### Title
Learning with Limited Attention




### ISBN
9798189276637



### Author
Shepherd, Sarah Stolp



### Publication Date
2026-05-01

Peer reviewed|Thesis/dissertation

Learning with Limited Attention

by

Sarah Stolp Shepherd

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Celeste Kidd, Chair
Professor Jan Engelmann
Professor Stephen Hinshaw
Professor Steven Piantadosi

Spring 2026

Learning with Limited Attention

Abstract

Learning with Limited Attention

by

Sarah Stolp Shepherd

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Celeste Kidd, Chair

This dissertation presents empirical studies on how both informational utility and visual salience guide attention in learners across species. An individual may attend to something due to its usefulness for learning or due to its 'bright and shiny' features like bold colors and high contrast. A long history of work in developmental psychology establishes these features as foundational organizing principles of the human attention system—both predict looking time even in infancy. Digital information environments increasingly leverage the psychology of attention to maximize engagement and ad revenue. In this dissertation I investigate these foundational mechanisms of attention in digital information environments. Chapters 2 and 3 demonstrate that engagement does not predict learning in young children, as adult observers strongly assume, with implications for education, policy, and technology design. We find that attentional capture due to visual salience can disrupt learning in young children, despite capturing more attention. In Chapter 4, we demonstrate that low-income children consume more low-quality media content using a novel tool to quantify a metric of problematic visual salience, flicker, and discuss the implications of this inequity on children's learning. Chapter 5 investigates how media consumption preferences impact learning and processing in young adults. We find that processing is likely shaped by media consumption preferences, in line with an adaptive account of attentional bandwidth. Finally, Chapter 6 investigates how general informational utility is as an organizing principle of attention by testing a distantly related mammal, dogs. Children and nonhuman primates preferentially attend to events of intermediate surprisal, likely due to these events offering the greatest potential for learning, while dogs do not. This work investigates how attention interacts with technology engineered to capture and sustain attention, and demonstrates that the mechanism through which attention is captured may shape processing and learning. This dissertation provides a perspective on human attention in digital information environments and poses broader questions about the evolutionary origins of attention in supporting complex processes like curiosity and learning.

# Contents

# List of Figures

# List of Tables

## Acknowledgments

This work was made possible through the collaboration and generous support of many people along the way. I think I drafted these acknowledgments a thousand times in my head as a means of procrastination or to remind myself that one day I would actually be writing them after completing the daunting task of a PhD. At the end of this five-year journey, I'm overwhelmed with gratitude for everyone who has guided my learning and lent support.

To my dissertation committee—Steve, Steve (" the Steves"), Jan, and my advisor Celeste—thank you for your generous and thoughtful guidance. You created an environment for me to explore my curiosity and challenged me to do so through a rigorous scientific lens. The knowledge you shared shapes both this body of work and me as a person. Thank you for investing in these ideas and my education.

To Celeste especially, thank you for teaching me how to do good science, driven by rigor, integrity, and curiosity. You model how to complement scientific rigor with high impact and joy. Academic work can at times feel tedious and impersonal, and I will forever be grateful for an advisor who heard "cool" and "fun" as valid justification for pursuing ideas. Thank you for the many coffees, breakfasts, bowls of udon, and countless hours of thoughtful guidance.

To the Kidd Lab for bringing together people who became lifelong friends and valued collaborators. Holly Palmeri, our outstanding lab manager, thank you for making all this work possible. Thank you to the research assistants who supported data collection and the families who participated in these studies. Evan and Alex, thank you for your friendship and camaraderie over the years. International travel, 3 am deadlines, and navigating the uncertainty of grad school were infinitely more fun with you both.

To my family and friends—especially Mom, Dad, Matt, my grandparents Heidi and Gerry, and my extended Stolp and Shepherd families—thank you for all your enthusiastic support. Loving a PhD student can be difficult at times as experiments fail unpredictably and we learn how to communicate why our work is so important that it consumes half a decade of our lives—you made it seem easy. This dissertation represents not just the five years of graduate training but a love of learning I inherited from a family of educators. To Matt, who enthusiastically uprooted our life in Wānaka to move to Berkeley and kept me sane through the ups and downs of grad school. Thank you for all your love and support.

To all my teachers, mentors, and peers who invested in my education through the years. The public schools in Eugene, Oregon, the University of Otago and my New Zealand family, and UC Berkeley. Learning is a gift; thank you to everyone who shared it with me.

# Chapter 1

# Introduction

The attention system is well-adapted to support learning. A long history of developmental psychology research demonstrates that attention is organized in ways that guide people towards information useful for learning (e.g., Berlyne, 1960; Kidd et al., 2012; Oudeyer and Smith, 2016; Schulz and Bonawitz, 2007.) These patterns of attention are widely appreciated as foundational for cognitive development, supporting skills from language acquisition to causal learning (Baldwin, 1991; Oudeyer & Smith, 2016; Schulz & Bonawitz, 2007). Empirical evidence demonstrating learning value as an organizing principle of attention requires other factors that predict attention to be controlled—like contrast, hue saturation, motion, and other sources of visual salience—factors that can be superficial cues unrelated to a material's information value. However, in real-world environments these features are not controlled for and may therefore influence attention in ways that are not supportive of learning. In this dissertation I investigate how visually salient features guide attention and how attentional capture due to visual salience impacts learning.

## 1.1 What guides attention?

There are two important features that guide attention that I will focus on in this dissertation: information value and visual salience. This dissertation will focus on these two features in the applied context of learning from digital materials, where these features are often manipulated to maximize engagement and reach.

### Information value

Information value for learning guides attention. Information-seeking behavior is guided by material that is expected to reduce the gap between what is known and what is unknown, relative to an individual's current set of mental representations (Berlyne, 1954, 1960; Dember & Earl, 1957; Hunter & Ames, 1988; Kidd et al., 2012, 2014). People look at things that are useful for learning, as demonstrated by decades of research on infant attention, but also

Maria Montessori's concept of "readiness" or Vygotsky's "zone of proximal development" (Montessori, 1917; Vygotsky, 1978). Infants, young children, and primates allocate their visual attention in ways that are expected to support learning and curiosity-driven exploration (Cubit et al., 2021; Kidd et al., 2012, 2014; Poli et al., 2020; Twomey & Westermann, 2018; S. Wu et al., 2022). This work demonstrates that individuals allocate their attention in ways that are expected to adaptively regulate information-gathering to support learning when other features are controlled for. Visual attention can reasonably be expected to reflect learning value in these cases. Chapter 6 investigates information value as a general organizing principle of attention by testing a more distantly related mammal, dogs. We find evidence that information value does not guide attention uniformly across mammals.

### Visual salience

Attention also selects perceptually salient visual material such as motion onset, high visual contrast, and hue saturation (Aslin, 2007; Cohen, 1972; Haith, 1980; Salapatek & Kessen, 1966). Empirical evidence demonstrating learning value as an organizing principle of attention requires these factors to be controlled. Visually salient features—like contrast, hue saturation, and motion—can be superficial cues unrelated to a material's information value. Disengagement with material may become harder in environments that are especially visually salient, potentially creating competition between low-level attentional capture and learning-optimal disengagement with content. Here, we investigate the impact of increasing attentional capture due to visual salience on subsequent learning.

## 1.2 Attention in digital contexts

Herbert Simon observed that an abundance of information creates a scarcity of attention (Simon et al., 1971). Simon's theory of the attention economy is salient in current digital environments. The limited capacity of human information processing drives our modern attention economy: a virtually endless amount of information is available in digital environments and must compete for limited attention.

The scarcity of attention available to consume the abundance of digital content has become an economic problem. The competition for attention is driven by ad revenue and influence, where content that captures the most attention is rewarded (Hayes, 2025; T. Wu, 2016). Algorithms on popular platforms rely on engagement metrics to sort or otherwise suggest content for users (Lewandowsky et al., 2023). The highest-grossing content is therefore often the content with the highest engagement. Disengagement—or shifting attention away—on these platforms is therefore an obstacle to ad revenue and influence.

**Kids and the attention economy** The competition for attention in digital environments shapes not only the distribution but also the creation of content. The emphasis on most-watched content encourages the creation of content aimed at attaining high visual engagement. This is especially true for the largest children's media platforms, like YouTube Kids, Roblox, and Twitch. Creators of popular children's media frequently use engagement metrics to guide content production, specifically minimizing moments when children shift attention away (Segal, 2022). To this end, digital content increasingly leverages low-level perceptual features known to capture and sustain visual attention (Goodrich et al., 2009; Lewandowsky et al., 2023; Segal, 2022; T. Wu, 2016). However, disengagement may be an important signal for learning, and attempts to minimize it may undermine learning, especially in young children.

## 1.3 Overstimulation & visual salience

Promoting content expected to minimize disengagement may foster patterns of attention that are not just detrimental to learning but harmful to cognitive development. Higher visual salience is theorized to disrupt learning due to overstimulation (Christakis et al., 2018; Essex et al., 2022; Loschky et al., 2020; Shepherd & Kidd, 2024). Fast-paced and fantastical children's content is known to reduce executive function in children immediately after viewing, likely because the high visual salience of this type of material disrupts processing (Anderson et al., 1977; Essex et al., 2022; Geist & Gibson, 2000; Lillard, Li, & Boguszewski, 2015; Lillard & Peterson, 2011). Young children's still-developing executive function makes them more at risk for worse cognitive and behavioral outcomes associated with this content. Early exposure to television viewing is associated with disordered attention, hyperactivity, reduced executive function, and language delays (Kostyrka-Allchorne, Cooper, Gossmann, et al., 2017; Lissak, 2018; Madigan et al., 2019). Due to the nature of this work, evidence on overstimulation is largely observational and relies on general screen time measures to draw conclusions. Intervention studies with mice demonstrate that excessive exposure to highly salient audio-visual content during development can lead to disordered attention, cognitive deficits, and behavioral changes (Christakis et al., 2012). One proposed mechanism for these outcomes is that fast-paced, visually salient content captures exogenous attention, which biases bottom-up processing, overloads executive function, and ultimately disrupts goal-directed learning and attention (e.g., Christakis et al., 2018; Loschky et al., 2020). Here, we investigate the conflict created between exogenous and endogenous attention present in current digital environments and its impact on processing. Chapters 2 and 5 demonstrate that attentional capture due to visual salience can disrupt learning in young children and young adults, likely due to overstimulation.

## 1.4 Learning with limited attention in a digital world

In this dissertation, I investigate how features of the environment impact the relationship between attention and learning. I use digital environments as a test case for how increased visual salience may disrupt learning, despite capturing more attention. Chapter 2 demonstrates that attentional capture due to visual salience can disrupt learning in young children, despite capturing more attention. This work is followed up with an adult study in Chapter 5, where we provide evidence for an adaptive account of attentional bandwidth in which perceptual salience is not objective but rather is sensitive to an individual's optimal range of stimulation for learning. I employ behavioral eye tracking with young children and adults to investigate attentional signatures of overstimulation in digital contexts. Eye tracking findings are presented in concert with survey data on participants' media habits and preferences, as well as behavioral responses while viewing material, like fidgeting. Chapter 3 presents the *lack* of fidgeting as a possible behavioral indicator of overstimulation in young children, contrasted with the strong expectations adults hold that the absence of fidgeting signals learning. Chapter 4 presents an applied computer vision analysis of children's media, demonstrating that poor children are disproportionately exposed to visually salient digital content expected to be overstimulating and therefore most at risk for the harms this content could cause. Finally, Chapter 6 focuses on information value as an organizing principle of attention. We demonstrate that information value does not uniformly guide attention across all mammals through an eye-tracking study with dogs. This work is foundational for thinking more broadly about how different attention systems may differentially prioritize features to solve the problem of information overload.

Together, this work provides a nuanced picture of how the shift to digital information environments is shaping attention and learning across development, with implications for education, policy, and technology design. This dissertation draws on a long history of attention research with young children to investigate these concepts in digital environments. Changes to how information is consumed reflect the need for a new understanding of how human cognitive capacities interact with and are shaped by the environment. This work is essential because many modern information environments are designed to leverage findings from psychology to maximize engagement, often with little regard for learning. This work represents the first account of how attention in digital contexts may be shifting with individuals' media preferences and experiences, possibly leading to distinct attentional profiles. We take an applied approach, grounded in developmental psychology, to investigate learning and attention in digital environments.

# Chapter 2

# Visual engagement is not synonymous with learning in young children

## 2.1 Introduction

Equating sustained visual attention with learning is a popular trend among creators of children's media content. However, decades of research has demonstrated that learning value is not the sole driver of attention—or even the primary one (Haith, 1980). Attention may be guided by informational utility or highly salient perceptual features (Berlyne, 1954). Despite this literature, visual attention in popular media settings is often assumed to result from the material's learning value. For example, visual engagement is used in testing children's entertainment materials for suitability in promoting language development and literacy (Albahiri & Alhaj, 2020; Anggraini et al., 2022; Toleuzhan et al., 2022).

The trend of children's edutainment content creators conflating high visual engagement with high learning value is not new (Ryan, 2014); however, the impact and reach of the conflation have increased substantially due to recent changes in the quantity, availability, and financial incentives driving the creation of content for young children (Neumann & Herodotou, 2020).

Understanding the distinction between visual engagement and learning is pressing because more young people are watching this content more frequently for longer durations. This content is more readily available around the world via smartphones, and it's pushed harder by marketing tactics due to increased financial incentives in the new attention economy. Here, we apply classic methods from developmental psychology to test the relationship between visual engagement and learning in young children. We find no evidence of increased attentional engagement leading to greater learning in 3- to 6-year-old children, and in fact that higher rates of visual engagement for visually complex displays lead to *less* learning in participants. We draw insights and inspiration from relevant developmental psychology, connect it to recent changes in why and how content is created for children, and then present empirical evidence from a novel eye-tracking study we designed that suggests how the conflation of attentional engagement and learning in children is flawed.

## The psychology behind children's attention

Attention is necessary for learning, and previous work demonstrates that it can reflect a readiness to learn (e.g., Maria Montessori's concept of "readiness" or Vygotsky's "zone of proximal development", (Montessori, 1917; Vygotsky, 1978)). Supporting empirical evidence demontrates children and non-human primates selectively attend to events of intermediate surprisal, which are theorized to hold greater informational utility (Cubit et al., 2021; Kidd et al., 2012, 2014; Oudeyer & Smith, 2016; S. Wu et al., 2022).

At the same time, attention does not *only* reflect learning value. Attention also selects perceptually salient visual material such as motion onset, high visual contrast, and hue saturation (Aslin, 2007; Cohen, 1972; Haith, 1980; Salapatek & Kessen, 1966).

While it's possible that attentional selection due to perceptual salience could lead to greater learning gains in young children—as edutainment producers commonly suggest—this has remained an untested claim. We investigate here.

## Attention in economic context

The widespread adoption of the internet as a primary source of information, communication, and entertainment has monetized attention. In the modern economy, attention is valuable. By its very definition, attention is a finite resource. The intrinsic limits of our information processing abilities drive the attention economy: as the world becomes increasingly information rich, it must necessarily become attention poor (Simon et al., 1971). This dynamic creates competition for engagement in many information-rich but attention-scarce environments, and is evident on many popular streaming platforms.

The competition for attention is also driven by substantial financial incentives. In the US alone, digital video ad spending grew by 21% to 47.1B in 2022 (IAB, 2023). Not only is digital media advertising revenue growing rapidly, approximately half of digital video ad purchasers are quantifying viewers' visual engagement with biometric measures of attention, such as eye-tracking, demonstrating the value placed on consumer attention (IAB, 2023). The emphasis on most-watched content encourages a practice of creating content aimed at attaining high visual engagement. In line with this significant growth in online advertising revenue, many popular platforms employ recommending algorithms that prioritize content expected to maximize sustained visual attention (Lewandowsky et al., 2023; T. Wu, 2016). This includes many popular children's media platforms, like YouTube Kids, Roblox, and Twitch. The content itself, beyond the advertising within the platforms, also emphasizes attention as an important gauge of success. Creators of children's media frequently use engagement metrics to guide content production. For example, CoComelon Nursery Rhymes is the most popular children's YouTube channel, with over 167 billion views as of August 2023, and is known to test how effectively their content visually engages children as young as two years old when placed near another screen showing naturalistic scenes (Segal, 2022). This practice is not necessarily unique to the streaming era; a review of children's DVDs noted the high concentration of low-level perceptual features being used in educational children's content nearly 15 years ago (Goodrich et al., 2009).

## 2.2 Methods

### Participants

Fifty-five 3- to 6-year-old children completed the study across two locations. Twenty-five participants volunteered at a local science museum near the UC Berkeley campus and thirty participated in a developmental lab space on campus. Four additional participants were excluded for parents not complying with experimenter instructions, and a further three additional participants were excluded due to technical issues. Testing location had no significant impact on results. All participants were compensated for their time either with $10 or a small prize.

### Stimuli

Six novel objects (toys) were pseudo-randomly paired with novel words. The novel objects were small figurines of colorful monsters (see Figure 3.1). All toys were photographed in the same lighting conditions and backgrounds were removed. These photos of the toys were used for the screen-based portion of the task. Six familiar objects were selected from the MacArthur-Bates Communicative Development Inventory of earliest learned words (Dale & Fenson, 1996). The familiar objects were only used in the screen-based portion of the task, so simple, recognizable illustrations of these objects were selected.

Videos were used in the complex condition only. Videos were selected to be of colorful, age-appropriate naturalistic scenes with varied movement that did not include a clear singular subject. All audio was removed from videos.

### Procedure

Participants were seated on their parents lap for the duration of the experiment. Exceptions were made for children who requested to sit on a booster seat or when parents were unable to hold the child on their lap. In these cases, parents sat behind their child or in the waiting room if requested by the child. To avoid parents influencing their child's behavior, all parents wore blindfolds or closed their eyes and listened to music through headphones while children participated in the study. In cases where parents closed their eyes experimenters confirmed that they complied throughout the study. Visual stimuli were presented on a monitor connected to a Gazepoint GP3 eye-tracker approximately 60cm away from the participant. The stimuli were presented from a Windows laptop using Psychopy software (Peirce et al., 2019). The eye-tracker was calibrated using a nine-point display of a shrink-

Figure 2.1: The learning trials in the simple condition presented the toys on a static gray background. The learning trials in the complex condition presented the toys in front of colorful videos of naturalistic scenes. Audio of child-directed speech labeled the toys with their novel names and was the same in both conditions.

ing dot that the participant fixated as it moved around the screen. An identical five-point validation sequence followed, allowing the experimenter to verify the calibration had been successful. If necessary, the experimenter repeated the calibration to achieve a satisfactory validation.

## Learning Trials

The experiment consisted of both a screen-based eye tracking portion and a short test using the novel objects on the table in front of the participant. The screen-based portion consisted of both a learning and a test phase. The learning phase consisted of 12 trials during which the novel objects were presented with their corresponding novel word pairings. For each trial the novel object was presented with two carrier phrases "Do you see the X?" and "What a nice X!". Each novel object was presented twice in a random order during the learning trials.

## Test Trials

After the learning trials the experimenter introduced the test trials as a game with toys that participants "might know the names of". The first set of test trials were the screen-based implicit test trails. Each implicit test trial presented two objects, a target and a distractor, on screen with an auditory prompt "Can you find the X?". The first 12 test trials asked participants to find a named novel object. Test trials 13-24 asked participants to find the known objects using the same format. These known-object test trials were included to check that participants were completing the task as expected. The known-object trials always occurred after the novel-object test trials. All test trials were pseudo-randomized with each


Figure 2.2: We asked participants to identify names of novel toys. Participants learned these novel word-object associations via screens. We tested their learning via screen and with physical objects.

object used twice as the target and twice as the distractor. Each trial displayed a unique pairing. Participant's test trial was deemed correct if their looking time to the correct object was greater during a two-second window following a 200ms delay from the time of the target word onset to account for any saccade latency. This window of analysis and criteria were the same for both novel-and known-object implicit test trials.

The final portion of the experiment consisted of explicit test trials. Children were asked to identify physical versions of the novel objects, providing an explicit measure of their learning. Physical stimuli were kept out of view from the participant until the screen-based eye-tracking portion of the study was completed. The experimenter then presented all six novel objects on the table and asked the participant to hand them each toy by name, one at a time. Between trials the experimenter replaced the selected toy and shuffled all six toys to prevent participants from selecting toys in serial order, and to make it more difficult to track which toys had already been selected. Experimenters told participants it was okay to guess if they were not sure of their answer and gave no feedback between physical trials. Participant responses were recorded on video for later coding.

