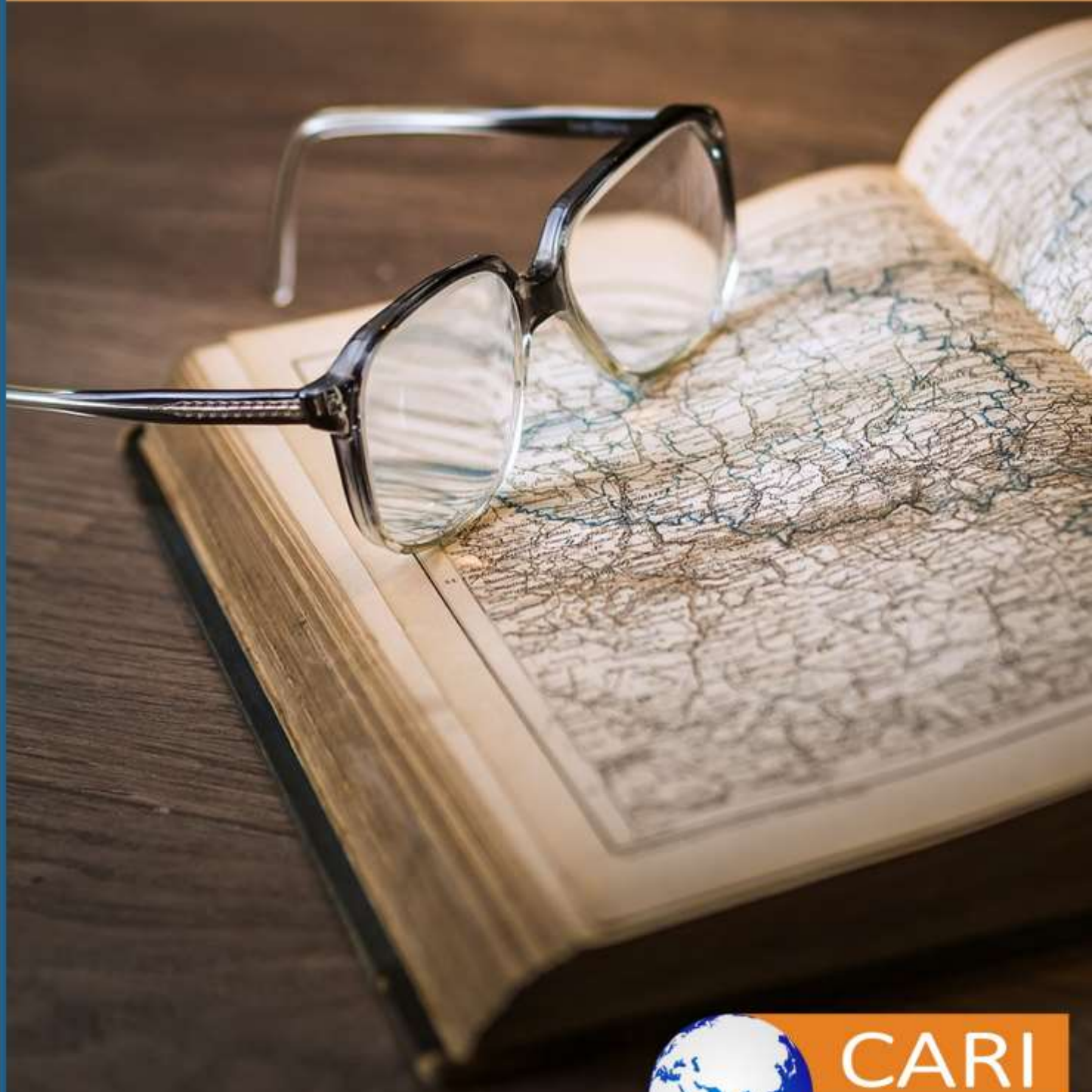


International Journal of Humanity and Social Sciences

(IJHSS) **The Thumb-Object Rule: Neuroscience Strategies for
Enhancing Viewer Attention in Vertical Short Videos**



CARI
Journals

The Thumb-Object Rule: Neuroscience Strategies for Enhancing Viewer Attention in Vertical Short Videos

 Dmitrii Zenkov

Hi Agency, Sheridan, Wyoming, USA.

