.90	1
Boundin Chunk3	44.45	47.40	1
Boundin Chunk3	61.20	63.70	3
Dustin Chunk1	46.25	48.35	1
Dustin Chunk2	0.00	3.20	2
Dustin Chunk2	77.05	81.90	1
Dustin Chunk3	0.00	2.25	4
Dustin Chunk3	16.55	20.35	1
Dustin Chunk3	69.30	73.05	1
Lifted Chunk3	82.30	88.25	2
Soar Chunk1	0.00	3.70	2
Soar Chunk1	29.95	33.80	1
Soar Chunk1	71.30	74.45	1
Soar Chunk2	0.00	2.55	4
Soar Chunk2	25.35	28.75	1
Soar Chunk2	67.85	70.25	2
Soar Chunk2	135.90	139.95	1

Table 1. A list of time windows where significant differences in pupil synchrony between adults and children were found in the pupil synchrony analysis and were also found in the randomized data analysis.