### Conditions

Participants were randomly assigned to either the simple or complex condition. In the complex condition, a unique video played in the background of each learning trial. Videos did not repeat such that participants should not associate a particular referent in a video with a novel word label. Each participant in the complex condition saw 12 videos in a different random order across all 12 learning trials. Stimuli in the simple condition were presented on a static gray background. Test trials and physical object trials were the same for both conditions.

### Eye-tracking data processing

Eye tracking data was excluded from all trials with greater than 90% trackloss. After removing these trials with extreme trackloss, we averaged each participant's looking time across their validly tracked trials to calculate their average looking time for the entire screen and areas of interest. This allowed us to compare a participant's looking behavior in the learning trials to their performance on the test trials. We further excluded participants who were missing data from either all learning trials or all test trials. This resulted in excluding a further 4 trials. After processing the data in this manner, we were left with complete observations from 49 participants. 26 in the simple condition with a mean age of 4.955, and 23 in the complex condition with mean age 4.996.

### Behavioral analysis: fidget ratings

We investigated a behavioral signal often associated with learning and engagement, fidgeting. Participant videos were viewed and rated for fidgeting behavior by two coders naive to condition. Fidgeting during the learning trials was rated on a scale from 1 to 5, with 1 indicating minimal movements throughout and 5 indicating frequent or pronounced movements (e.g., shifting, wiggling, large body movements). Ratings were given to all participant videos, Five participants were excluded due to missing or incomplete videos, resulting in a final sample of 43 for the fidgeting analysis.

## 2.3 Results

### Increased visual engagement in complex condition

Participants looked to the screen during learning trials significantly longer in the complex condition (Figure 2.3). This result was expected as the complex condition was designed to increase visual engagement through the use of low-level attentional attractors. On average, participants looked to the learning trials 87.41% of the time in the complex condition, compared to 81.04% in the simple condition. This difference was confirmed to be significant by a Wilcoxon rank sum exact test (W = 463, $p$ <0.001)

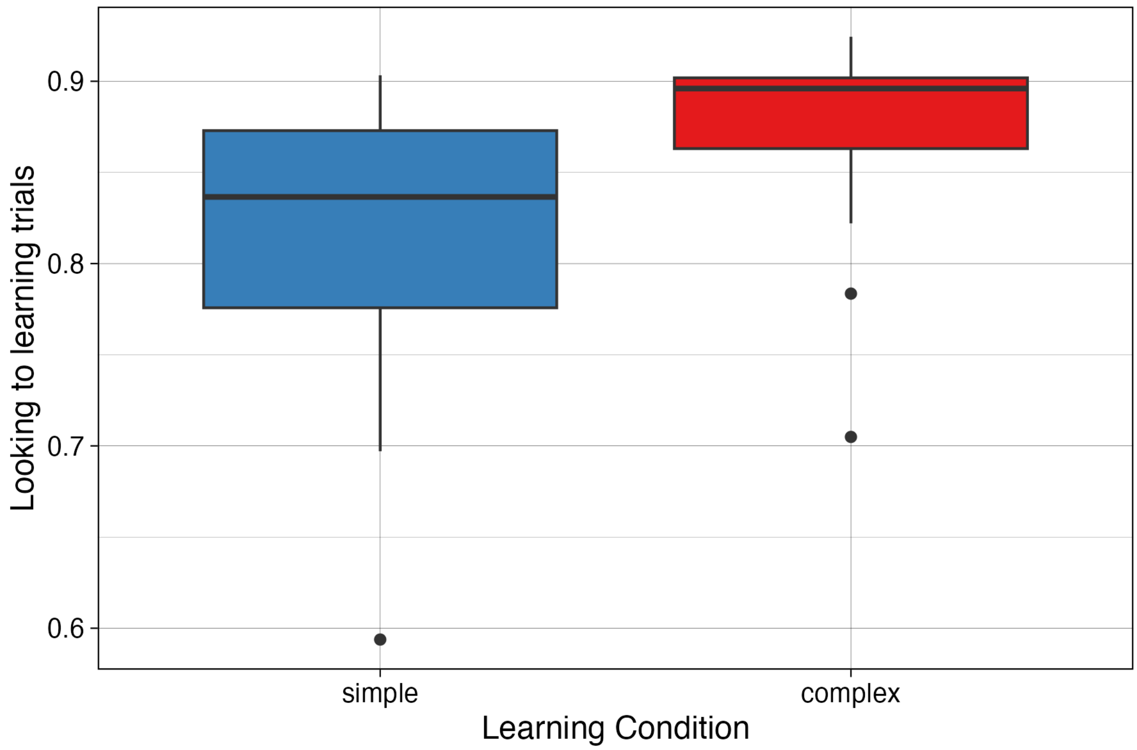


Figure 2.3: Participants attended to the screen more in the complex condition, where increased low-level perceptual features were present during learning trials.

## Visual salience did not predict learning outcomes

As we expected, learning was not improved by increased visual engagement. Overall, participants demonstrated that they learned the names of the toys above chance in both the explicit and implicit test trials (see Figure 2.4). There was no significant difference in learning between conditions on either implicit or explicit word learning. On average, participants in the simple condition correctly identified the correct toy on 53% of explicit trials, compared to 47% in the complex condition. This difference was not significant in a Wilcoxon rank sum exact test (W = 260.5, $p$ = 0.44). Implicit learning was nearly identical across conditions. Participants looked to the correctly named item in 72.39% of trials in the simple condition and 71.64% of implicit trials in the complex condition.

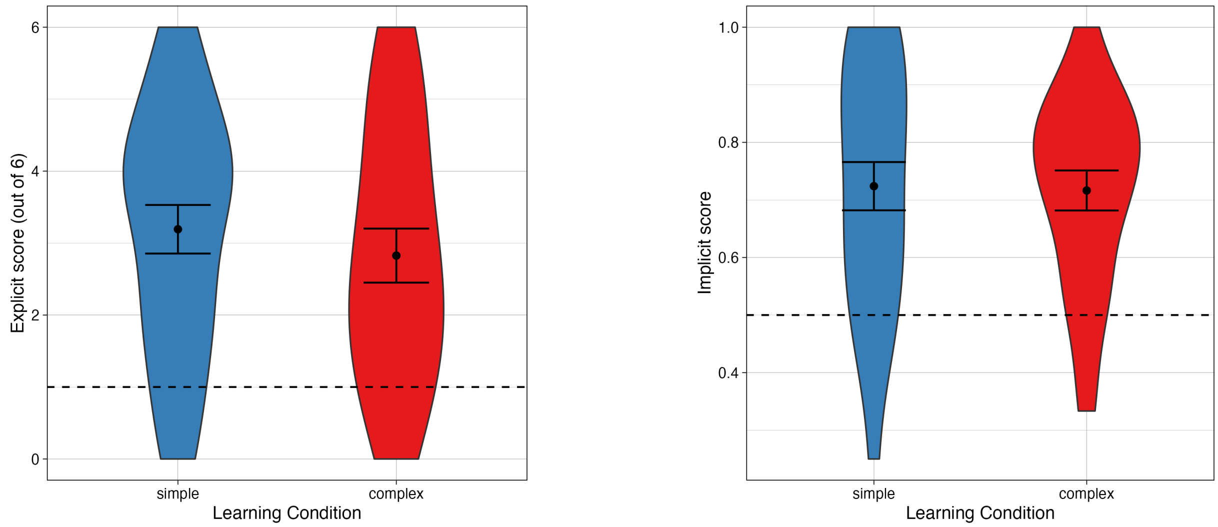


Figure 2.4: Participants learned the toy names similarly in both conditions, despite increased attention to the learning trials in the complex condition.

## Increased visual engagement predicted worse learning when distractors were present

A linear regression controlling for age revealed that the interaction between a participant's looking time to the screen during learning trials and the condition they were in was a significant predictor of their performance on explicit learning trials ($\beta = 1.823$, $p = 0.0017$). For participants in the complex condition, increased looking to the screen during learning trials predicted worse outcomes on the explicit trials. In the simple condition looking time to the learning trials was associated with better explicit learning scores. The difference in this interaction between conditions is shown in Figure 2.5. This interaction was not observed in the implicit learning trials. The implicit learning trials were forced choice, and participants were largely successful in identifying the correct labeled toy.

## Increased focus to target object did not explain learning trends

We found no evidence that participants who had better focus to the target object during the learning trials learned more word-object associations. Both overall looking time to the target and looking to the target as a proportion of all looking to the screen did not predict explicit or implicit learning in either condition.

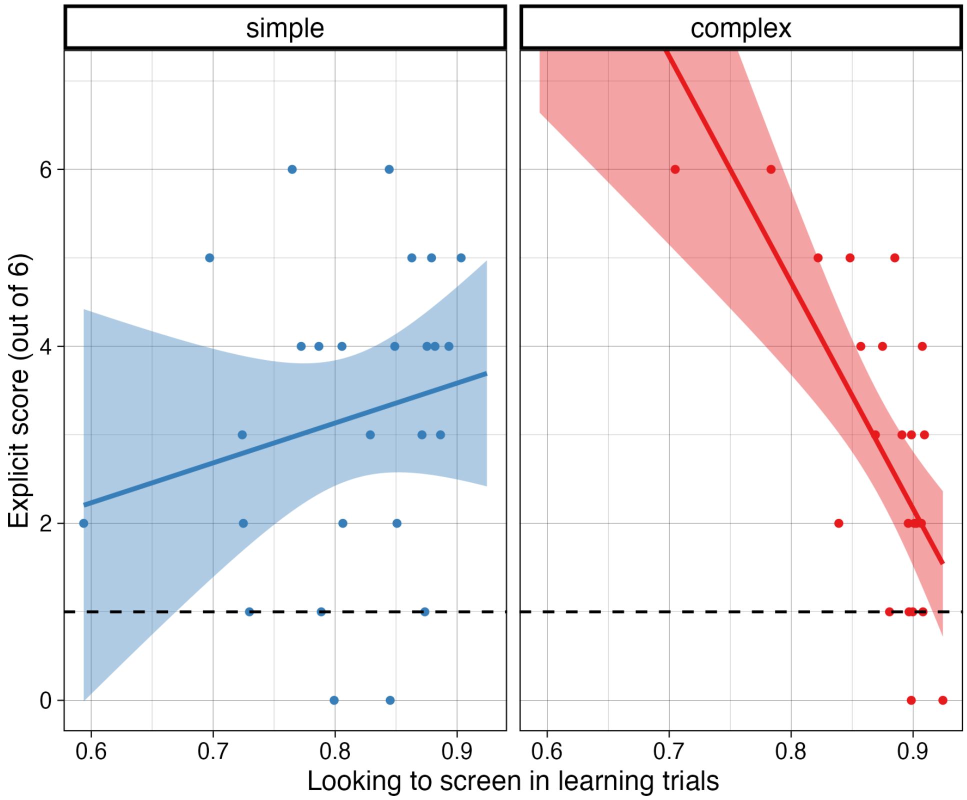


Figure 2.5: Increased visual attention during learning trials in the complex condition was associated with worse learning in the explicit trials. Participants who engaged the most learned the fewest toy names compared to less-attentive participants in the same condition.

## Increased fidgeting in simple condition

Inter-coder reliability for fidget ratings was high, with Cohen's kappa (k = 0.84) indicating strong agreement between the coders. Participants fidgeted more in the simple condition (M = 3.70, SD = 1.27) than in the complex condition (M = 2.85, SD = 1.38) (see Figure 2.6). A two-sample t-test confirmed this difference to be statistically significant ($t(41) = -2.31$, $p = 0.03$, Cohen's d = -0.71). Fidgeting did not predict learning in either condition (see Figure 2.7). Children who were more physically engaged during learning did not learn more word-object pairings. These findings challenge the assumption that physical stillness reflects meaningful engagement; despite being instructed to sit still, lower fidgeting was not associated with greater learning.

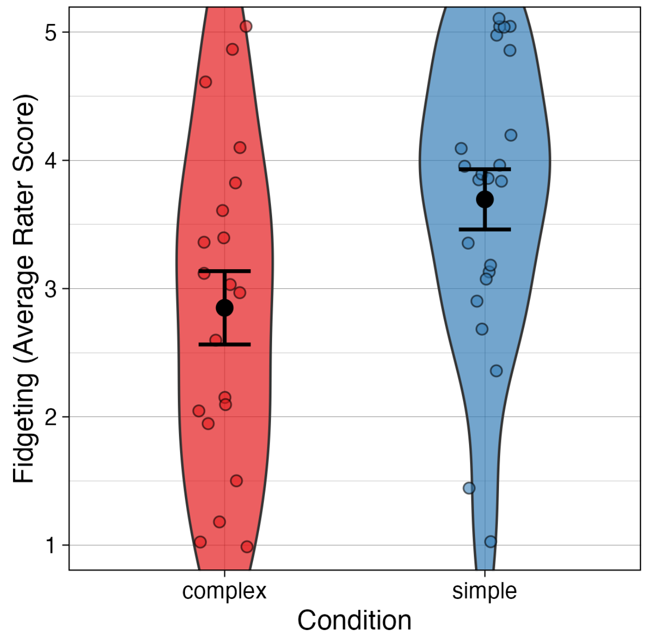


Figure 2.6: Participants fidgeted more during learning trials in the simple condition.

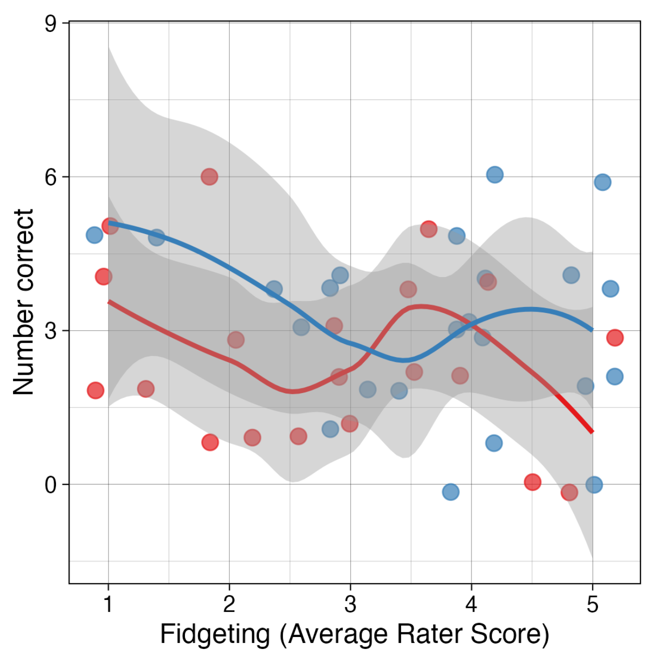


Figure 2.7: Fidgeting did not predict learning in either condition.

## 2.4 Discussion

We found that increasing children's visual attention through heightened salient, low-level perceptual attentional attractors resulted in worse learning outcomes on a word learning task. In fact, these low-level cues elicited greater attention while hindering word learning for some children within the same condition. These highly engaged children may have learned something about the backgrounds that is not captured in our learning measures. However, their poor performance on a relatively simple and clearly defined learning task is still notable considering less-attentive children in the same condition performed better on the task. If anything, highly engaged children should have learned the names of the toys as well as something about the background videos. While we cannot speak to what they may have learned about the ambient background scenes, their increased attention did not benefit them on the word learning task and demonstrates the risk of highly engaging content that creates competition for learning, especially when this content aims to be educational. Although highly engaged children may be learning less or learning at a slower rate, they appear more willing to watch the content that is more likely to disrupt their learning for longer. Whether longer viewing sessions would offset learning disruptions for these highly engaged children would be an interesting future direction.

Children watching the visually complex stimuli fidgeted significantly less than children in the visually simple condition. However, despite higher rates of physical and visual engagement, these children did not learn more word-object pairings. Crucially, sustained visual attention predicted learning deficits in the visually complex condition. Across both conditions, fidgeting did not significantly predict learning. This suggests that children may be adaptively self-regulating, fidgeting more when watching understimulating simple content and less when watching highly stimulating content. Decreased fidgeting in the complex condition may indicate overstimulation, making children *appear* highly focused while actually impairing their ability to disengage. These results further support the theory that attention captured through low-level salience may overstimulate children, overloading their executive function and ultimately disrupting processing. This pattern of fidgeting behavior suggests that children's ability to disengage is further shaped by the visual salience of stimuli.

The high-stakes competition to capture and sustain children's attention in online environments is driving the production of highly visually engaging content that may undermine well-intentioned educational messages. Using highly salient perceptual features to captivate young viewers not only interferes with their potential learning but may also increase the risk for adverse cognitive and behavioral outcomes. Our findings align with theories of overstimulation that link the adverse effects of screen time to the excessively salient audiovisual properties common in highly engaging children's content.

The potential harms of screen-time during early childhood are well documented, with known associations between excessive early screen use, developmental delays, and worse academic performance (Madigan et al., 2019). Specifically, children's screen use has been tied to disordered attention, impaired vision, sleep disruption, and poor mental health outcomes (Lissak, 2018). However, the research on children's screen usage and its potential harms is largely observational, making it difficult to determine the causal structure of this relationship despite the well-established associations. One theory suggests that overstimulating audio-visual content, like screen-based media, may be responsible for the negative cognitive and behavioral outcomes associated with children's screen time (Christakis et al., 2018). This theory began in part with observations of a relationship between fast-paced television content and disordered attention in young children (Anderson et al., 1977; Geist & Gibson, 2000). Further observational evidence has shown that fast-paced and fantastical children's television reduces young children's executive functioning for at least a short time after exposure (Essex et al., 2022; Lillard & Peterson, 2011). The theory that developing brains are susceptible to potential negative effects of overstimulation is also supported by animal models: mice exposed to heightened audio-visual stimulation early in life display both behavioral and cognitive deficits (Christakis et al., 2012; Christakis et al., 2018).

Our findings support this theory of overstimulation. Within the complex condition alone, maximized attention may be a sign of overstimulation, explaining the poor learning outcomes for children that barely looked away from the screen. These findings suggest that overstimulating children's media that maximizes visual engagement through low-level salient visual features can be detrimental to children's learning and development. It's important to consider that while our experimental design consisted of colorful, engaging videos, they were all of naturalistic scenes which we expect to be less stimulating than much of the content young children are engaging with on popular media platforms. This suggests that the potential negative impact of highly stimulating children's content may be even more severe in common online environments.

This study cannot directly speak to the level at which the presence of distractors would detract from learning. However, this experimental design provided participants with a better chance to learn simple associations than most children's content that is designed to be highly engaging. There was no interfering audio, and all background distractor videos were of ambient scenes that did not repeat with a corresponding toy to avoid confusion about which object was being labeled. Moreover, children in both conditions performed at the same level, further demonstrating that visual engagement, not the presence of distractors, was the moderating factor in the relationship between condition and explicit learning outcomes.

Recommendations for children's screen time remain general and often unclear to parents, in part due to the lack of evidence demonstrating a causal link between screen time and its potential detrimental effects. In the streaming era, where anyone can easily create and upload a video to YouTube or TikTok, placing the burden on caretakers to assess the quality of the media their children consume is overwhelming. Highly stimulating children's content poses a serious risk to the learning and development of young children that can be hard for adult caretakers to uncover because it encourages sustained visual engagement, which may appear as learning or enjoyment from the viewer.

Children's content, especially that which claims to be educational, needs to be created with the potential detrimental impacts of overstimulation in mind. We should require more careful and rigorous evaluation of children's media. In the absence of formal guidelines, caregivers should be cautious of media that may be highly engaging and perceptually salient.

# Chapter 3

# Adults equate stillness with learning

## 3.1 Introduction

Learning is notoriously hard to infer from behavior, despite decades of effort by developmental psychologists trying to address this problem. Even formal measures of learning, like tests, are imperfect and often impractical to gauge in-the-moment learning. Researchers and educators have relied on proxies like sitting still or looking at learning materials to estimate engagement or learning (Sinatra et al., 2015). However, empirical work demonstrates that behavioral indicators of engagement are imperfect predictors for student performance (Li et al., 2022). Inferring internal processes like attention and learning from behavior is a challenge for caregivers, educators, and psychologists seeking to better understand and support kids' learning.

Efforts to quantify learning and attention in real time often rely on physical movement and gaze location (Godwin & Fisher, 2018; Parambil et al., 2022; Sanders et al., 2021). However, these measures do not perfectly capture student performance and over-index on behaviors weakly linked to learning outcomes (Aslin, 2007). Measures of student engagement, even in formal, well-controlled educational contexts, remain a challenge (Henrie et al., 2015).