<https://orcid.org/0009-0002-8219-2060>

Accepted: 17th March, 2026, Received in Revised Form: 30th March, 2026, Published: 6th April, 2026

ABSTRACT

Purpose: This work provides directors and content creators with data-driven strategies to capture and hold viewer attention in 9x16 vertical short videos. The goal is to align creative choices with how the human brain naturally processes visual information on mobile screens.

Methodology: The research is based on a review of 14 sources, including computational models of "saliency" (what stands out), neuroimaging (brain scans like fMRI and EEG), and eye-tracking experiments specifically for vertical formats. The analysis focuses on how low-level visual features—like contrast and size—affect a viewer's brain during the first 800 milliseconds of a video.

Findings: The study shows that the brain's primary visual cortex quickly creates a "saliency map" that prioritizes large objects located in the center of the frame. The Thumb-Object Rule—placing a main object larger than a thumb in the center of the frame—reduces the effort a viewer's eyes must take to find the subject, which significantly lowers skip rates. This "neural pop-out" effect is strongest when the object lines up with the vertical midline of the phone.

Unique contribution to theory, practice and policy (recommendations):

This research turns complex neuroscience into a practical checklist for artistic and video directors to replace vague creative directions with measurable data.

Unique Contribution to Theory, Practice and Policy: Make the main subject take up approximately 8–12% of the vertical height. Second, should ensure the object is visually larger than a thumb on a standard smartphone screen. Place the subject in the central third (the vertical midline) of the 9x16 frame. Maintain this focused layout for the first 0.3–0.8 seconds to allow the brain to process the image before it decides to skip. Use color contrast, lighting, or small movements to make the object stand out.

Keywords: *Vertical video, Visual saliency, Thumb-Object Rule, Neuro-design, Viewer retention, Artistic direction*

1. Introduction

This study connects cognitive neuroscience to digital content creation. It aims to aid directors, content creators, visual storytellers, and arts researchers making quick decisions about editing and framing for short vertical videos. The ideas come from the author's 15 years of directing entertainment and internet videos. Having led over 250 projects in different markets, the author-researcher found a 18–20% increase in profits by using science in visual composition. A key method is reducing mental strain on creative staff. Short-form vertical videos have become a dominant format across social media platforms. Visual design and structural clarity of such videos are increasingly considered crucial for audience engagement (Astani, 2018). The growth of short-form videos and their use in advertising campaigns is also mentioned by research on content virality and the choice of YouTube as a distribution platform (Saputra, 2019). By using data-based parameters instead of unclear directions (e.g., “make it look good”), production becomes more efficient. The Thumb-Object Rule is central to this. It suggests that the main object in the first frame should be bigger than a thumb on a phone screen. Since users often use thumbs to scroll through social media, a smaller object might be missed, causing viewers to skip. The Thumb-Object Rule comes from an analysis of fourteen sources, like computational saliency models (Itti et al., 1998), visual cortex mapping (Zhang et al., 2012; Liu et al., 2025), visual search (Nothdurft, 2002; Dolci et al., 2024), and eye-tracking studies of vertical frames (Baig et al., 2025). This paper translates these findings into advice for micro-data video (Sallam et al., 2025). The brain quickly makes representations that decide where attention goes within 300–800 milliseconds. By using the vertical frame to focus on a main element, directors can cut down on search effort and raise pop-out (Baig et al., 2025; Thayer; Sprague, 2023). Because this paper is for practitioners and researchers, neuroimaging data is less important than clear choices in blocking, lighting, and editing. By matching creative sense with how people see, directors can make sure their work is noticed in fast digital environments.

2. Methodology

This research used a review method, bringing together fourteen papers from many fields to link brain science with filmmaking. The papers were picked to give a broad view of attention, from basic brain activity to how stories grab us. To keep things both careful and helpful, the papers were put into four groups:

1. **Basics of Neuroscience and Models:** Key papers on how we notice things (Itti et al., 1998; Zhang et al., 2012) and recent brain scans showing that the first visual area (V1) is key for attention (Liu et al., 2025; Wang et al., 2022).