Physical and visual engagement can reflect myriad different underlying processes (Aslin, 2007). For example, visual engagement—meaning uninterrupted visual attention—does not predict learning in young kids, and in some cases, greater visual engagement actually predicts worse learning (Shepherd & Kidd, 2024). Similarly, greater fidgeting does in fact predict better learning in kids with ADHD, likely because it increases arousal which facilitates sustained attention for these kids (Hartanto et al., 2016; Son et al., 2024). Fidgeting also supports creativity more broadly without impairing performance on traditional academic tasks (Oppezzo et al., 2025).

Stillness might also be a difficult signal to interpret. Kids and teens are significantly more still during sedentary screen-based activities like gaming than during non-screen sedentary activities like drawing (Ekris et al., 2018). If we assume that stillness reliably signals attention and learning, then we would predict that kids playing video games would reliably be learning more than kids reading or drawing.

Here, we test whether adults perceive stillness as indicating greater focus and better learning through a simple experiment that asks adults to estimate the learning of a child in videos. Our videos use a confederate child imitating real videos of children engaged in learning who are either still or fidgety for ecological validity with the benefit of allowing us to control for the specific features of the child in the video by holding them constant. We find that adults highly associate stillness with better attention and learning, even though empirical evidence does not support this belief.

This finding has important implications for adult educators and caregivers who may act as gatekeepers to learning opportunities for children. A child who is perceived as inattentive or a poor learner because they are fidgeting may not be offered the same learning opportunities as a still counterpart; likewise, a still child who may be overstimulated and thus learning less may not have this obstacle to their learning noticed and addressed.

This project is part of a larger body of ongoing work investigating the relationship between engagement and learning, particularly in the context of learning from screen-based media. This experiment represents the first in a line that aims to disentangle the complex relationship between engagement and learning. The larger project aims to provide evidence for educators and caregivers to better support children's learning given the complicated relationship between observable signals of engagement and learning outcomes.

## 3.2 Methods

### Participants

Two-hundred adult U.S. participants fluent in English were recruited on Prolific to participate in a two-minute survey involving watching short videos and answering multiple choice questions.

### Stimuli

Participants watched two five-second videos of the same child: one in which the child sat still while watching a video and one in which the child fidgeted while watching a video. Participants were told that the child was watching a video but otherwise had no information about what the child was watching or had been directed to do. Videos were reconstructions of participants watching videos in a separate lab-based task. A child actor reenacted videos that were selected of especially fidgety and still kids. Fidget movements were isolated to hand movements and the child actor maintained visual attention on the screen throughout in both conditions.

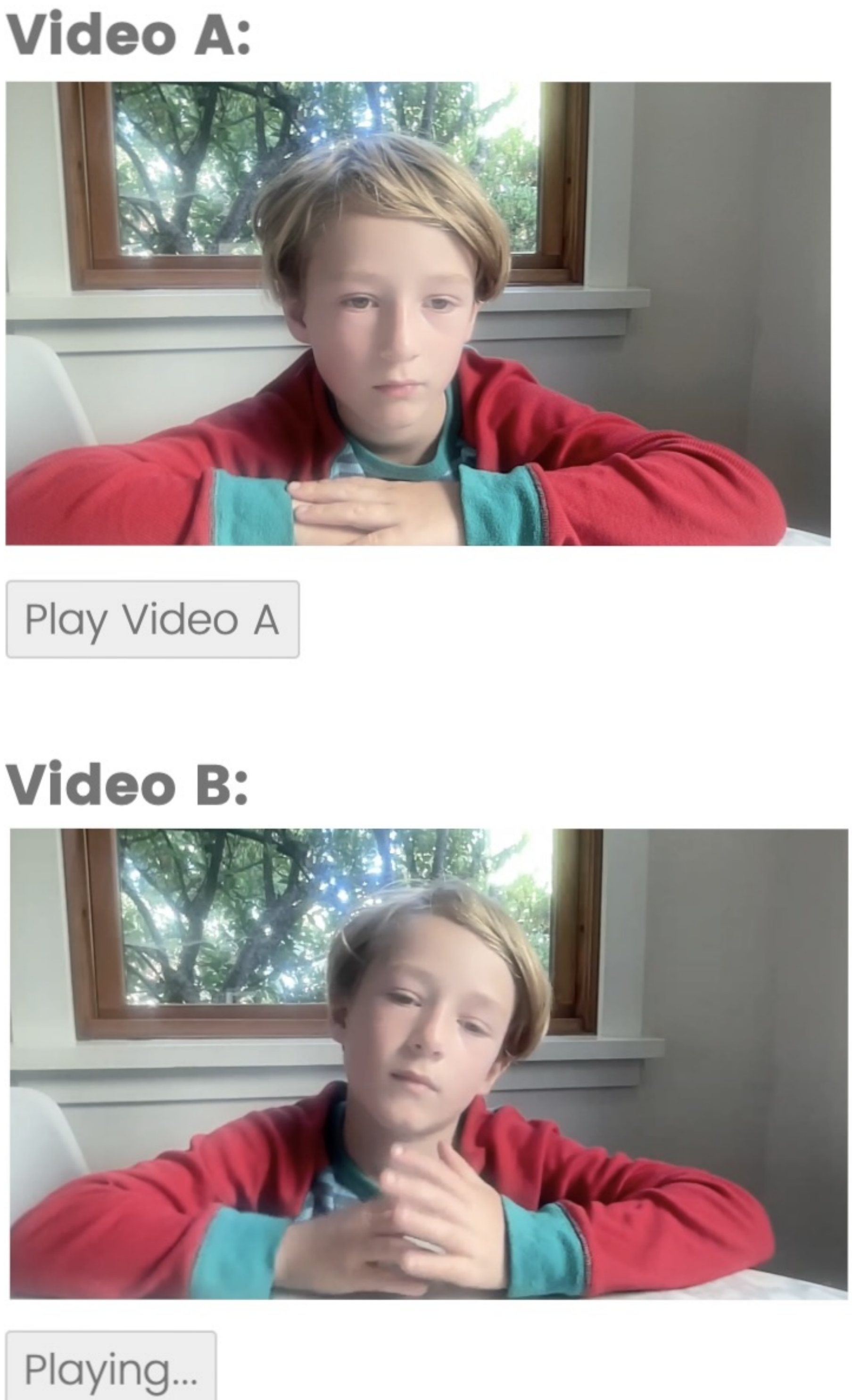


Figure 3.1: Example stimuli videos of a child confederate/actor sitting still (Video A) and fidgeting (Video B) while watching a video. The child actor imitated videos of participants who were watching videos themselves in past studies to increase the ecological validity of the experiment. Study participants were asked to select which video showed greater attention, learning, and focus. (See methods.)

### Procedure

Videos shown to participants were randomly selected from four fidgety and four still videos to control for any specific fidget behaviors. Video order was counterbalanced. Participants were required to watch both videos through before answering questions and then were allowed to re-watch videos at any time while answering. After watching each video, participants were asked five questions in random order and asked to answer by selecting either "Video A" or "Video B".

1. In which video do you think the child is paying the most attention to what they are watching?
2. In which video do you think the child is most focused on what they are watching?
3. In which video do you think the child is learning more from what they are watching?
4. In which video do you think the child is fidgeting more?
5. In which video do you think the child is moving more?

Questions four and five served as a manipulation check to see if participants correctly identified videos as fidgety and still. Seven participants did not accurately select the video with more movement and were removed from analysis, leaving a final sample of N=193 ($M_{age}$= 41.8, SD = 12.2, range = 39- 79) .

## 3.3 Results

### Participants think stillness indicates attention and focus

Participants selected the video of the still child as being more focused. 181 of 193 participants selected the still video when asked which video they thought the child was more focused on what they were watching. This was significantly above chance, confirmed by an exact binomial test (95% CI = $[0.89, 0.97]$, $p$ ¡ .001).

Participants also selected the video of the child being still as more attentive. 178 out of 193 participants selected the still video when asked in which video the child was paying the most attention to what they were watching. This was significantly above chance, confirmed by an exact binomial test (95% CI = $[0.88, 0.96]$, $p$ ¡ .001).

### Participants think stillness indicates learning

The child sitting still while watching a video was also selected as learning more from what they were watching, 171 participants selected the still child out of 193. This was again confirmed significant with an exact binomial test (95% CI = $[0.83, 0.93]$, $p$ ¡ .001).

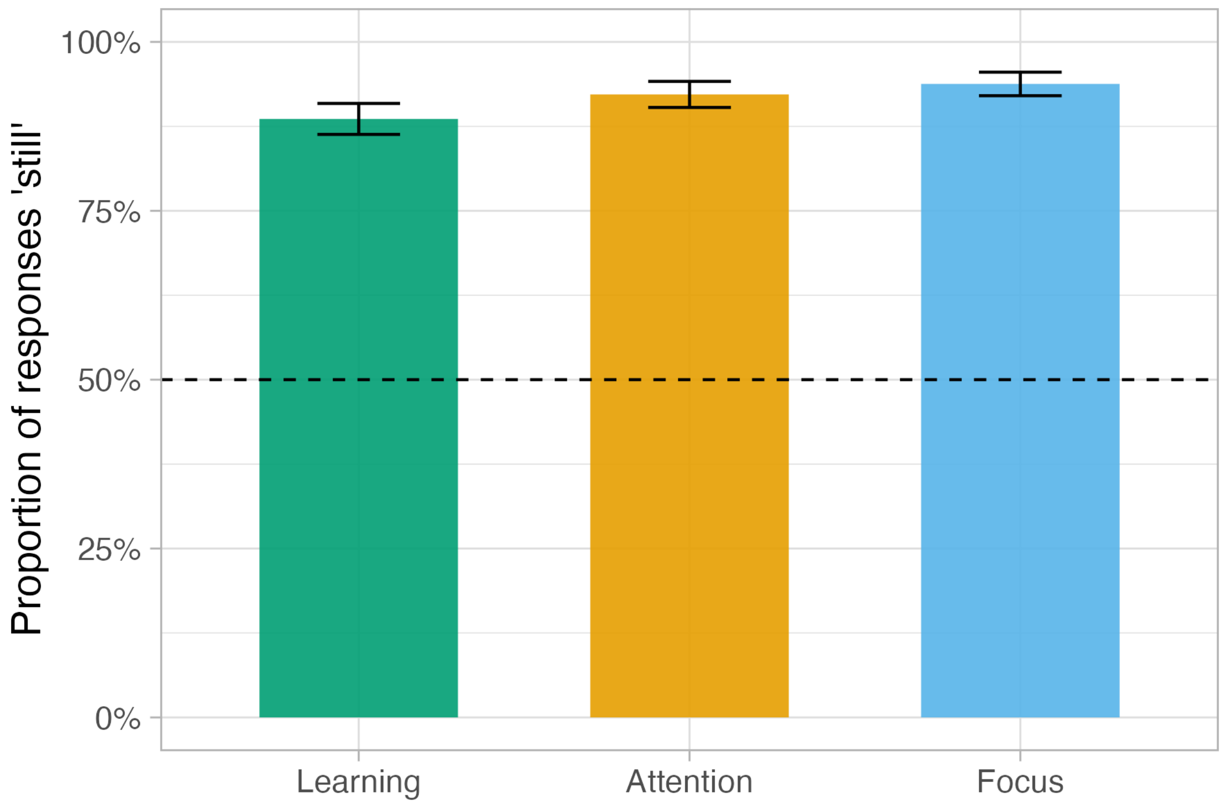


Figure 3.2: Participants rated sitting still as showing greater learning, attention, and focus. Error bars represent SE. The dotted line represents chance. That each of these bars is significantly over at chance and approaching ceiling demonstrates that adults strongly associate stillness with better learning, attention, and focus despite the lack of empirical evidence to support this intuitive association.

## 3.4 Discussion

Our results demonstrate that adults strongly interpret a child's physical stillness as indicating that they are attending and learning better—though empirical evidence suggests otherwise. This fact could have profound implications for inequities in learning opportunities offered to children by their teachers and caretakers. Children who may be more inclined towards movement due to natural variation in their sensory preferences or processing systems may be overlooked for advanced learning opportunities and, likewise, children who are still from overstimulation may have the impediment to their learning go unnoticed.

This is particularly relevant as informal screen-based learning is increasingly happening on tablets, smart TVs, and other screens at home. Adults may be less likely to intervene when they observe a still child and interpret stillness as high engagement and learning from screen-based media. This risks conflating signs of overstimulation with learning. Young kids whose attentional control is not fully developed are especially susceptible to the visually engaging features in popular kids media (Blakley et al., 2021). These low-level attentional cues may therefore induce stillness by capturing attention through means entirely unrelated to learning value. As demonstrated in Chapter 2, children fidget *more* while watching visually salient material that is less conducive to learning. Stillness while watching visually salient content may reflect a low-level orienting response, supporting the theory that perceptual salience may bias bottom-up processing and overload executive function. We propose that similar to visual engagement, physical engagement is often equated with learning but may reflect overstimulation in some cases.

We highlight the strong belief that physical engagement indicates attention and learning as a potential source of inequity in learning opportunities for kids with different sensory needs. Kids who fidget more in classroom or childcare settings where sitting still is emphasized may lose out on learning opportunities if the adults present assume their fidgeting indicates lack of focus and learning.

Our findings also have important implications for the study of learning and engagement. Fidgeting supports learning and creative thinking. Although empirical results are still unclear, we suggest that the relationship between physical engagement and learning is highly complex, likely depending on both the context and individual differences in need for arousal. While the exact nature of the relationship between engagement and learning remains somewhat opaque, the belief that stillness equates to attention and learning is widely held.

Finally, these results have important implications in cognitive science and psychology, where stillness of a child's body or eyes are sometimes used as a proxy for useful attention or learning (Aslin, 2007). This study is the first in a line of in-progress and planned research that investigates the exact relationship between stillness and learning considering individual differences in optimal arousal levels for learning, and we hope it will inspire other researchers to reconsider assumptions that stillness means greater focus in future work.

# Chapter 4

# Bright, shiny garbage: Video content shown to low-income children is characterized by higher flicker

## 4.1 Introduction

Interventions to bridge the gap in school performance between children from low- and high-socioeconomic households have largely focused on increasing learning opportunities for low-income kids (Camilli et al., 2010; Dietrichson et al., 2017; Gorey, 2001; Gormley et al., 2008; Kim & Quinn, 2013). Interventions have inspired programs like Head Start and No Child Left Behind that improve language, reading, and math outcomes by increasing the quantity of learning opportunities for low-income children (Gormley et al., 2008; Gormley Jr. et al., 2005; Kline & Walters, 2016; Weiland & Yoshikawa, 2013). These programs provide the costly types of learning experiences that lower-income families lack the resources to provide to their children without aid (Alexander & Entwisle, 1988; Crosnoe et al., 2010; Lareau, 2018; Magnuson et al., 2004).

Here, we analyze an overlooked source of inequity in learning opportunities for children, which is those that occur in informal settings like video viewing on phones, tablets, laptops, and smart TVs at home. Children in lower-income families have more screen time—more than double in recent reports—than children from higher-income families (Christakis et al., 2004; Hedderson et al., 2023; Rideout & Robb, 2020). We provide the first evidence from a machine vision model to demonstrate that the *quality* of content served to low-income children is an additional source of inequity. Our results show that low-income children are more often served content that is visually engaging but known to be associated with learning disruptions, disordered attention, and cognitive deficits.

Screen time recommendations may be unrealistically low for families who rely on screens to manage caregiving demands (Hiniker et al., 2016). The American Academy of Pediatrics advises no screen time for children under 2 years and one hour per day for children ages

2 to 5 years, with an emphasis on co-viewing (COUNCIL ON COMMUNICATIONS AND MEDIA et al., 2016). However, children under 2 years average more than one hour of screen time per day (Rideout & Robb, 2020). Children are spending more time than recommended on screens, often out of necessity to supplement caregiving, especially in low-income families. What children watch also varies by household income; for example, half of 0- to 8-year-old children from low-income families watch YouTube daily compared to just 29% of children from high-income families (Rideout & Robb, 2020). We investigate how the *quality* of media available to children in families of different income levels may be a further source of inequity.

We provide evidence that children's media available to low-income families is higher in visual salience, which likely disrupts learning and can lead to worse cognitive and behavioral outcomes. Our findings also show that the prevalence of visual salience in children's media is increasing, suggesting that platforms that prioritize engagement metrics over educational value are driving trends in children's media consumption and production. Our findings suggest that ad-supported media available to lower-income families relies more on low-level "bright and shiny" features to capture children's attention.

Our findings suggest that we could do more to protect vulnerable populations of kids' learning opportunities beyond the existing high-cost, high-effort interventions. They suggest that these populations could additionally benefit from programs and policies to protect children's access to high-quality entertainment and educational content. Common Sense Media publishes detailed reports on children's programming and the suitability of all mainstream television shows and movies for viewing by children. If we better understood the relationship between features like flicker and learning outcomes, and could easily extract these from media sources, published guides made available to parents, educators, and the public at large could be used to guide lower-income children towards higher-quality content, which would in turn put pressure on existing producers of the low-quality content to improve their products.

## Visual salience disrupts learning

Understanding visual salience across digital media is crucial because it disrupts learning in young children. Higher visual salience is theorized to disrupt learning by triggering overstimulation (Christakis et al., 2018; Shepherd & Kidd, 2024). Fast-paced and fantastical children's content reduces executive function immediately after viewing, likely because the high visual salience of this type of material disrupts processing (Essex et al., 2022; Lillard, Drell, et al., 2015). Preschool-aged children are more likely to experience disrupted learning from these types of materials because their executive function is less developed than that of older children (Best & Miller, 2010). This means young children are also more at risk of worse cognitive and behavioral outcomes associated with overly salient content. Early exposure to television viewing is associated with disordered attention, hyperactivity, reduced executive function, and language delays (Kostyrka-Allchorne, Cooper, & Simpson, 2017). Further, we know that excessive exposure to highly salient audio-visual content can lead to disordered attention, cognitive deficits, and behavioral changes from intervention studies in animals (Christakis et al., 2012). This extensive literature underscores the importance of examining how visual salience varies across different types of children's media.

### Quantifying media quality: visual salience

To evaluate the quality of media available to children of different socioeconomic backgrounds, we analyzed flicker in ad-supported, paid, and publicly funded children's media. Ad-supported media is available at no cost to families, whereas paid media on streaming platforms costs U.S. consumers an average of $54 per month (J.D.Power, 2022). We focus here on flicker as a measure of visual salience. Flicker captures changes in luminance over time and is the strongest predictor of gaze location in leading models of visual salience (Itti & Koch, 2001; Mital et al., 2011). Flicker is often conflated with editorial features, like cuts and scene changes, and closely matches pacing measures that predict children's reduced executive function after TV viewing (Cooper et al., 2009; Essex et al., 2022; Irwin et al., 2000; Kostyrka-Allchorne, Cooper, & Simpson, 2017; Kostyrka-Allchorne et al., 2019; Lillard, Drell, et al., 2015).

If ad-supported media contains higher levels of flicker than paid content, these systematic differences in visual salience would disproportionately affect children from low-income households who watch significantly more of this content (Rideout & Robb, 2020). This is in part because ad-supported media is available at no cost to families. By analyzing flicker in children's media, we aim to determine whether differences in visual salience contribute to socioeconomic disparity in the quality of learning material and opportunities.

## 4.2 Methods

### Data collection

Videos were randomly selected from popular children's media advertised for preschool-aged children across three categories: ad-supported, paid, and publicly available educational media. Ad-supported media was sampled from the most popular children's YouTube Kids channels based on both views and subscribers, including Kids Diana Show, Cocomelon, and El Reino Infantil, for example (Socialblade, 2024). Individual videos from YouTube Kids were randomly sampled, relying on current views, video length, and date posted to maintain a balanced sample that represents what many children are likely watching. Importantly, videos from YouTube were recorded using a premium subscription so no ad breaks were included. Paid media was selected from streaming platforms Disney+, Netflix, Hulu, and Amazon Prime; however, the content may be offered by more than one streaming platform at any given time. These paid TV episodes were selected at random from highly popular children's programs based on viewership data and colloquial reports of currently popular children's TV—including shows like Bluey, Paw Patrol, Cocomelon, and Peppa Pig (Doperalski, 2024; LuminateData, 2024; Netflx, 2023). Where available, paid TV episodes were sampled from multiple different production years. Public educational media was selected from the Public Broadcasting Service (PBS) online streaming platform, which is available at no cost in the U.S.. PBS TV episodes were randomly sampled from the recommended episodes available to local PBS viewers from popular shows like Arthur, Sesame Street, and Alma's Way (DIRECTV, 2025). None of the videos were viewed by the researcher team prior to selection. Samples from all three sets include both animated and live-action videos. All videos sampled are included in the final analyses. For the full list of media included and available metadata see Appendix A. All data is available on the Open Science Framework, including links to YouTube videos used in analysis. Anonymous link: https://osf.io/yfbsj/overview?view_only=5590fcb5a86f48c1b4040118db936664.

## Data preparation and cleaning

All videos were screen-recorded on an Apple M2 Mac built-in 14-inch display at the highest resolution available, up to 1080HD50p. After recording, introduction sequences (e.g., "theme songs") and end credits or advertised "subscribe now" sequences were removed. This was done to ensure the comparison was on the content itself, which children may be expected to learn from. Title slides were not removed. Each final recorded video was reviewed by researchers to ensure there were no errors in the recordings. Finally, all videos were downsampled to 256×256 resolution for analysis. Final videos are matched on frame rate and resolution.

## Analysis

Videos were processed using OpenCV, an open-source computer vision library in Python to get our final measure of mean flicker. Each frame of each video was changed to CIELab color space to separate color from luminance. The absolute difference between the luminance channels of adjacent frames was calculated and summed across all pixels in each frame. Finally, luminance differences were averaged across all frames in the current video to calculate mean flicker. After calculating mean flicker for each video, we weighted these means by video length in frames for the analyses reported in our results.

For a video made up of a series of frames $F_1, F_2, \ldots, F_n$, we calculate the mean flicker across all frames. First, we convert each frame, $F_i$, to CIELab Color Space and extract the L channel (luminance) as $L_i$. Next, for each frame $L_i$, we compute the absolute difference between the current, $L_i$, and previous, $L_{i-1}$, luminance channels (if not the first frame). This is done for all pixels in the frame:

$$\Delta L_i = |L_i - L_{i-1}|$$

We then sum the absolute luminance differences across all pixels for the current frame:

$$\text{TotalDiff}_i = \sum \Delta L_i$$

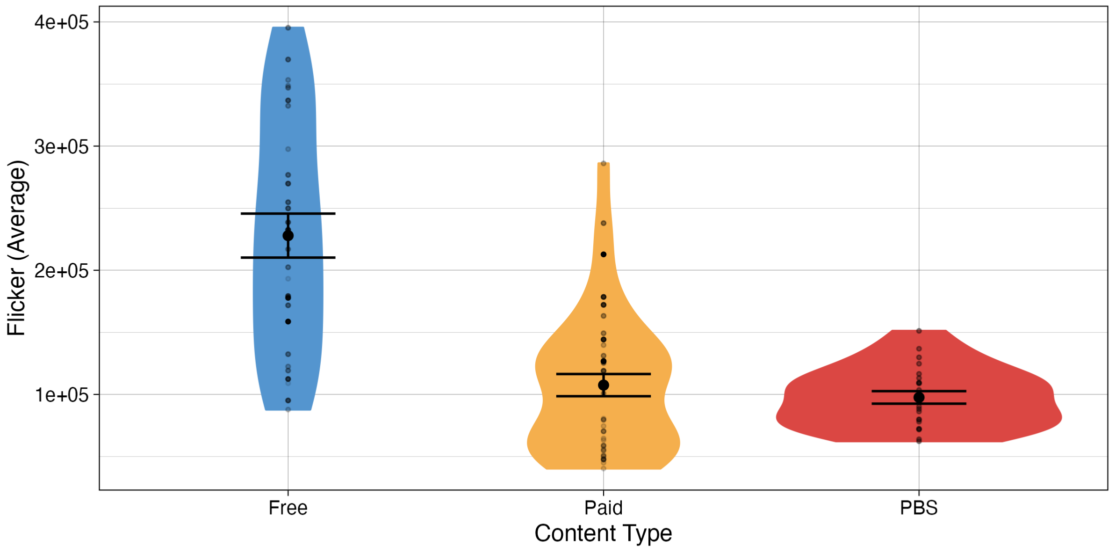


Figure 4.1: Ad-supported children's media had the highest average flicker compared to paid or public media, making ad-supported media the most visually salient. Point transparency reflects video weight (video length in frames). Error bars show standard error.

Finally, we calculate the total luminance difference across all frames and take the mean of these differences to get our final measure Mean Flicker:

$$\text{Mean Flicker} = \frac{1}{n-1}\sum_{i=2}^{n} \text{TotalDiff}_i$$

Where: $n$ is the total number of frames and $\text{TotalDiff}_i$ is the total luminance difference for frame $i$.

## 4.3 Results

### Flicker is higher in ad-supported media

Mean weighted flicker is higher in videos from ad-supported platforms compared to both paid and public platforms (see Figure 4.1). See Table 4.1 for mean flicker and mean flicker weighted by video length in frames. A linear model controlling for year of publication with weights for video length in frames confirmed that content type was a significant predictor of mean flicker (adjusted $R^2 = 0.4326, p < .001$). Compared to ad-supported videos, both paid videos ($\beta = -80, 62, p < .001$) and PBS videos ($\beta = -106, 341, p < .001$) had significantly less flicker when year of publication was stable.

### Flicker increases by year

Year of publication significantly predicted the level of flicker in videos of all content types ($\beta = 6,234, p < .005$). This increasing prevalence of high-flicker content suggests a shift in children's media consumption towards more visually salient and disruptive media (see Figure 4.2).

### Paid media contains lower flicker than ad-supported media made by the same producer

Our sample contained both ad-supported and paid versions of "Cocomelon" (available on YouTube and Netflix, respectively). The sample of ad-supported Cocomelon content has nearly four billion views in all (3,811,390,814 at time of writing), and the sample of paid content represents 14% of all Cocomelon that is currently available on Netflix.

Ad-supported episodes of Cocomelon had higher weighted mean flicker ($wM = 237971.9, wSD = 9411.574$) compared to paid episodes ($wM = 154371.1, wSD = 21025.447$). A one-way ANOVA weighted by video length in frames confirmed that the difference between Ad-supported and paid Cocomelon was statistically significant ($p < .001$). This finding underscores that content offered to lower-income viewers is more likely to disrupt learning, even when offered by the same content producer.

## 4.4 Discussion

We found that flicker, a measure of visual salience, is significantly higher in free, ad-supported children's media compared to paid and educational public programming. Our findings suggest that children from lower-income households are disproportionately exposed to more visually salient media. Children in low-income households watch more YouTube and less paid streaming media than children in high-income households, and watch more than *twice*

Table 4.1: Summary of flicker statistics by content type.

| ContentType | TotalFrames | MeanFlicker | WeightedMean | WeightedSD |
|---|---|---|---|---|
| Ad-supported | 2,770,532 | 227,921 | 219,728 | 83,581 |
| PBS | 1,990,918 | 97,615 | 98,158 | 23,669 |
| Paid | 2,820,525 | 107,505 | 132,736 | 56,618 |

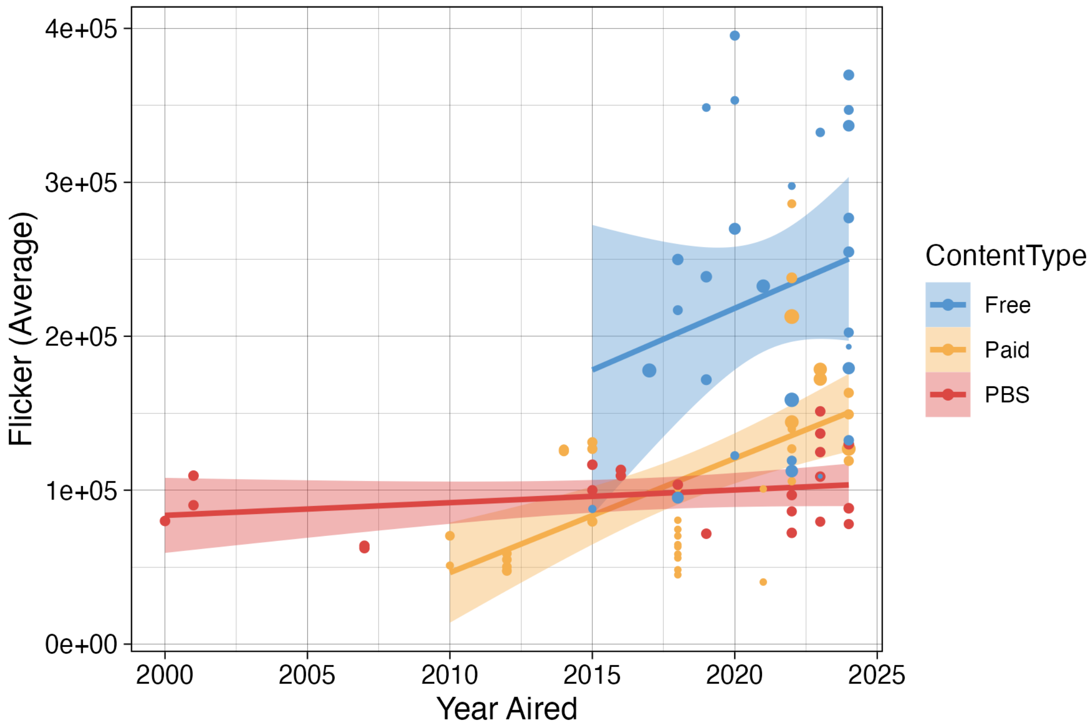


Figure 4.2: Average flicker was higher for more recent content. Point size reflects video weight (video length in frames).

the amount of screen time overall (Rideout & Robb, 2020). This places low-income children at greater risk for negative outcomes linked to highly salient media.

This is concerning because highly salient visual content disrupts young children's learning, likely by placing greater demand on their cognitive resources (Christakis et al., 2018; Essex et al., 2022; Lillard, Drell, et al., 2015; Shepherd & Kidd, 2024). This type of content is also linked to disordered attention, developmental delays, and adverse long-term impacts on cognitive and behavioral development, all of which are all linked to the overstimulating properties of high visual salience (Christakis et al., 2012; Kostyrka-Allchorne, Cooper, & Simpson, 2017; A. Lang et al., 1999). Excessive exposure to highly salient visual content should be a concern not just as a disparity in learning opportunities, but also because of the potential harm it poses to cognitive development. Based on our results, we expect that kids from lower-income families will experience greater disruptions to their learning from these materials, and are at greater risk for adverse cognitive and behavioral outcomes.

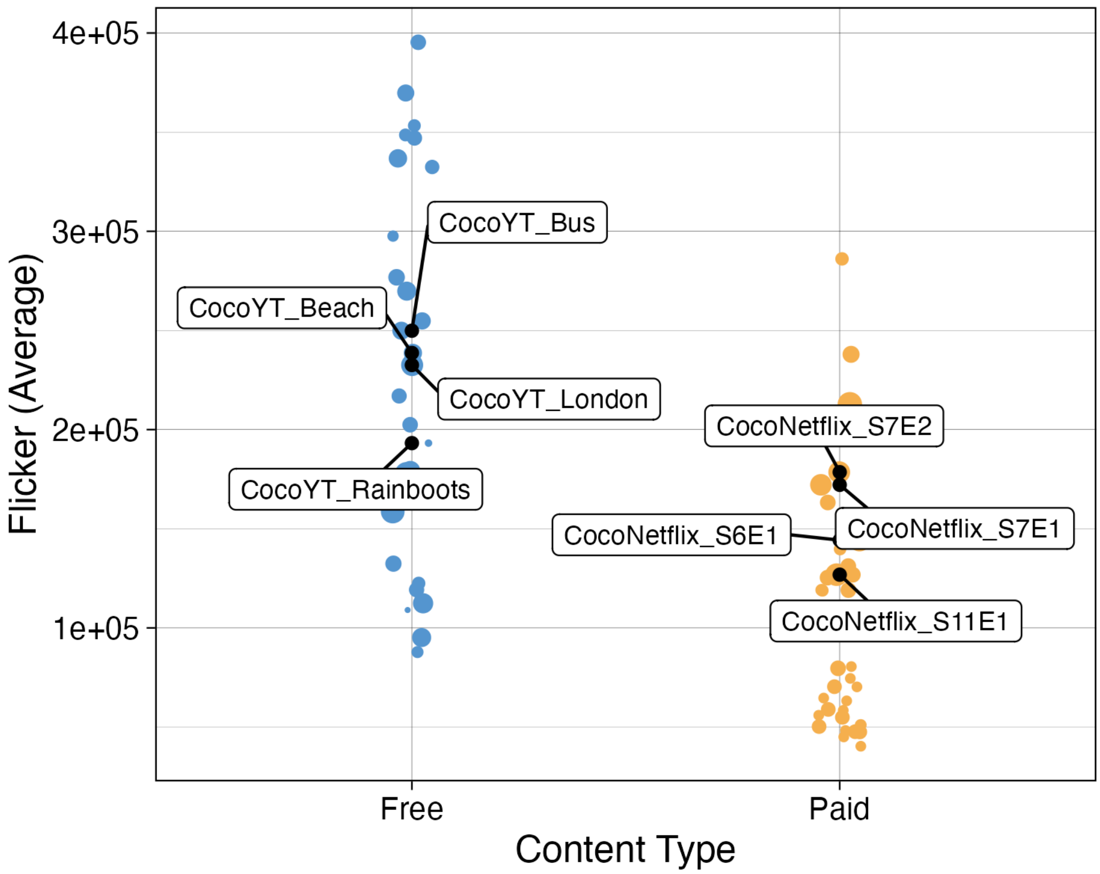


Figure 4.3: Ad-supported content made by the same media producer was more visually salient than their paid content.

## Implications for education

Children will likely habituate to the level of visual salience in the media they regularly consume. Exposure to content that is more visually salient, like the ad-supported media discussed here, may make it more difficult to engage with educational material. We found that content designed to be educational, produced for PBS in this instance, is less visually salient compared to ad-supported content, which is not necessarily designed to be educational or developmentally appropriate. This suggests that material designed to support learning contains fewer salient features. Additionally, we found that the prevalence of flicker is increasing over time. This trend highlights a challenge for the development of educational material, which takes longer to develop and is updated less frequently, making it difficult to keep up with rapidly changing media environments on platforms like YouTube. This disparity in pace of media production is increasing with the use of artificial intelligence to generate children's

YouTube content automatically.

### Supporting family decision making

Caregivers consistently report that screen-time is valuable for their child's learning. Over 70% of parents of children aged zero- to eight-years-old report learning as a key reason for allowing screen time (Rideout & Robb, 2020). However, highly salient media is designed to capture children's sustained visual attention, making it difficult for caregivers and educators to tell whether a child is meaningfully engaged or overstimulated. This creates a risk that children appear to be learning when they are actually overstimulated. By quantifying features like flicker, which are linked to disruptions in learning and adverse cognitive and behavioral outcomes, we can develop tools to empower families to make more informed decisions about their child's media consumption in an increasingly complex media landscape.

## 4.5 Conclusion

We demonstrate that platform design can perpetuate disparities in the quality of media available to children from different socioeconomic backgrounds. We find that ad-supported media, watched more by low-income children, contains more flicker—a key feature of visual salience. This finding suggests that low-income children are disproportionately exposed to the harms associated with highly salient media. This work highlights the need for regulation that ensures equitable access to high quality media.

# Chapter 5

# The Brain Rot Effect: The case for adaptive attentional bandwidth

## 5.1 Introduction

Trends in digital media favor fast paced, visually stimulating features to capture attention (Cutting, 2016; Simon et al., 1971). The emergent attention economy has given rise to extreme formats like "brain rot", where two unrelated videos are presented simultaneously. Although brain rot content might sound as though it should introduce distraction into learning, its popularity might suggest otherwise. We use brain rot to test whether attention adapts to increased perceptual load.

### Features of brain rot are viewed as harmful

Both popular discourse and academic research link short-form and high-stimulation media trends to shorter attention spans, digital addiction, diminished critical thinking, and attentional disorders (Christakis et al., 2018; Hayes, 2025; Lissak, 2018; Nguyen et al., 2025; Ward et al., 2017; T. Wu, 2016). Leading theories suggest the heightened audiovisual stimulation disrupts processing and learning (Christakis et al., 2012; Lillard & Peterson, 2011; Nguyen et al., 2025; Shepherd & Kidd, 2024).

### Then why is brain rot popular?

Brain rot was the Oxford Word of the Year in 2024, and refers to simultaneous video streams designed to capture and hold attention on social media platforms like Instagram and TikTok (Heaton, 2024). Brain rot can also refer to the assumed cognitive decline due to overconsumption of this type of content(Yousef et al., 2025). Here, we operationalize brain rot as two videos presented simultaneously via split screen, with a primary video alongside a second visually engaging but irrelevant video.

The recent appearance of study tools that convert students' notes into brain rot suggests that at least some people believe this content facilitates learning (Upton-Clark, 2024). Individuals who need greater stimulation for optimal learning could theoretically improve processing for these individuals, in line with some work on learning from accelerated playback speeds (Chen et al., 2024; D. Lang et al., 2020; Tharumalingam et al., 2025). We test the idea that brain rot may not uniformly impair processing and downstream learning for all learners, and could instead provide greater stimulation for learners who need more, thus facilitating their learning.

## Distinct sources of cognitive load

Human attention is limited; thus, the system must allocate resources to support effective perception, processing, and learning with respect to resource limitations. Cognitive load refers to the demands that a task places on these limited resources. The specific features of a task impact cognitive load. Perceptual complexity–like a busy visual scene–increases the need to focus on task-relevant information and ignore task-irrelevant distractors. On the other hand, executive control requires resources to stay on task and hold relevant information in working memory, among other executive functions(Kahneman, 1973; Reese et al., 2016; Sweller et al., 1998). Current models of processing distinguish between perceptual load and executive load.

The load theory of attention provides a framework that accounts for the distinct effects of perceptual and executive load on task performance (Lavie et al., 2004). Task performance may improve under higher perceptual load, counter to traditional accounts of cognitive load. These findings are explained by a two-stage account. Under high perceptual load, an early perceptual mechanism filters out distractors before they are processed. Under low perceptual load, distractors are processed and must be suppressed by a late-stage executive control mechanism. As a result, high perceptual load reduces distractor interference, while high executive control load increases interference, a pattern widely supported by both theoretical and empirical work (Lavie, 2005, 2010).

These differential effects demonstrate that processing is not only sensitive to the degree of cognitive load, but also to the source of load. In this view, people may seek higher perceptual load (like brain rot) to maintain a balance between perceptual and executive load. Variation in preferences for highly stimulating media, as well as mixed findings in the playback-speed literature, may be explained by variation in this optimal range of stimulation.

## Brain rot as an adaptive attentional strategy

Here, we investigate brain rot as a potential case study of how attentional capacity may be adaptive rather than fixed. Individuals may vary in the range of perceptual stimulation that optimally supports engagement and processing. Too much overstimulates and impairs processing, resulting in reduced performance, while too little under-stimulates and requires greater executive effort, also reducing performance. If true, this would predict that high perceptual load would not uniformly impair learning. Rather, the impact would depend on an individual's optimal range of stimulation.

Here, we test this account by manipulating perceptual load with brain rot stimuli during a learning task. We find that the same material disrupts processing for some participants while improving processing for others, and that the effect is predicted by their preference for brain rot. Eye-tracking data further supports our adaptive attentional bandwidth framework by differentiating perceptual and executive load.

## 5.2 Methods

### Participants

We recruited seventy-five undergraduate students ($M_{age}$ = 20.8, SD = 2.7, range = 18-35) from a university subject pool. We excluded fourteen additional participants due to technical issues (n = 9), excessive environmental noise (n = 4), or experimenter error (n = 1). We compensated participants with course credit and a $5 bonus.

### Stimuli

Primary stimuli were two excerpts of educational physics videos lasting 2 minutes and 15 seconds.[1] The physics videos were presented on the bottom half of a split-screen visual display (see Figure 5.1). The top half consisted of either a static gray background or an unrelated, visually engaging gameplay video[2], according to a counterbalanced, within-subject trial condition (control vs. brain rot). We paired each video with a 30-item multiple-choice test assessing comprehension of the material.

### Procedure

Participants completed the study in a quiet, darkened room to control for distractions. We presented the visual stimuli on a monitor connected to a desktop-mounted EyeLink 1000 eye tracker (SR Research Ltd. 2005). The EyeLink monocularly recorded gaze location and pupil diameter at 1000Hz. We positioned participants 70 cm from the display. We calibrated the eye tracker using a standard nine-point calibration procedure followed by an identical validation sequence before each trial. We repeated the validation sequences until satisfactory, ¡ 0.5 degrees of average error.

Participants watched a video and completed a corresponding comprehension test in each of two within-subject blocks. Participants were required to view each video in its entirety at least once, but were allowed to re-watch the video continuously until they felt ready to

[1] Adapted from 'Windmills Are NOT Like Dams' and 'How Modern Light Bulbs Work' by @minutephysics on YouTube.