2. **How Attention Works and Learning:** Studies using behavior and brain waves to study how the brain picks what to focus on (Nothdurft, 2002), how we map different features (Thayer & Sprague, 2023), and how what we've learned affects what we notice now (Dolci et al., 2024).

3. **Vertical Video and Its Design:** Research on where people look on the screen in 9:16 videos (Baig et al., 2025) and design tips for short videos on phones (Sallam et al., 2025; Amini et al., 2015).

4. Media Analysis and Examples: Data on what competes for our attention (Nardo et al., 2016), local info from Jurnal Ars on TikTok and ads (Ramadini et al., 2024; Fitaloka et al., 2025), and thoughts on creative impact (Peng, 2025).

The analysis was done in three steps. First, each paper was looked at with three questions in mind: (a) How do we pick what to pay attention to first? (b) What proof is there about the first 800ms of watching vertical videos? (c) How does this help a director on set or when editing?

Second, the review focused on operational translations. When papers gave numbers—like how important being in the middle is for vertical videos (Baig et al., 2025)—these were used to set exact numbers and times (like the 8–12% height rule) in the framework.

Third, differences were kept in mind. The author knew that lab tests (using simple things to look at) are not the same as the busy visuals of today's short videos. This was used to make the Thumb-Object Rule a suggestion for creative teams to test, not a hard rule. This way, the method is based on science, helpful, and fits what today's digital design needs.

3. Results and discussion

This study suggests that having a big, centered, and visually distinct object in the first frame greatly raises the chance of early viewing and lowers the chance of early skipping. This idea is backed by findings from studies of computations, the brain, and perception.

The model of visual importance (by Itti et al., 1998) gives us a way to talk about this framework. This model says that simple traits, like brightness, color, and direction, come together to form a saliency map. A director can use this to control attention by mixing these traits. When something is big, different, and in the middle, it makes a strong point on the saliency map that the brain notices.

Brain scans show how fast this happens. Detailed fMRI and EEG data show that the top layers of the primary visual cortex (V1) create these saliency maps very quickly. Then, they send them to the parietal cortex to guide attention (Zhang et al., 2012; Liu et al., 2025; Wang et al., 2022). This automatic process happens in the first 300–800ms of seeing something. For a director, this means Frame Zero is a key time: the traits chosen for the very first moment decide if the brain's automatic systems start working or if the viewer skips.

Action studies link these brain tendencies to pop-out effects. These effects mean we notice important things no matter how complex the scene is (Nothdurft, 2002). This effect is stronger in the 9:16 vertical format. Because the vertical frame is thin, the middle line is very important. Eye-tracking data for vertical content shows that putting things in the center (Central AOI) makes people look at them faster and pays attention better (Coefficient K) than putting them to the side (Baig et al., 2025). By centering the object, the director helps the viewer avoid looking around the screen, which saves time for understanding the story.

But importance works with statistical learning (Dolci et al., 2024). While basic importance (size and center) grabs the viewer at first, seeing things repeatedly helps people learn common patterns. This means series creators should use a two-part plan: use the Thumb-Object Rule to

hook viewers at the start, then add some variety to keep things interesting (Thayer; Sprague, 2023).

The Thumb-Object Rule is a simple way to use these ideas. Field studies of TikTok and micro-data videos (Ramadini et al., 2024; Sallam et al., 2025; Amini et al., 2015) agree with the research: clear visuals at the start lead to more viewer involvement. When something takes up 8–12% of the vertical space and is bigger than the thumb used for scrolling, it becomes a clear visual anchor.

This rule is especially important in crowded spaces. Studies of natural videos show that when many things compete for attention, each important thing matters less (Nardo et al., 2016). So, the director needs to make sure the opening is very focused, like a close-up in the center, before showing more complex scenes.

To use this idea in practice, directors should plan the first 0.8 seconds carefully. This means:

Blocking: Put the main object in the center, either vertically or horizontally.

Scaling: Make sure the object looks bigger than a thumb on a normal phone screen.

Reinforcement: Add other cues like small movements or focused lighting.