[2] Adapted from Subway Surfers Gameplay [4K 16x9] No Copyright by @subwaysurfers on YouTube.

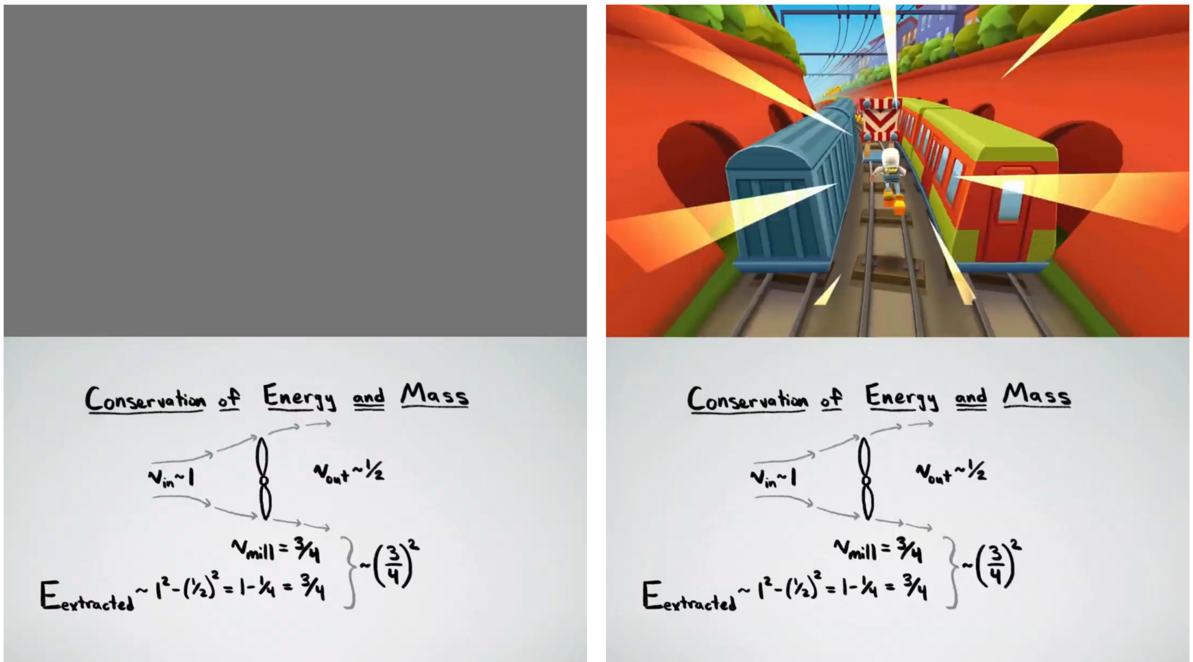


Figure 5.1: Screenshots of matched videos for control condition (left) and brain rot condition (right).

complete the test. After the initial viewing, the video looped automatically until participants looked away from the screen for more than one second. Following each video, participants completed a 30-item test assessing comprehension of the video content. Following both blocks, participants gave their preferences for the two video presentations and answered questions about their media consumption habits. We counterbalanced video content across participants and randomized trial condition across trials.

## Brain rot preference score

We created a composite score of brain rot preference from five equally weighted, binary questions. We asked participants which of the two videos they found more engaging, more enjoyable, more effective for learning, easier for sustaining their attention, and more preferred overall. Higher composite scores indicate a stronger preference for the brain rot video presentation over the control presentation.

## Media habit score

We created a media habit score to quantify participants' prior consumption of online media. This composite score was composed of 9 range-normalized and equally weighted questions, including 1 binary and 8 ordinal items. These items asked participants about their screen time, time spent watching short-form content, frequency of watching sped-up content, frequency of watching brain rot content, frequency of consuming multiple streams of content simultaneously, frequency of consuming unrelated media while working or studying, attitude toward engagement with brain rot content, experience using brain rot study tools, and interest in using brain rot study tools. Higher composite scores correspond to heavier online media consumption, stronger proclivities for highly stimulating content, and stronger interest in educational applications of brain rot.

## Eye-tracking data pre-processing

### Pupil dilation

We used pupil size as a proxy measurement of executive load. To quantify pupil size, we removed 150ms before and after a blink event to account for blink-induced pupillary response (Bentivoglio et al., 1997). We then normalized pupil size for each participant using the first 10 seconds of the current trial to accounted for different individual baselines for each trial type (brain rot and control). Finally, we scaled pupil change from baseline for clearer interpretation.

### Eye blink behavior

Endogenous eye blink behavior reflects the regulation of sensory and information processing (for review see Callara et al., 2023; Stern et al., 1984). We analyzed blink behavior to test the effect of increased perceptual load in the brain rot condition as a function of participants' preference scores using two metrics: total blink time and blink count.

Total blink time is the total number of eye-tracking samples occurring during a blink event. Since blink event samples and missing data cannot easily be differentiated, all blink events longer than one second were removed, a standard threshold for identifying endogenous eye blinks (Stern et al., 1984). Participants used a chin rest, so any blink events longer than one-second are likely due to drowsiness rather than missing data. This data filtering resulted in the removal of about 0.6% of all data. The final blink sample count was summed and divided by remaining samples. The final measure is total blink time per sample for a given trial. Blink count was also quantified as distinct blink events. Both total blink time and blink count were log-transformed prior to analysis to meet model assumptions (Bentivoglio et al., 1997).

## 5.3 Results

### Brain rot has heterogeneous effects on learning and metacognition

We assessed whether the simultaneous, brain rot-style presentation of unrelated video game content during viewing of physics videos affected performance on a subsequent comprehension test. We ran a generalized linear mixed-effects model predicting trial-level accuracy, with condition (control vs. brain rot) and condition order as fixed effects and random slopes and intercepts per participant. This model revealed a significant effect of condition order ($\beta$ = 0.29, $t$ = 3.35, $p$ ¡ .001) such that performance was better in the second within-subjects condition. There was no significant effect of condition ($\beta$ = -0.14, $t$ = -1.56, $p$ = .12). There was no group-level effect of brain rot on test performance. However, the random slope variance for condition (Var = 0.33) indicated substantial heterogeneity in the effect of brain rot. Consistent with this variation, the mean absolute difference in by-participant average accuracy between conditions was 15.5% (SD = 0.12, range = 0-0.5).

We also tested for condition differences in participants' metacognitive sensitivity, or the relationship between certainty and accuracy. We quantified metacognitive sensitivity using the pROC package in R to calculate the area under the Type 2 ROC curve, or AUC2. This curve captures the hit rate vs. false alarm rate for whether higher trial-level certainty predicts accuracy. An area of 0.5 corresponds to at-chance performance, and higher AUC2 values indicate stronger metacognitive sensitivity.

A linear mixed-effects model predicting AUC2, with condition (control vs. brain rot) and condition order as fixed effects and random intercepts per participant, revealed no significant effect of condition ($\beta$ = 0.03, $t$ = 1.22, $p$ = .23) or of condition order ($\beta$ = 0.03, $t$ = 1.50, $p$ = .14). The null effect of condition reflects bidirectional effects on the individual level. The mean by-participant difference in AUC2 between conditions was 0.15 (SD = 0.13, range = 0.003-0.6).

### Brain rot preferers learn better from brain rot, worse from traditional content

We hypothesized that participants' reported preferences for the brain rot-style video presentation or the control presentation may capture the variation in task performance across conditions. Brain rot preference scores were right-skewed (M = 1.49 out of 5, SD = 1.79). Nearly half (n = 35, or 46.7%) of the participants had a score of zero, meaning that they chose the control presentation without brain rot content as more engaging, more learnable, and generally preferred across all five questionnaire items. Eight (or 10.7%) participants had a brain rot preference score of 5, reflecting a universal preference for the brain rot presentation.

We assessed whether participants' content preferences predicted different learning outcomes, and found a crossover effect (see Figure 5.2). Participants with a stronger brain rot preference performed better in the brain rot condition, whereas participants with a weaker

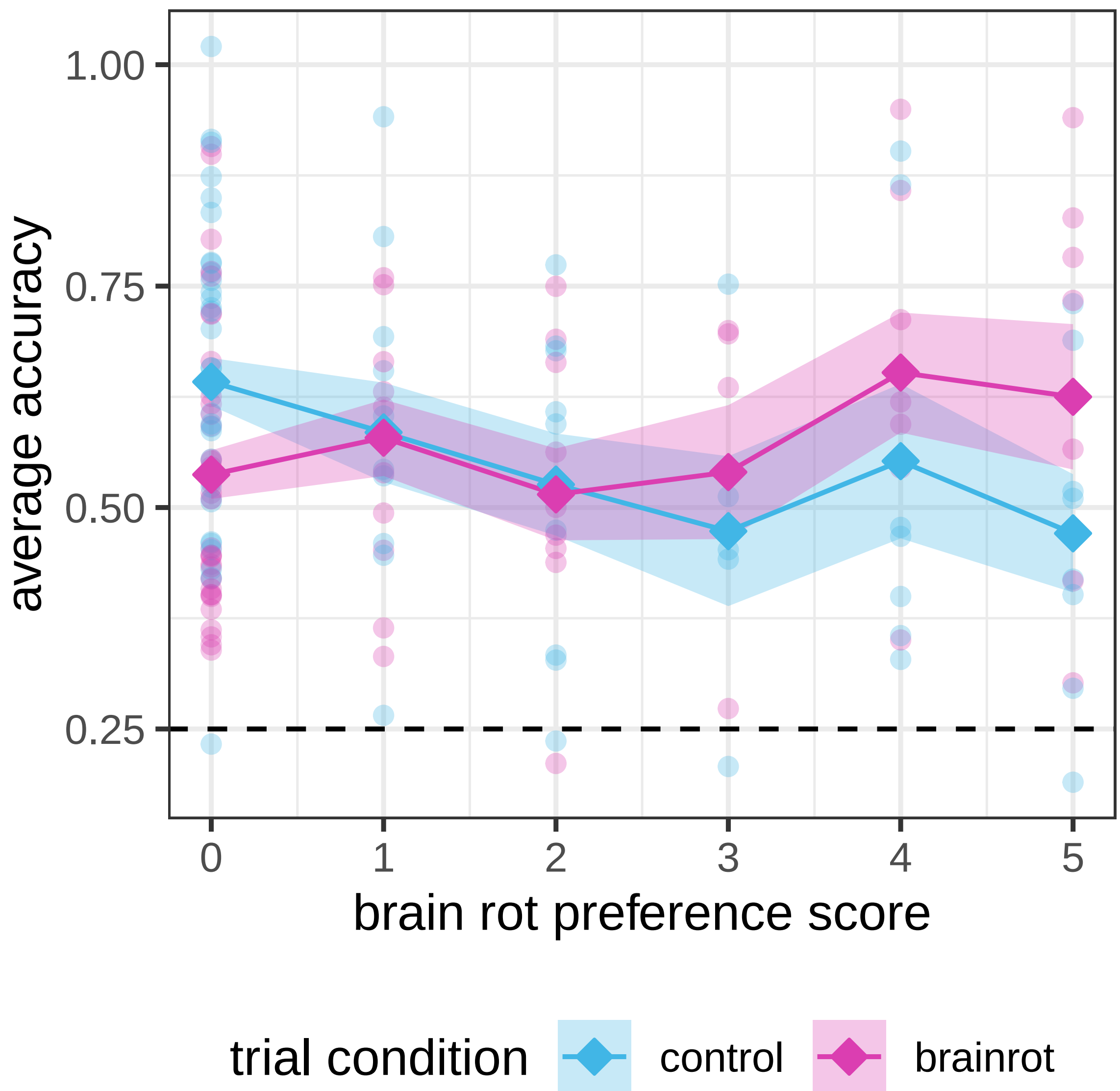


Figure 5.2: Preference for brain rot content predicts differential comprehension test performance across conditions. Dashed line represents chance performance. Points are participant averages by condition and are jittered vertically. Diamonds are condition means by preference score. Ribbons represent SE.

brain rot preference performed better in the control condition. We fit a linear mixed-effects model predicting participants' average accuracy, with condition (control vs. brain rot), scaled brain rot preference score, and condition order as fixed effects and random intercepts per participant. This model revealed a significant effect of condition order ($\beta$ = 0.04, $t$ = 2.08, $p$ = .04), a significant effect of scaled brain rot preference score ($\beta$ = -0.05, $t$ = 2.55, $p$ = .01), and a significant interaction between condition and scaled brain rot preference score ($\beta$ = 0.08, $t$ = 3.59, $p$ ¡ .001). The effect of the brain rot condition was not significant ($\beta$ = -0.02, $t$ = -1.22, $p$ = .22). Higher brain rot preference scores predicted lower test performance in the control condition, but higher test performance in the brain rot condition.

## Brain rot preferers have poor metacognition for traditional content

Stronger brain rot preference predicted worse metacognitive sensitivity in the control condition and better metacognitive sensitivity in the brain rot condition (see Figure 5.3). We fit a linear mixed-effects model predicting AUC2, with condition (control vs. brain rot), scaled brain rot preference score, scaled average accuracy, and condition order as fixed effects and random intercepts per participant. The model revealed a significant effect of brain rot preference score ($\beta$ = -0.04, $t$ = -3.02, $p$ = .003), a significant effect of scaled average accuracy ($\beta$ = 0.06, $t$ = 6.42, $p$ ¡ .001), and a significant interaction between condition and brain rot preference score ($\beta$ = 0.06, $t$ = 2.97, $p$ = .004). Better task performance was associated with higher AUC2 scores, as is common with many measures of metacognition (Rahnev, 2025). Critically, the interaction between condition and brain rot preference remained after controlling for task performance. Higher brain rot preference was associated with enhanced metacognition in the brain rot condition, but diminished metacognition when learning from the less stimulating control presentation.

## Brain rot preference is not captured by gaze patterns

Looking behavior was analyzed within the brain rot condition alone since looking to the video was at ceiling during the control condition. A linear regression predicting the proportion of time spent looking to the content video during brain rot trials (as opposed to the brain rot video) revealed no effect of brain rot preference score ($\beta$ = -0.02, $t$ = -1.16, $p$ = 0.25) or condition order ($\beta$ = 0.03, $t$ = 0.85, $p$ = 0.4). A linear regression predicting the number of switches made between the brain rot and content videos similarly showed no significant effect of brain rot preference score ($\beta$ = 6.4, $t$ = 1.16, $p$ = 0.25) or condition order ($\beta$ = -19.24, $t$ = -1.7, $p$ = 0.09). These results indicate that preference for brain rot does not predict differences in looking behavior in our task.

Lastly, the duration of watching did not differ between conditions or preference scores. A linear mixed-effects model predicting total viewing time confirmed no effects of condition ($\beta$ = .018, $t$ = 0.12, $p$ = 0.9) , brain rot preference score ($\beta$ = 0.07, $t$ = 0.58, $p$ = 0.57), or condition order ($\beta$ = -0.2, $t$ = -1.26, $p$ = 0.21), and no interaction between trial condition

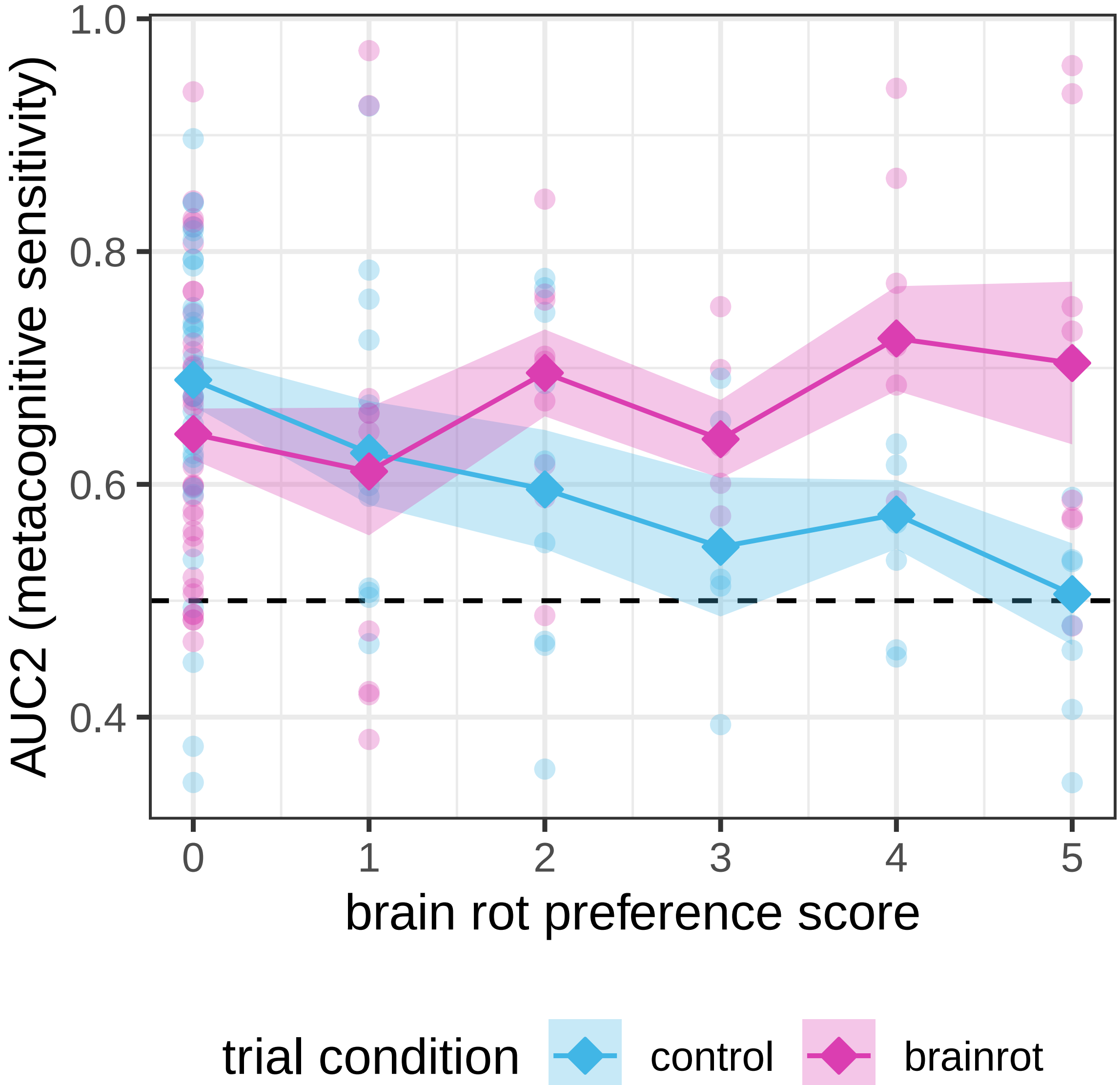


Figure 5.3: Preference for brain rot predicts differential metacognitive sensitivity across conditions. Dashed line represents at-chance metacognition. Points are participant averages by condition and are jittered vertically. Diamonds are condition means by preference score. Ribbons represent SE.

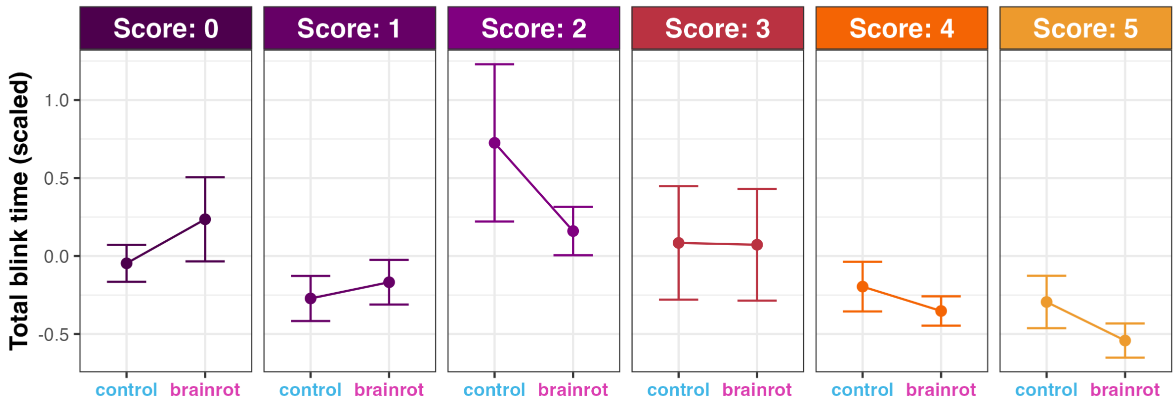


Figure 5.4: Preference for brain rot predicts differential blink behavior across conditions. Points are average blink time by condition. Error bars represent SE.

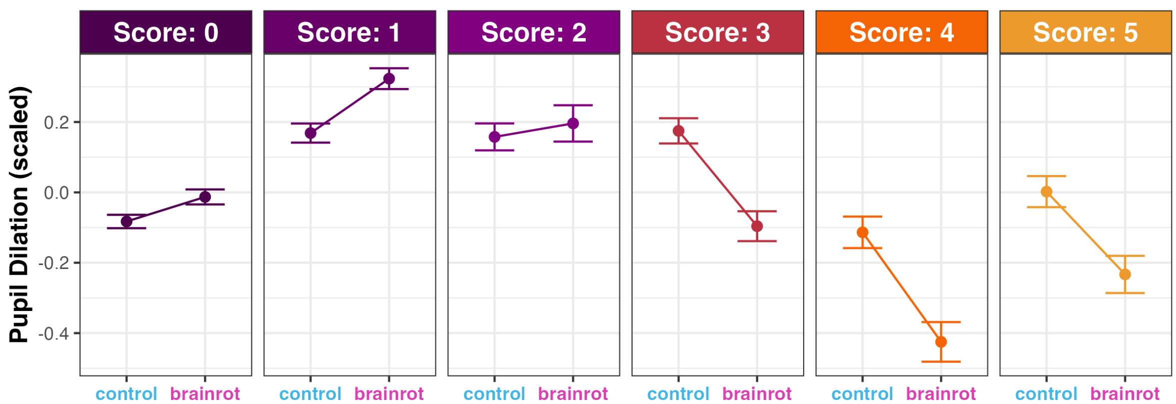


Figure 5.5: Preference for brain rot predicts differential pupil behavior across conditions. Points are average change in pupil size from baseline by condition. Error bars represent SE.

and brain rot preference score ($\beta = 0.05$, $t = 0.37$, $p = 0.71$). Since preference score does not predict watching duration in the brain rot condition (M = 2.46, SD = 1.08) or the control condition (M = 2.51, SD = 1.1) we restrict our analysis of eye blink and pupil dilation to the first loop of each video.

## Eye blink behavior supports variation in experience of perceptual load

A linear mixed-effects model predicting the total blink time (log-transformed blink samples) with condition, scaled brain rot preference score, and condition order as fixed effects and random intercepts per participant, revealed a significant interaction between condition and preference score, ($\beta$ = -0.24, $t$ = -2.31, $p$ = .024). This suggests that only individuals with lower preference for brain rot experience greater perceptual load while watching brain rot, reflected by their increased total blink time. The effect of preference score was significant in the brain rot condition ($\beta$ = -0.28, $t$ = -2.44, $p$ = .02), but not in the control condition ($\beta$ = 0.04, $t$ = -0.35, $p$ = 0.73). Total blink time is reduced among brain rot preferers in the brain rot condition only.

We also analyzed distinct eye blink events as opposed to total eye blink durations. A linear mixed-effects model predicting log-transformed blink count confirmed that higher preference results in fewer blinks in the brain rot condition ($\beta$ = -0.23, $t$ = -2.19, $p$ = .03). In the control condition, preference did not significantly predict blink count ($\beta$ = -0.11, $t$ = -1.08, $p$ = .28). The interaction between preference and trial condition was not significant for distinct blink events ($\beta$ = -0.12, $t$ = -1.47, $p$ = .15). These effects remain for both blink measures after controlling for task performance and metacognitive sensitivity.

## Pupil dilation suggests increased executive load during traditional content for brain rot preferers

We used a linear mixed-effects model to analyze the effects of condition and scaled brain rot preference score on participants' normalized pupil size, controlling for condition order with a random intercept for participants. The main effects of condition ($\beta$ = -3.66, $t$ = -1.04, $p$ = 0.29) and of brain rot preference score ($\beta$ = 7.77, $t$ = 0.45, $p$ = 0.66) were not significant. Condition order significantly predicted pupil dilation ($\beta$ = 21.64, $t$ = 5.14, $p$ ¡ 0.001). A significant interaction between brain rot preference score and condition ($\beta$ = -40.65, $t$ = -9.85, $p$ ¡ 0.001) revealed that a stronger preference for brain rot is associated with decreased pupil size while watching brain rot. These results are consistent with decades of empirical work demonstrating that pupil size increases with general cognitive load, as well as more recent evidence that pupil dilation specifically reflects increased task-related executive control demands (Gilzenrat et al., 2010; Kahneman, 1973). In line with this interpretation, increased pupil dilation in the control condition among brain rot preferers indicates increased executive demand. This suggests that attending to the control condition was more effortful for brain rot preferers.

## Brain rot preferers consume more media

We assessed whether participants' digital media habits predict their preferences between the video presentations in our task. Media habit scores were approximately normally distributed

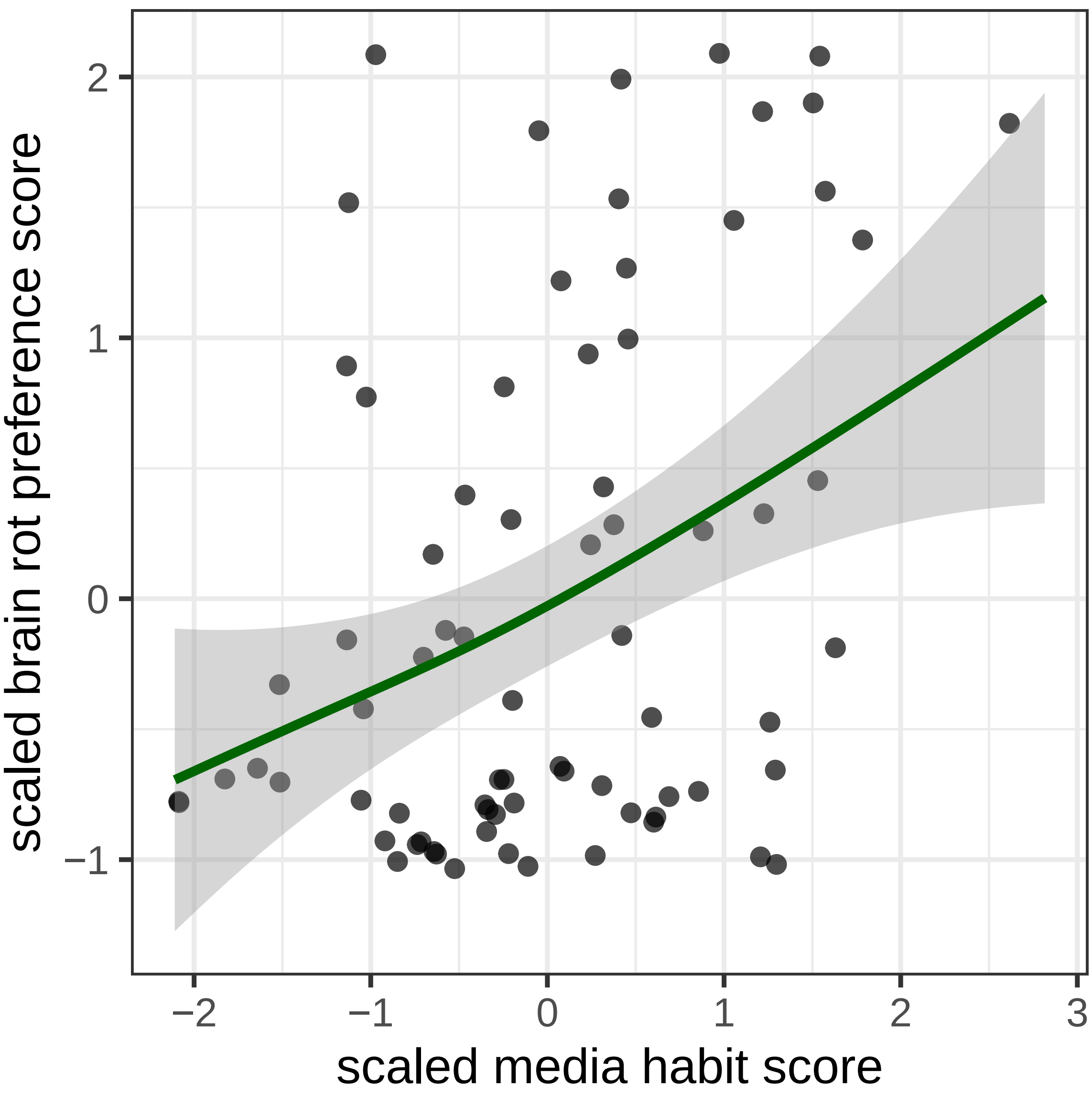


Figure 5.6: Preference for brain rot is correlated with media habits. Individuals who prefer brain rot consume more, and more stimulating media.

(M = 3.83 out of 9, SD = 1.39, range = 0.90-7.75). Media habit score was positively correlated with brain rot preference score (Spearman's $\rho = 0.35$, $p = .002$). Media habits characterized by heavier online media consumption, higher tolerance for attentional load, and increased interest in or use of brain rot study tools were associated with a higher likelihood of preferring to learn from the brain rot video presentation in our task. Notably, media habit scores did not independently predict task performance or metacognition in subsequent analyses.

## 5.4 Discussion

The incentives of digital media environments are rapidly changing the nature of the online content that we consume. Attention-grabbing "brain rot" content exemplifies this change. The term brain rot itself codifies a set of lay hypotheses, largely untested, about the potential (harmful) effects of brain rot on attention and learning. Here, we show that counter to widespread assumptions, brain rot-style presentations of otherwise identical content led to improved comprehension and metacognitive sensitivity for a meaningful subset of learners. These results are in line with hypotheses that suggest that individuals learn best from stimuli that meet their specific optimal loads.

Our data on individual preferences for learning from brain rot suggest that participants are at least somewhat well calibrated in understanding the attentional demands that are optimal for their own learning. In this sense, our composite score of brain rot preference may be seen as a measure of global metacognition for brain rot. Global metacognition represents broad, more abstract self-beliefs about performance. This contrasts with local metacognition, which refers to momentary confidence in trial-by-trial performance and has overlapping but distinct neurocognitive underpinnings (Rouault & Fleming, 2020). Awareness of one's ability to learn is a double-edged sword. The growing preference for and use of brain rot even in educational settings may be less concerning if the primary users genuinely benefit. At the same time, a perceived dependence on high stimulation for learning may lead to disengagement from traditional educational content.

### Attentional signatures of brain rot learners

Eye-tracking findings highlight two key attentional signatures associated with improved learning from brain rot. Importantly, these signatures are not a function of visual search. Neither proportion of looking time nor frequency of switches to the brain rot predicted whether participants learned better from brain rot. Instead, these performance patterns co-vary with two key differences in visual processing mapping onto the two-stage load theory of attention (Lavie et al., 2004). Blink responses indicate divergence in perceptual load, while pupil dilation patterns point to variation in executive load.

First, people who learn better from brain rot are less overstimulated by brain rot. More extreme blink responses correspond to higher perceptual load, or overstimulation, and brain rot learners blinked less in the brain rot condition (Bentivoglio et al., 1997). By contrast, people who learn better in the control condition demonstrate stronger signs of overstimulation in the brain rot condition, as expected (see Figure 5.4). Brain rot learners' apparent immunity to this overstimulation is consistent with an attentional system adapted to a higher objective ceiling of perceptual stimulation. Crucially, this variation exists despite the two groups viewing identical content, indicating that perceptual load is not objectively determined by stimulus characteristics alone. Our findings suggest that observer factors also affect perceptual load.

Second, people who learn better from brain rot exhibit higher executive load when watching content *without* brain rot. Pupil dilation indexes executive load (Beatty, 1982; Gilzenrat et al., 2010). We observed a crossover effect (see Figure 5.5) in which participants tended to show higher pupil dilation when watching the video from which they learned less. Brain rot learners showed higher pupil dilation in the control condition, reflecting increased expenditure of executive control to maintain attention on the control video. Speculatively, this control video was too boring for brain rot learners, i.e., outside of their optimal range of stimulation. This is striking given that the videos in the control condition were still very stimulating, with constantly changing visuals corresponding to the concepts being taught. Thus, our task may underestimate the extent of variation in attentional bandwidth within our sample. Simpler content would likely produce stronger effects than we observed here.

Taken together, these findings lend preliminary support to an adaptive attentional bandwidth account. The range of stimulation or load that is optimal for learning is not fixed, but varies across individuals. Individual adaptation shifts, but does not expand, this optimal window. This property is reflected in the performance tradeoffs we observe in our task, where adaptation to perceptually demanding content is associated with poorer outcomes when learning from less stimulating content.

## Connection to real-world digital media habits

Preliminary evidence suggests that real-world digital media habits relate to preference for learning from brain rot. Specifically, we found that preference for learning from brain rot was associated with a composite score indicative of higher online media consumption, tolerance for high stimulation, and intentions to use brain rot study tools. This correlation does not establish causality, although we cautiously speculate that the relationship is complex and bidirectional. It is likely that increased short-form media consumption breeds familiarity with, and liking of, brain rot content, which in turn reinforce digital habits. Other more stable individual differences may also affect people's propensity to engage with these forms of content. Future work should use longitudinal and intervention designs to clarify these relationships.

It is worth noting that our sample is limited to undergraduate students whose media habits are unrepresentative of the broader population. We suspect, for example, that other age groups may differ, both in the proportion of users whose learning would benefit from highly stimulating content and in the malleability of their attentional systems.

Public discourse increasingly emphasizes "resetting" attention and building tolerance to boredom through incremental exposure to "boring" tasks, like reading or even a viral TikTok challenge to do nothing (Nazaryan, 2026; Smith, 2023). These trends suggest that attention can be adapted in the opposite direction than observed in our brain rot learners, lowering the baseline stimulation needed to maintain engagement. Beyond individual practices, this discourse is part of a broader "attention activism" movement, including initiatives of groups like the School of Radical Attention, where disengagement with content is reframed as a form of resistance to algorithmically-curated and feeds and content competing for attention (Burnett & Mitchell, 2025; The Friends of Attention, 2026). The adaptive account that we provide evidence for here would suggest that these movements and practices could be effective.

## 5.5 Conclusion

We proposed an adaptive account of attentional bandwidth in which processing is best facilitated within an individual's optimal range of stimulation. We found evidence in line with this account using the brain rot trend as a test case. In general, our results indicate that an individual's optimal range of stimulation may shift, perhaps as a result of a combination of media habits and more stable individual differences.

# Chapter 6

# Information value does not guide attention uniformly across mammals

## 6.1 Introduction

Popular theories of curiosity suggest that uncertainty guides attention and exploration to support learning (Berlyne, 1960; Dubey & Griffiths, 2020; Kidd & Hayden, 2015; Loewenstein, 1994). Empirical evidence supports these theories, demonstrating that attention is organized in ways that support curiosity-driven learning, meaning that information-seeking behavior is guided by material that is expected to reduce the gap between what is known and what is unknown, relative to a learner's current set of mental representations (Berlyne, 1954; Dember & Earl, 1957; Hunter & Ames, 1988; Kidd et al., 2012; Oudeyer & Smith, 2016).

Work with human infants, young children, and rhesus macaques demonstrates a U-shaped relationship between uncertainty and attention (Cubit et al., 2021; Kidd et al., 2012; T. Wu, 2016). This U-shape supports the theory that both humans and nonhuman primates preferentially attend to information that will best support learning. Information that is too simple offers little learning value, and information that is too complex relative to what a learner knows is too difficult to process. These early-emerging patterns of attention are widely appreciated as foundational for cognitive development, supporting skills from language acquisition to causal learning (Baldwin, 1991; Oudeyer & Smith, 2016; Schulz & Bonawitz, 2007). Here, we test how generalized this organizing principle of attention is by testing a more distantly related species, dogs.

Dogs' attentional preferences are not well-studied. The orienting response has long been studied in comparative and developmental psychology as a foundation of the attention system across many mammals (e.g., Lynn, 2013; N, 1960). For example, influential early work established that dogs attend to novel events in their environment, termed the 'investigatory reflex' by Pavlov (Pavlov (1927), 2010). This pattern of attention is also observed in rodents and humans (Berlyne, 1950, 1958). This bottom-up response to novel stimuli in an

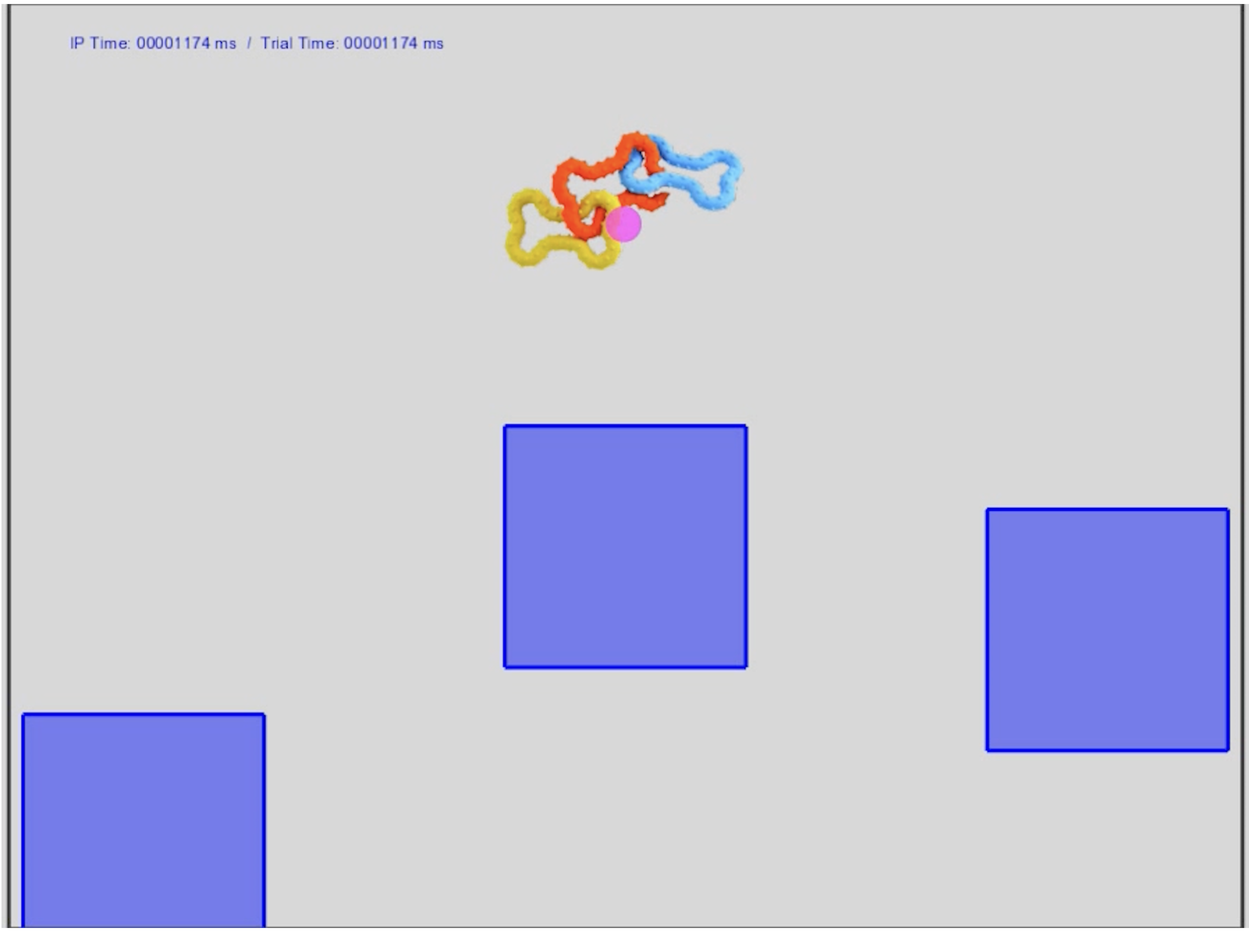


Figure 6.1: Example of stimuli. Gaze location is shown with a pink dot.

environment is important for detecting possible threats or other visually salient stimuli. We know that dogs share at least some features of attention with humans beyond this simple orienting response. For example, dogs and infants both look longer to events that violate expectations about physical object properties, like an object disappearing behind an occluder and not reappearing (Stahl & Feigenson, 2015; Völter et al., 2023). However, it remains an open question whether the attentional mechanisms that support curiosity-driven learning are unique to humans and nonhuman primates.

We employ a variation of the primate and infant paradigms with trained research dogs to test whether this adaptive regulation of information-seeking behavior is common across distantly related mammals. We demonstrate that dogs' visual attention is different from humans and nonhuman primates, showing a preference for events of high surprisal rather than intermediate surprisal.