Directors can test this by comparing a Thumb-Object version with a normal version. They can measure how quickly people look at it in a lab and how often people skip it in real life. Even small tests can give good data for specific viewers (Baig et al., 2025).

In the video by Zack D. Films (Peng, 2025), the director puts the performer in the center against an empty background, making a strong vertical shape that the brain notices easily. In commercial design (Fitaloka et al., 2025), putting the product in the center with strong contrast makes it stand out, helping the Informing-Persuading-Reminding process that keeps customers coming back (Ramadini et al., 2024).

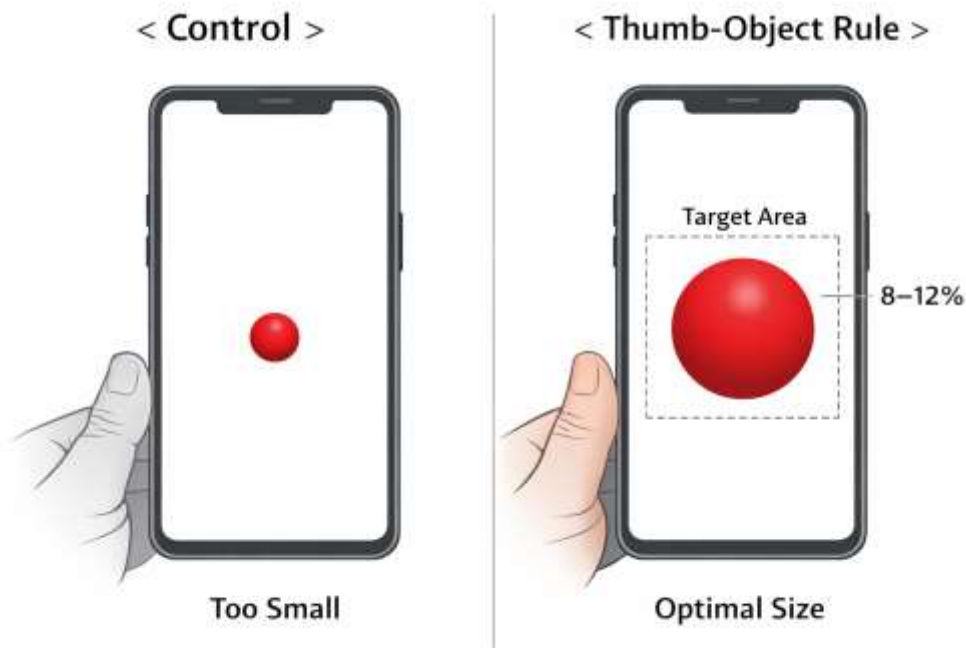


Illustration 1. How Thumb-Object Rule works

4. Conclusion

Evidence from models of computational salience, brain mapping of the visual cortex, and eye-tracking studies suggests a rule for video production: to keep viewers watching, the initial object in a 9x16 short video should be big, centered, and easy to see. This Thumb-Object Rule, shown in Illustration 1, links early production choices to how the brain quickly processes visual info and how people watch videos on phones. By treating the first few tenths of a second of a video as key for grabbing attention, producers can use this simple trick to lower the chance that people will skip the video.

The Thumb-Object Rule provides a data-driven production framework for vertical video, recommending that directors place a visually distinct subject—highlighted by color, lighting, or motion—directly in the center of the 9x16 frame for the first 300–800ms. To maximize engagement and minimize skip rates, this central object should occupy approximately 8–12% of the vertical frame, ensuring it is large enough to be seen behind the viewer's scrolling thumb. While current research is limited by small sample sizes and requires broader cultural validation, this approach serves as a practical tool that moves filmmaking beyond guesswork by aligning creative choices with the brain's rapid visual processing systems. Ultimately, this method helps creators produce clearer, high-impact openings that are more likely to succeed on modern mobile platforms.

References

- Amini, F., Henry Riche, N., Lee, B., Hurter, C., & Irani, P. (2015). Understanding data videos: Looking at narrative visualization through the cinematography lens. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15) (pp. 1459–1468). ACM. <https://doi.org/10.1145/2702123.2702431>
- Astani, A. R. (2018). Perancangan video profil khasanah batik “Adi Purwo” Kabupaten Purworejo. *Ars: Jurnal Seni Rupa dan Desain*, 21(1). <https://doi.org/10.24821/ars.v21i1.2526>
- Baig, A. Y., Riaz, M., Baig, M. A., & Hasan, M. Z. (2025). Dynamics of viewer’s attention: Making content more engaging in a vertical frame. *The Regional Tribune*, 4(3), 193–208. <https://doi.org/10.55737/trt/SR25.138>
- Dolci, S., Rashal, E., Santandrea, E., Ben Hamed, S., Chelazzi, L., Macaluso, E., & Boehler, C. N. (2024). The dynamics of statistical learning in visual search and its interaction with salience processing: An EEG study. *NeuroImage*, 286, 120514. <https://doi.org/10.1016/j.neuroimage.2024.120514>
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259. <https://doi.org/10.1109/34.730558>
- Liu, C., Liu, C., Huber, L., Zhaoping, L., & Zhang, P. (2025). The superficial layers of the primary visual cortex create a saliency map that feeds forward to the parietal cortex. *PLoS Biology*, 23(10), e3003159. <https://doi.org/10.1371/journal.pbio.3003159>
- Nardo, D., Console, P., Reverberi, C., & Macaluso, E. (2016). Competition between visual events modulates the influence of salience during free-viewing of naturalistic videos. *Frontiers in Human Neuroscience*, 10, 320. <https://doi.org/10.3389/fnhum.2016.00320>
- Nothdurft, H.-C. (2002). Attention shifts to salient targets. *Vision Research*, 42(10), 1287–1306. [https://doi.org/10.1016/S0042-6989\(02\)00016-0](https://doi.org/10.1016/S0042-6989(02)00016-0)
- Peng, S.-H. L. (2025). Dangling in the sky: A thought experiment about fear, reality and letting go. *Information Matters*, 1–3. <https://informationmatters.org/2025/07/dangling-in-the-sky-a-thought-experiment-about-fear-reality-and-letting-go/>
- Fitaloka, K., Suryani, E., & Firdaus. (2025). Konstruksi representasi perempuan Muslim modern melalui elemen desain komunikasi visual dalam iklan Wardah “Beauty Moves You”. *Ars: Jurnal Seni Rupa dan Desain*, 28(3), 171–180. <https://doi.org/10.24821/ars.v28i3.18231>
- Ramadani, L., Rahman, Y., & Putra, I. D. A. D. (2024). Analisis video TikTok untuk mempertahankan perhatian konsumen terhadap layanan online jasa Aqiqah. *ARS*, 27(2), 85–92. <https://doi.org/10.24821/ars.v27i2.12255>
- Sallam, S., Sakamoto, Y., Herath, A., Petrie, J., Brussoni, M., Jacob, J. M., & Irani, P. (2025). Unpacking micro data videos: Key elements and design practices in minute-long data videos for mobile usage. *Proceedings of the ACM on Human-Computer Interaction*, 9(5), Article MHCI036. <https://doi.org/10.1145/3743722>

Saputra, P. S. (2019). Analisis wacana kritis iklan film pendek LINE versi “Ada Apa Dengan Cinta?” Ars: Jurnal Seni Rupa dan Desain, 22(1),16–24.

<https://doi.org/10.24821/ars.v22i1.2764>

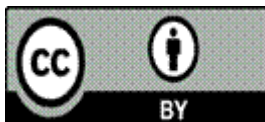
Thayer D.D, Sprague T.C. Feature-Specific Saliency Maps in Human Cortex. J Neurosci. 2023 Dec 13;43(50):8785-8800. <https://doi.org/10.1523/jneurosci.1104-23.2023>

Wang, L., Huang, L., Li, M., Wang, X., Wang, S., Lin, Y., & Zhang, X. (2022). An awareness-dependent mapping of saliency in the human visual system. NeuroImage, 247, 118864.

<https://doi.org/10.1016/j.neuroimage.2021.118864>

Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012). Neural activities in V1 create a bottom-up saliency map. Neuron, 73(1), 183–192.

<https://doi.org/10.1016/j.neuron.2011.10.035>



©2026 by the Authors. This Article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>)