## 6.2 Methods

We conducted a screen-based eye-tracking study to investigate the dogs' gaze behavior in response to video stimuli. Data was collected using an EyeLink 1000 eye-tracker (SR Research, Canada) in a controlled environment at the Clever Dog Lab. All data collection was conducted by certified dog trainers experienced in running eye-tracking experiments with dogs.

During a single trial, dogs were presented with a sequence of 28 events. A single event consisted of an image of an object (e.g., a dog toy or other inanimate dog-related items like beds and leashes) popping out of one of three identical blue boxes on screen, then returning to its original location behind the box. Images were randomized, and each sequence presented three novel items, so an individual dog never saw an item repeat. The location of each box on the screen was randomized for each sequence, such that there was a box within each horizontal third of the screen. Each event was 2.15 seconds in duration: one second to pop out of the box and one second to return to the box, with a 150 ms delay between each event. Each sequence varied in the probability of the three possible events occurring (items popping out of their respective boxes). These unique sequences thus vary in event complexity, where simple sequences contain many highly predictable events (e.g., A, A, A, A, . . . ) and complex sequences contain more unpredictable ones (e.g., A, C, A, C, B, . . . ). Each trial presented one sequence in full, comprised of 28 individual events. All dogs were presented with 18 unique sequences in a random order across test sessions. The ideal information rate for dogs is not known, so we tested a range of information rates based on the results from both human and primate versions of this task (Cubit et al., 2021; Kidd et al., 2012; S. Wu et al., 2022).

All dogs will see the same 18 sequences, each made up of 28 events. Each full sequence represents one trial. The 18 sequences were presented in a random order for each dog. Dogs were presented with 1-3 full sequences trials per data collection session. Based on a previous study that used the same design and stimuli with human infants we tested 30 individuals. Due to the limited dog research in this field, a power analysis was not possible.

### Eye-tracking measures

Arrival time of gaze to the events as they occur and dwell time to events were analyzed. Arrival time was calculated as the latency to look to an event after an object appeared from behind an occluder. As such, arrival time may also be a negative value, determining whether a participant is looking to the area where an event will occur before the object appears from behind an occluder. Thus, arrival time is effectively a measure of reaction time that also considers predictive looks. Dwell time was caluclated as the time spent looking at an object after it appeared (i.e., during the 2.15 s of each event). Sequences in which participants looked for fewer than four events were discarded; this resulted in the exclusion of only one total event sequence.

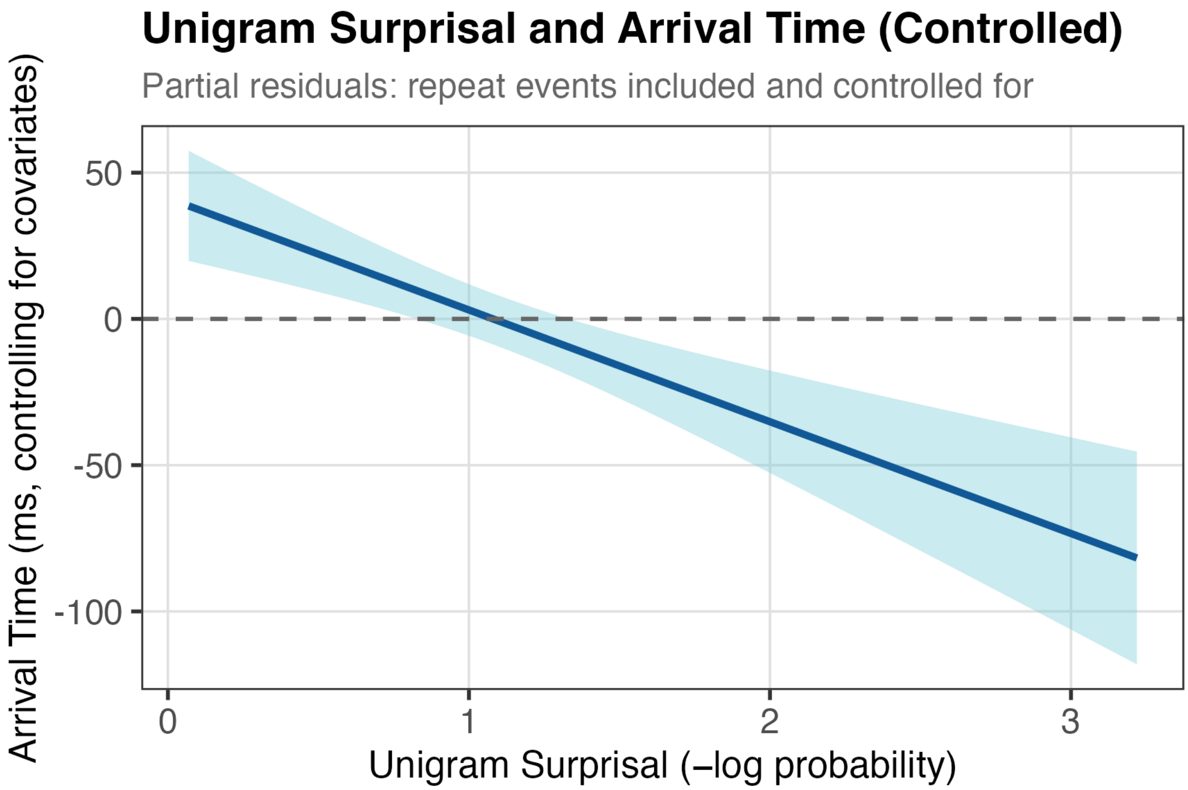


Figure 6.2: Dogs look the fastest to events of high surprisal, indicating a preference for higher surprisal. The trend line represents the effect of unigram surprisal on arrival after statistically controlling for repeat events, sequence position, spatial distance between events, and first appearances. The line shows the model-estimated linear trend and the shaded error shows the 95% confidence interval of that trend, fitted through the covariate-adjusted residuals of all 11,948 observations.

## 6.3 Results

### Surprisal model

We will analyze looking behavior as a function of the surprisal value of each individual event in a sequence, as done in previous work with infants, young children, and nonhuman primates (Cubit et al., 2021; Kidd et al., 2012; S. Wu et al., 2022). Following these previous analyses, we will calculate surprisal value as the negative log probability of each event's occurrence, this will be determined by both unigram (independent) and transitional (dependent on event order) Markov Dirichlet-multinomial (ideal observer) models (MDMD). The unigram surprisal model considers each event to be statistically independent. The transitional model

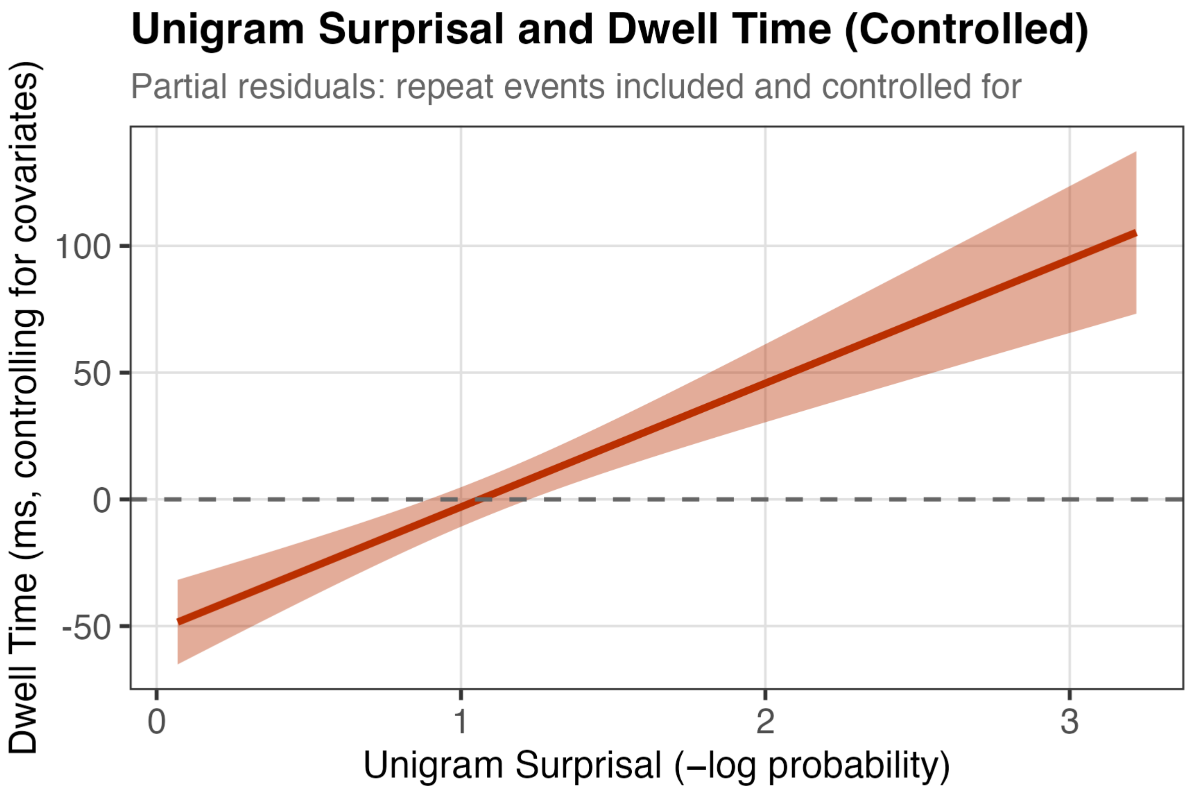


Figure 6.3: Dogs look the longest to events of high surprisal, indicating a preference for higher surprisal. The trend line represents the effect of unigram surprisal on dwell time after statistically controlling for repeat events, sequence position, spatial distance between events, and first appearances. The line shows the model-estimated linear trend and the shaded error shows the 95% confidence interval of that trend, fitted through the covariate-adjusted residuals of all 11,948 observations.

does *not* assume independence, and instead tracks conditional probability based on the event immediately preceding the current event. Both models begin with uninformative uniform priors, as a learner likely would before making any observations. The model then estimates the surprisal value of the current event for each event in a given sequence. The simple prior is combined with the learner's previous observations from the sequence to form a posterior belief. Thus, the next event conveys the surprisal value according to the probabilistic expectations of the learner's updated belief. This method to calculate surprisal and the sequences used in the present study are adapted directly from the infant and Rhesus Macaque versions of the task (Kidd et al., 2012; S. Wu et al., 2022).

## Pre-registered Goldilocks analysis

The pre-registered model specified linear and quadratic surprisal terms to test for a U-shaped Goldilocks relationship between surprisal and two measures of attention, arrival time and dwell time, following the analysis done with infants, young children, and Rhesus Macaques (Cubit et al., 2021; Kidd et al., 2012; S. Wu et al., 2022). Likelihood ratio tests comparing the pre-registered quadratic models with linear models revealed that the quadratic term did not significantly improve fit for surprisal (unigram or transitional) in the dwell time analysis. For unigram surprisal, the quadratic term showed a trend in the predicted Goldilocks direction in arrival time ($\beta = 960.904, p = 0.078$), but did not reach statistical significance. The quadratic term was non-significant for arrival time in the transitional model ($\beta = -32.12, p = 0.95$).

A likelihood ratio test comparing the pre-registered quadratic model to the linear-only model indicated that the quadratic term did not significantly improve fit ($\chi^2 = 3.10, p = .08$). Analysis of AIC and BIC further confirmed that the quadratic term did not meaningfully improve model fit; BIC favored the linear model ($\Delta$ AIC = 1, $\Delta$ BIC = 6). Based on these comparisons, the linear models are reported as our primary analysis. Pre-registered analyses are included for transparency. The non-significant quadratic term across all models indicates that dog attention may be organized by different principles than human and primate attention.

The significant negative linear trend across all models, in which dogs arrive earlier and look longer to more surprising events, is consistent with a novelty or uncertainty preference. We discuss this pattern in relation to prior findings with humans and primates and the role of uncertainty for guiding attention across species more broadly.

## Dogs are faster to look to high suprisal events

Dogs were faster to look to events of high surprisal. The relationship between both unigram and bigram surprisal is a negative linear relationship, confirmed by an ANOVA comparing the pre-registered GAM and linear mixed-effects models.

Specifically, we will analyze the looking behavior using a mixed-effect linear model with random intercepts as done by Wu and colleagues, using standardized linear and quadratic surprisal terms as predictors. We will include the random effect of Dog ID as well as random slopes to account for repeated observations. We will control for covariates, such as whether an event is the first time an object is shown, trial number, and the distance between two consecutive events. We will also use a generalized additive model (GAM) to visualize the relationship between the surprisal values generated by the MDMD and the looking behavior. We will do this for all participants, both as a group and individually.

All mixed-effects models include random effects of Dog ID and control for the following covariates: whether an event is the first time an object is shown, trial number, the distance between two consecutive events, and whether an event is a repeat of the previous event.

### Negative linear relationship between unigram surprisal and arrival time

Dogs were quickest to look to events of higher surprisal (see Figure 6.2). We fit a linear mixed-effects model predicting gaze arrival time to each event with standardized surprisal of each event (negative log probability), whether the event was the first appearance of an object in that sequence (binary, 1 or 0), scaled distance from the event immediately prior to the current event, and whether the current event was a repeat of the previous event (binary, 1 or 0). The model revealed a significant effect of unigram surprisal on gaze arrival time ($\beta = -21.81, p < .001$), a significant effect of distance from the previous event ($\beta = 27.76, p =< .001$), and whether an event is a repeat of the previous event ($\beta = -342.85, p =< .001$). There is no significant effect of whether an event is the first appearance of an object ($\beta = 21.15, p = .31$), and a potentially marginal effect of the event number within a sequence ($\beta = 1.58, p = .013$).

### Negative linear relationship between transitional surprisal and arrival time

Dogs were also quickest to look to events of higher transitional or bigram surprisal. We fit a linear mixed-effects model predicting gaze arrival time to each event with standardized bigram surprisal of each event (negative log probability), whether the event was the first appearance of an object in that sequence (binary, 1 or 0), scaled distance from the event immediately prior to the current event, and whether the current event was a repeat of the previous event (binary, 1 or 0). The model revealed a significant effect of bigram surprisal on gaze arrival time ($\beta = -12.91, p = .008$), a significant effect of distance from the previous event ($\beta = 28.55, p =< .001$), and whether an event is a repeat of the previous event ($\beta = -331.87, p =< .001$). There is no significant effect of whether an event is the first appearance of an object ($\beta = -5.56, p = .77$), and a potentially marginal effect of the event number within a sequence ($\beta = 1.20, p = .05$).

### Pre-registered Goldilocks analysis

**Unigram surprisal** Our pre-registered analysis including both quadtratic and linear terms did not significantly improve fit for unigram surprisal in our arrival time analysis. We fit a mixed-effects model with both quadratic and linear terms. The model confirms the significant linear effect of unigram surprisal on arrival ($\beta = -2394.87, p < .001$), distance

($\beta = 28.09, p < .001$), whether an event is an immediate repeat ($\beta = -345.78, p < .001$), and the event number within a sequence ($\beta = 1.41, p < .05$). The quadratic surprisal term was non-significant ($\beta = 960.904, p = 0.078$), as well as whether an event was the first appearance of object ($\beta = -8.61, p = .70$).

**Transitional surprisal** Our pre-registered analysis including both quadtratic and linear terms also did not significantly improve fit for transitional surprisal in our arrival time analysis. We fit a mixed-effects model with both quadratic and linear terms. The model confirms the significant linear effect of transitional surprisal on arrival ($\beta = -1432.30, p < .001$), distance ($\beta = 28.53, p < .001$), whether an event is an immediate repeat ($\beta = -331.93, p < .001$). The quadratic transitional surprisal term was non-significant ($\beta = 960.904, p = 0.078$), as well as whether an event was the first appearance of object ($\beta = -5.36, p = 0.78$), and the event number within a sequence ($\beta = 1.21, p = .06$).

## Dogs look to high suprisal events longer

Dogs looked longer to high surprisal events, again confirmed to be linear by model comparison. The re-registered quadratic terms were non-significant for dwell time and therefore are not included here. Linear mixed-effects model results are reported as our main dwell time analysis.

### Positive linear relationship between unigram surprisal and dwell time

We fit a linear mixed-effects model predicting dwell time to the object that appeared during each event with standardized unigram surprisal of each event, whether the event was the first appearance of an object in that sequence (binary, 1 or 0), scaled distance from the event immediately prior to the current event, and whether the current event was a repeat of the previous event (binary, 1 or 0). The model revealed a significant effect of unigram surprisal on dwell time to the event ($\beta = 22.77, p < .001$), a significant effect of distance from the previous event ($\beta = -21.08, p =< .001$), whether an event is a repeat of the previous event ($\beta = -53.84, p =< .001$), and the event number within a sequence ($\beta = -2.05, p < .001$). There is no significant effect of whether an event is the first appearance of an object ($\beta = -12.19, p = .51$),

### Positive linear relationship between transitional surprisal and dwell time

We fit a linear mixed-effects model predicting dwell time to the object that appeared during each event with standardized transitional surprisal of each event, whether the event was the first appearance of an object in that sequence (binary, 1 or 0), scaled distance from the event immediately prior to the current event, and whether the current event was a repeat of the previous event (binary, 1 or 0). The model revealed a significant effect of transitional surprisal on dwell time to the event ($\beta = 12.24, p < .01$), a significant effect of distance from the previous event ($\beta = -21.84, p =< .001$), whether an event is a repeat of the previous event ($\beta = -65.29, p =< .001$), and the event number within a sequence ($\beta = -1.63, p < 0.01$). There is no significant effect of whether an event is the first appearance of an object ($\beta = -16.97, p = .31$),

## 6.4 Discussion

Dogs are quicker to look at, and look longer at events of high surprisal. These results indicate a preference for novelty or uncertainty. Human and primate attention is organized to regulate information intake by favoring intermediately surprising events, demonstrated by the U-shape relationship between engagement and surprisal (Kidd et al., 2012; S. Wu et al., 2022). Here, we demonstrate that dogs' looking behavior indicates a preference for high surprisal events, unlike infants and primates. Our results indicate that information value does not guide attention uniformly across mammals.

The U-shaped pattern between engagement and surprisal may have evolved more recently. The last common ancestor between dogs and humans can be traced back to about 100 million years ago (Lindblad-Toh et al., 2005). Dogs also occupy a particular niche, having been domesticated and bred for desirable traits like hunting, protection, and sensitivity to human cues. Given this evolutionary history—that dogs are distantly related to humans and primates, and the fact that humans domesticated dogs for protection—we cautiously suggest that this attentional preference for high surprisal may be a product of human selection. Preferential looking to high surprisal events may be indicative of an attention system adapted to threat detection. Further work investigating the effect of domestication would help disentangle whether information value guides attention similarly in species that have not been domesticated but are distantly related to primates.

It is possible that a preference for intermediately surprising information may be unique to humans and nonhuman primates, though more work is needed to test this theory. An attention system evolved to adaptively regulate the intake of information supports efficient learning and underpins information seeking behaviors like curiosity. We raise the possibility that this feature of attention may support sophisticated learning observed across human and nonhuman primates, like tool use and complex social systems. Future work is needed to determine whether the Goldilocks effect of attention is unique to humans and nonhuman primates.

# Bibliography


Albahiri, M., & Alhaj, A. (2020). Role of visual element in spoken English discourse: Implications for YouTube technology in EFL classrooms.

Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first 2 years of school: Patterns and processes [Place: US Publisher: Univ of Chicago Press]. *Monographs of the Society for Research in Child Development*, *53*(2), 157–157.

Anderson, D. R., Levin, S. R., & Lorch, E. P. (1977). The effects of TV program pacing on the behavior of preschool children. *Educational Communication & Technology*, *25*(2), 159–166.

Anggraini, P., Apriliani, N., Supeni, I., & Handrianto, C. (2022). The Use of the Cocomelon YouTube Channel as a Medium for Introducing Children's English Vocabulary. *SAGA Journal of English Language Teaching and Applied Linguistics*, *3*, 81–90.

Aslin, R. N. (2007). What's in a look? [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-7687.2007.00563.x]. *Developmental Science*, *10*(1), 48–53.

Baldwin, D. A. (1991). Infants' Contribution to the Achievement of Joint Reference [_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8624.1991.tb01577.x]. *Child Development*, *62*(5), 875–890.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources [Place: US Publisher: American Psychological Association]. *Psychological Bulletin*, *91*(2), 276–292.

Bentivoglio, A. R., Bressman, S. B., Cassetta, E., Carretta, D., Tonali, P., & Albanese, A. (1997). Analysis of blink rate patterns in normal subjects. *Movement Disorders: Official Journal of the Movement Disorder Society*, *12*(6), 1028–1034.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour [Num Pages: 14 Place: London, etc., United Kingdom Publisher: Cambridge University Press]. *British Journal of Psychology. General Section*, *41*(1), 68–81.

Berlyne, D. E. (1954). A theory of human curiosity [Place: United Kingdom Publisher: British Psychological Society]. *British Journal of Psychology*, *45*, 180–191.

Berlyne, D. E. (1958). The influence of complexity and novelty in visual figures on orienting responses [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology*, *55*(3), 289–296.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity* [Pages: xii, 350]. McGraw-Hill Book Company.

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function [_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8624.2010.01499.x]. *Child Development*, *81*(6), 1641–1660.

Blakley, E. C., Duggan, N., Gaspelin, N., & Gerhardstein, P. (2021). The Development of Oculomotor Suppression of Salient Distractors in Children. *Journal of Vision*, *21*(9), 2704.

Burnett, D. G., & Mitchell, E. (2025). Attention sanctuaries: Social practice guidelines and emergent strategies in attention activism. *Annals of the New York Academy of Sciences*, *1546*(1), 5–10.

Callara, A. L., Greco, A., Scilingo, E. P., & Bonfiglio, L. (2023). Neuronal correlates of eyeblinks are an expression of primary consciousness phenomena [Publisher: Nature Publishing Group]. *Scientific Reports*, *13*(1), 12617.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development [Publisher: SAGE Publications]. *Teachers College Record*, *112*(3), 579–620.

Chen, A., Kumar, S. E., Varkhedi, R., & Murphy, D. H. (2024). The Effect of Playback Speed and Distractions on the Comprehension of Audio and Audio-Visual Materials. *Educational Psychology Review*, *36*(3), 79.

Christakis, D. A., Ramirez, J. S. B., & Ramirez, J. M. (2012). Overstimulation of newborn mice leads to behavioral differences and deficits in cognitive performance [Publisher: Nature Publishing Group]. *Scientific Reports*, *2*(1), 546.

Christakis, D. A., Ramirez, J. S. B., Ferguson, S. M., Ravinder, S., & Ramirez, J.-M. (2018). How early media exposure may affect cognitive function: A review of results from observations in humans and experiments in mice [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *115*(40), 9851–9858.

Christakis, D. A., Zimmerman, F. J., DiGiuseppe, D. L., & McCarty, C. A. (2004). Early television exposure and subsequent attentional problems in children. *Pediatrics*, *113*(4), 708–713.

Cohen, L. B. (1972). Attention-Getting and Attention-Holding Processes of Infant Visual Preferences [Publisher: [Wiley, Society for Research in Child Development]]. *Child Development*, *43*(3), 869–879.

Cooper, N., Uller, C., Pettifer, J., & Stolc, F. (2009). Conditioning attentional skills: Examining the effects of the pace of television editing on children's attention [_eprint: https://onlinelibrary.wiley.co 2227.2009.01377.x]. *Acta Paediatrica*, *98*(10), 1651–1655.

COUNCIL ON COMMUNICATIONS AND MEDIA, Hill, D., Ameenuddin, N., Reid Chassiakos, Y. (, Cross, C., Hutchinson, J., Levine, A., Boyd, R., Mendelson, R., Moreno, M., & Swanson, W. S. (2016). Media and Young Minds. *Pediatrics*, *138*(5), e20162591.

Crosnoe, R., Leventhal, T., Wirth, R. J., Pierce, K. M., Pianta, R. C., & Network, N. E. C. C. R. (2010). Family Socioeconomic Status and Consistent Environmental Stimulation in Early Childhood [_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8624.2010.01446.x]. *Child Development*, *81*(3), 972–987.

Cubit, L. S., Canale, R., Handsman, R., Kidd, C., & Bennetto, L. (2021). Visual Attention Preference for Intermediate Predictability in Young Children [_eprint: https://srcd.onlinelibrary.wiley.com *Child Development*, *92*(2), 691–703.

Cutting, J. E. (2016). The evolution of pace in popular movies. *Cognitive Research: Principles and Implications*, *1*(1), 30.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children [Place: US Publisher: Psychonomic Society]. *Behavior Research Methods, Instruments & Computers*, *28*(1), 125–127.

Dember, W. N., & Earl, R. W. (1957). Analysis of exploratory, manipulatory, and curiosity behaviors [Place: US Publisher: American Psychological Association]. *Psychological Review*, *64*(2), 91–96.

Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A.-M. (2017). Academic Interventions for Elementary and Middle School Students With Low Socioeconomic Status: A Systematic Review and Meta-Analysis [Publisher: American Educational Research Association]. *Review of Educational Research*, *87*(2), 243–282.

DIRECTV. (2025, January). 10 Most Popular PBS Kids Shows on Now.

Doperalski, D. (2024, March). The Variety Streaming Originals Charts.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity [Place: US Publisher: American Psychological Association]. *Psychological Review*, *127*(3), 455–476.

Ekris, E. V., Chinapaw, M. J. M., Rotteveel, J., & Altenburg, T. M. (2018). Do Young People Ever Sit Still? Variations in Accelerometer Counts, Muscle Activity and Heart Rate across Various Sedentary Activities in Youth. *International Journal of Environmental Research and Public Health*, *15*(5).

Essex, C., Gliga, T., Singh, M., & Smith, T. J. (2022). Understanding the differential impact of children's TV on executive functions: A narrative-processing analysis. *Infant Behavior & Development*, *66*, 101661.

Geist, E. A., & Gibson, M. (2000). The effect of network and public television programs on four and five year olds ability to attend to educational tasks [Place: US Publisher: Journal of Instructional Psychology]. *Journal of Instructional Psychology*, *27*(4), 250–261.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269.

Godwin, K. E., & Fisher, A. V. (2018). Wiggleometer: Measuring Selective Sustained Attention in Children. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *40*(0).

Goodrich, S. A., Pempek, T. A., & Calvert, S. L. (2009). Formal Production Features of Infant and Toddler DVDs. *Archives of Pediatrics & Adolescent Medicine*, *163*(12), 1151–1156.

Gorey, K. M. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity [Place: US Publisher: Guilford Publications]. *School Psychology Quarterly*, *16*(1), 9–30.

Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool Programs Can Boost School Readiness [Publisher: American Association for the Advancement of Science]. *Science*, *320*(5884), 1723–1724.

Gormley Jr., W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The Effects of Universal Pre-K on Cognitive Development [Place: US Publisher: American Psychological Association]. *Developmental Psychology*, *41*(6), 872–884.

Haith, M. M. (1980). *Rules that Babies Look by: The Organization of Newborn Visual Activity* [Google-Books-ID: QxpivgAACAAJ]. L. Erlbaum Associates.

Hartanto, T. A., Krafft, C. E., Iosif, A. M., & Schweitzer, J. B. (2016). A trial-by-trial analysis reveals more intense physical activity is associated with better cognitive control performance in attention-deficit/hyperactivity disorder. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *22*(5), 618–626.

Hayes, C. (2025). *The Sirens' Call: How Attention Became the World's Most Endangered Resource*. Penguin Press.

Heaton, B. (2024, December). 'Brain rot' named Oxford Word of the Year 2024.

Hedderson, M. M., Bekelman, T. A., Li, M., Knapp, E. A., Palmore, M., Dong, Y., Elliott, A. J., Friedman, C., Galarce, M., Gilbert-Diamond, D., Glueck, D., Hockett, C. W., Lucchini, M., McDonald, J., Sauder, K., Zhu, Y., Karagas, M. R., Dabelea, D., Ferrara, A., & Environmental Influences on Child Health Outcomes Program. (2023). Trends in Screen Time Use Among Children During the COVID-19 Pandemic, July 2019 Through August 2021. *JAMA Network Open*, *6*(2), e2256157.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, *90*, 36–53.

Hiniker, A., Suh, H., Cao, S., & Kientz, J. A. (2016). Screen Time Tantrums: How Families Manage Screen Media Experiences for Toddlers and Preschoolers. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 648–660.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli [Place: US Publisher: Elsevier/JAI Press Inc]. *Advances in Infancy Research*, *5*, 69–95.

IAB. (2023, May). *2022 Video Ad Spend & 2023 Outlook: Defining the Next Generation* (tech. rep.). Interactive Advertising Bureau.

Irwin, D. E., Colcombe, A. M., Kramer, A. F., & Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research*, *40*(10), 1443–1458.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention [Number: 3 Publisher: Nature Publishing Group]. *Nature Reviews Neuroscience*, *2*(3), 194–203.

J.D.Power. (2022). Cost of video streaming services in the U.S. 2022. *J.D. Power*.

Kahneman, D. (1973). Attention and Effort.

Kidd, C., & Hayden, B. Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron*, *88*(3), 449–460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex [Publisher: Public Library of Science]. *PLOS ONE*, *7*(5), e36399.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795–1804.

Kim, J. S., & Quinn, D. M. (2013). The Effects of Summer Reading on Low-Income Children's Literacy Achievement From Kindergarten to Grade 8: A Meta-Analysis of Classroom and Home Interventions [Publisher: American Educational Research Association]. *Review of Educational Research*, *83*(3), 386–431.

Kline, P., & Walters, C. R. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start*. *The Quarterly Journal of Economics*, *131*(4), 1795–1848.

Kostyrka-Allchorne, K., Cooper, N. R., Gossmann, A. M., Barber, K. J., & Simpson, A. (2017). Differential effects of film on preschool children's behaviour dependent on editing pace [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/apa.13770]. *Acta Paediatrica*, *106*(5), 831–836.

Kostyrka-Allchorne, K., Cooper, N. R., Kennett, S., Nestler, S., & Simpson, A. (2019). The Short-Term Effect of Video Editing Pace on Children's Inhibition and N2 and P3 ERP Components during Visual Go/No-Go Task [Publisher: Routledge _eprint: https://doi.org/10.1080/875656 *Developmental Neuropsychology*, *44*(4), 385–396.

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). The relationship between television exposure and children's cognition and behaviour: A systematic review. *Developmental Review*, *44*, 19–58.

Lang, A., Bolls, P., Potter, R. F., & Kawahara, K. (1999). The effects of production pacing and arousing content on the information processing of television messages [Publisher: Routledge _eprint: https://doi.org/10.1080/08838159909364504]. *Journal of Broadcasting & Electronic Media*, *43*(4), 451–475.

Lang, D., Chen, G., Mirzaei, K., & Paepcke, A. (2020). Is faster better? a study of video playback speed. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 260–269.

Lareau, A. (2018). Unequal Childhoods: Class, Race, and Family Life [Num Pages: 8]. In *Inequality in the 21st Century*. Routledge.

Lavie, N. (2005). Distracted and confused?: Selective attention under load [Publisher: Elsevier]. *Trends in Cognitive Sciences*, *9*(2), 75–82.

Lavie, N. (2010). Attention, Distraction, and Cognitive Control Under Load [Publisher: SAGE PublicationsSage CA: Los Angeles, CA]. *Current Directions in Psychological Science*.

Lavie, N., Hirst, A., Fockert, J., & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. *Journal of Experimental Psychology: General*, *133*, 339–354.

Lewandowsky, S., Robertson, R. E., & DiResta, R. (2023). Challenges in Understanding Human-Algorithm Entanglement During Online Information Consumption [Publisher: SAGE Publications Inc]. *Perspectives on Psychological Science*, 17456916231180809.

Li, L., Valiente, C., Eisenberg, N., Spinrad, T. L., Johns, S. K., Berger, R. H., Thompson, M. S., Southworth, J., Pina, A., Hernández, M. M., & Gal-Szabo, D. E. (2022). Longitudinal Relations between Behavioral Engagement and Academic Achievement: The Moderating

Roles of Socio-economic Status and Early Achievement. *Journal of school psychology*, *94*, 15–27.

Lillard, A. S., Drell, M. B., Richey, E. M., Boguszewski, K., & Smith, E. D. (2015). Further examination of the immediate impact of television on children's executive function [Place: US Publisher: American Psychological Association]. *Developmental Psychology*, *51*(6), 792–805.

Lillard, A. S., Li, H., & Boguszewski, K. (2015, January). Chapter Seven - Television and Children's Executive Function. In J. B. Benson (Ed.), *Advances in Child Development and Behavior* (pp. 219–248, Vol. 48). JAI.

Lillard, A. S., & Peterson, J. (2011). The Immediate Impact of Different Types of Television on Young Children's Executive Function. *Pediatrics*, *128*(4), 644–649.

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., Clamp, M., Chang, J. L., Kulbokas, E. J., Zody, M. C., Mauceli, E., Xie, X., Breen, M., Wayne, R. K., Ostrander, E. A., Ponting, C. P., Galibert, F., Smith, D. R., deJong, P. J., ... Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog [Publisher: Nature Publishing Group]. *Nature*, *438*(7069), 803–819.

Lissak, G. (2018). Adverse physiological and psychological effects of screen time on children and adolescents: Literature review and case study. *Environmental Research*, *164*, 149–157.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation [Place: US Publisher: American Psychological Association]. *Psychological Bulletin*, *116*(1), 75–98.

Loschky, L. C., Larson, A. M., Smith, T. J., & Magliano, J. P. (2020). The Scene Perception & Event Comprehension Theory (SPECT) Applied to Visual Narratives [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12455]. *Topics in Cognitive Science*, *12*(1), 311–351.

LuminateData. (2024). Film & TV Streaming Viewership.

Lynn, R. (2013, September). *Attention, Arousal and the Orientation Reaction: International Series of Monographs in Experimental Psychology* [Google-Books-ID: 1n1GBQAAQBAJ]. Elsevier.

Madigan, S., Browne, D., Racine, N., Mori, C., & Tough, S. (2019). Association Between Screen Time and Children's Performance on a Developmental Screening Test. *JAMA Pediatrics*, *173*(3), 244–250.

Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. (2004). Inequality in Preschool Education and School Readiness [Publisher: American Educational Research Association]. *American Educational Research Journal*, *41*(1), 115–157.

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.

Montessori, M. (1917). *The Advanced Montessori Method*. Frederick A. Stokes Company.

N, S. E. (1960). Neuronal models and the orienting reflex [Publisher: Josiah Macey Junior. Foundation]. *The central nervous system and behaviour*.

Nazaryan, A. (2026). They Want to Influence You to Do ... Nothing. *The New York Times*.

Netflx. (2023, December). What We Watched: A Netflix Engagement Report.

Neumann, M. M., & Herodotou, C. (2020). Young Children and YouTube: A global phenomenon [Publisher: Routledge _eprint: https://doi.org/10.1080/00094056.2020.1796459]. *Childhood Education*, *96*(4), 72–77.

Nguyen, L., Walters, J., Paul, S., Monreal Ijurco, S., Rainey, G. E., Parekh, N., Blair, G., & Darrah, M. (2025). Feeds, feelings, and focus: A systematic review and meta-analysis examining the cognitive and mental health correlates of short-form video use. *Psychological Bulletin*, *151*(9), 1125–1146.

Oppezzo, M., Schwartz, D. L., Aikens, R., Moore, J. B., Langford, J., Baiocchi, M., Gross, J. J., & Dohmen, I. (2025). Sit still or move more? The impact of fidgeting on creativity [Place: US Publisher: American Psychological Association]. *Journal of Educational Psychology*.

Oudeyer, P.-Y., & Smith, L. B. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12196]. *Topics in Cognitive Science*, *8*(2), 492–502.

Parambil, M. M. A., Ali, L., Alnajjar, F., & Gochoo, M. (2022). Smart Classroom: A Deep Learning Approach towards Attention Assessment through Class Behavior Detection. *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, 1–6.

Pavlov (1927), P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, *17*(3), 136–141.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning [Publisher: American Association for the Advancement of Science]. *Science Advances*, *6*(39), eabb5053.

Rahnev, D. (2025). A comprehensive assessment of current methods for measuring metacognition. *Nature Communications*, *16*(1), 701.

Reese, D., Pawluk, D., & Taylor, C. (2016, December). Engaging Learners Through Rational Design of Multisensory Effects. In *Emotions, Technology, and Design* (pp. 103–127).

Rideout, V., & Robb, M. (2020). The Common Sense Census: Media Use by Kids Age Zero to Eight, 2020 — Common Sense Media. *San Francisco, CA: Common Sense Media.*

Rouault, M., & Fleming, S. M. (2020). Formation of global self-beliefs in the human brain. *Proceedings of the National Academy of Sciences*, *117*(44), 27268–27276.

Ryan, E. (2014). Toward a Redefinition of "Educational" in Babies' and Toddlers' Media:Overt Visual Attention, Verbalization, and Other Measures of Engagement as Indicators of Learning. *Advances in Social Sciences Research Journal*, *1*, 105–114.

Salapatek, P., & Kessen, W. (1966). Visual scanning of triangles by the human newborn. *Journal of Experimental Child Psychology*, *3*(2), 155–167.

Sanders, A., Boswell, B., Walia, G. S., & Allen, A. (2021). Non-Intrusive Classroom Attention Tracking System (NiCATS). *2021 IEEE Frontiers in Education Conference (FIE)*, 1–9.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded [Place: US Publisher: American Psychological Association]. *Developmental Psychology*, *43*(4), 1045–1050.

Segal, D. (2022). A Kid's Show Juggernaut That Leaves Nothing to Chance. *The New York Times*.

Shepherd, S. S., & Kidd, C. (2024). Visual engagement is not synonymous with learning in young children. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*(0).

Simon, H. A., Deutsch, K. W., & Shubik, M. (1971). Designing organizations for an information-rich world. *Computers, communications, and the public interest*.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science [Publisher: Routledge _eprint: https://doi.org/10.1080/00461520.20
*Educational Psychologist*, *50*(1), 1–13.

Smith, D. G. (2023). How to Focus Like It's 1990. *The New York Times*.

Socialblade. (2024). Top 100 YouTubers made-for-kids Channels - Socialblade YouTube Stats — YouTube Statistics.

Son, H. M., Calub, C. A., Fan, B., Dixon, J. F., Rezaei, S., Borden, J., Schweitzer, J. B., & Liu, X. (2024). A quantitative analysis of fidgeting in ADHD and its relation to performance and sustained attention on a cognitive task [Publisher: Frontiers]. *Frontiers in Psychiatry*, *15*.

Stahl, A. E., & Feigenson, L. (2015). Cognitive development. Observing the unexpected enhances infants' learning and exploration [Place: New York, N.Y.]. *Science*, *348*(6230), 91–94.

Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The Endogenous Eyeblink [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8986.1984.tb02312.x]. *Psychophysiology*, *21*(1), 22–33.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296.

Tharumalingam, T., Roberts, B. R. T., Fawcett, J. M., & Risko, E. F. (2025). Increasing Video Lecture Playback Speed Can Impair Test Performance – a Meta-Analysis. *Educational Psychology Review*, *37*(2), 35.

The Friends of Attention. (2026, January). *Attensity!: A Manifesto of the Attention Liberation Movement*. Crown.

Toleuzhan, A., Sarzhanova, G., Romanenko, S., Uteubayeva, E., & Karbozova, G. (2022). The Educational Use of YouTube Videos in Communication Fluency Development in English: Digital Learning and Oral Skills in Secondary Education. *International Journal of Education in Mathematics, Science and Technology*, *11*(1), 198–221.

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/desc.12629]. *Developmental Science*, *21*(4), e12629.

Upton-Clark, E. (2024, November). 'PDF to brainrot' tools are bringing TikTok visuals to Gen Alpha's academics.

Völter, C. J., Tomašić, A., Nipperdey, L., & Huber, L. (2023). Dogs' expectations about occlusion events: From expectancy violation to exploration [Publisher: Royal Society]. *Proceedings of the Royal Society B: Biological Sciences*, *290*(2003), 20230696.

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain Drain: The Mere Presence of One's Own Smartphone Reduces Available Cognitive Capacity [Publisher: The University of Chicago Press]. *Journal of the Association for Consumer Research*, *2*(2), 140–154.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills [_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/cdev.12099]. *Child Development*, *84*(6), 2112–2130.

Wu, S., Blanchard, T., Meschke, E., Aslin, R., Hayden, B., & Kidd, C. (2022, April). Macaques preferentially attend to intermediately surprising information.

Wu, T. (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. Knopf Doubleday Publishing Group.

Yousef, A. M. F., Alshamy, A., Tlili, A., & Metwally, A. H. S. (2025). Demystifying the New Dilemma of Brain Rot in the Digital Era: A Review. *Brain Sciences*, *15*(3).