

A Critical Perspective on the “Swipe-Up” Interaction Design in Short-Form Video Interfaces

Li Ma*

Shandong Urban Service Vocational College, Yantai, Shandong, 264670, China

*Corresponding author: Li Ma, 782833233@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The proliferation of short-form video platforms has led to a distinctive interaction pattern known as the “swipe-up” gesture, which facilitates content consumption but raises concerns regarding user engagement, cognitive load, and ethical design. This study critically examines the interface design of the swipe-up interaction from both human-computer interaction (HCI) and sociotechnical perspectives. Through literature analysis, interface observation, and heuristic evaluation, the research identifies key usability challenges and cognitive risks, proposes design improvements, and discusses the implications for ethical and user-centered design in short-form video ecosystems. Two experimental analyses are presented, accompanied by visualizations and formula-based metrics. The findings provide actionable insights for designers and policy-makers aiming to balance engagement with user well-being.

Keywords: Short-Form Video; Interface Design; Swipe-Up Interaction

Published: Aug 28, 2026

DOI: <https://doi.org/10.62177/apemr.v3i7.1675>

1. Introduction: Critical Perspective on Swipe-Up Interaction

As for miniature video portals like Tiktok, instagram reels, youtube shorts they have rebuilt online media engagement. The heart of things, was really just this “Swipe UP,” offering up vertical swiping movements so people could freely move between all kinds of visual scenes indefinitely. But despite seeming simple, it has big sociotechnic consequence. Modern study focuses on the harm to focus times, thoughts tiredness, and habits, so we must thoroughly examine swipe-up functions.

This current study takes a negative attitude towards Interaction Design, it studies what kind of impact swipe-up layout gives to user’s behavior, as well as if some design changes can ease users’ cognitive pressure without losing function, it also find out usability advantage and limits, examine effects on users’ attention span, interest and decision-making power, explore if any redesign can reduce the risk and promote responsible use.

2. Critical Perspectives and Prior Research on Swipe-Up Interaction Design in Short-Form Video Platforms

2.1 Short-Form Video Interaction Patterns

The great majority of short-form video systems depend on a lot and simple use interaction so they can keep users active for a long time. And within them all these swipes ups become like that important navigation tool helping folks go smooth and uninterrupted between one clip after the other so its one long unbroken process of watching media. Like such involvement brings fast answers and makes one feel good with quick supplies of contents that might boost routine purchasing However,

many studies pointed out there could be negative aspects of such structural selection. Routine swipe – up actions lead to addiction as well, no control of self, more screen time. Furthermore, eye tracking shows people interacting with swipe-up layout will often spend more time looking at the main visual area and may not notice much of the surrounding graphics^[1]. And it also can show we have this small scope due to layout, and that might cause mental judgement on periphery data, and cause fatigue after long time. And more so: Speedy stuff rotating paired with steady side moves probably makes it harder to keep your focus and carefull thinks about things chewed on already - which results in passive participate - big worries ‘boud for moral shapechin’ and lastcliend health from this kind’a behavior^[2].

2.2 Critical Perspectives on Interface Design

Critical Design has certain key points like there is a big problem in the usability orientation in Interface Construction, Efficiency takes over very quickly. Conversely, Critical Design is about all the other possible moral, communal and mental consequences for humans of making certain choices in the designing process. For a given domain like swipe-ups on short-videos^[3], it makes us look into the outcomes of interface protocols which mainly aim at increasing how long people use something or interact with it. When it comes to stats about total screen time or swiping regardless of personal wellbeing can also lead towards promoting addiction, causing burnout and lowering mental health. Metrics like retaining doesn’t show these harmful things, it ignores this complicated relationship between looking at something and the mental stuff that comes when someone feels certain emotions. So, it becomes very important to have a common appraising setup which has both standard usability tests as well as ethics,cognition and community impact^[4]. And it allows developers to spot and reduce all sorts of potential dangers, meaning that participatory things will be not just useful or nice but also responsible for the well-being of people taking part over a long time.

2.3 Cognitive Load and Interaction Complexity

According to cognitive load theory (CLT), what the person must manage in order to be present with the information is not simply their own intrinsic load but also includes things like how the layout and visuals look (extrinsic load) as well as all the mental work put forth towards building and acquiring structure (germane load). In terms of short-form video apps, swipe-up is a persistent way to view the content and it has some influence on those cognitive elements. Although it is just a screen interface, this kind of new input sequence will increase extra mental load as there are few environmental clues that force people to keep switching their focus and guessing what comes next^[5]. And this would be less available faculties in our minds of interpreting important information, which may possibly induce cognitive tiredness, bad memorizing recording, and lesser involved thoughtfully^[6]. And then, the total cognitive burden borne by people can be written as follows:

Formula 1:

$$CL_{total}=CL_{intrinsic}+CL_{extraneous}+CL_{germane}$$

In this formula, $CL_{intrinsic}$ reflects the inherent complexity of the video content, including elements such as narrative structure, visual density, and the novelty of information. $CL_{extraneous}$ represents the additional cognitive burden imposed by the interface itself, including rapid content transitions, continuous scrolling, and the absence of natural pause points. $CL_{germane}$ corresponds to the mental effort dedicated to processing, integrating, and retaining information from the content. Reducing $CL_{extraneous}$ through careful interface design—by implementing features such as content grouping, visual guidance, or controlled pacing—can improve comprehension, alleviate cognitive load, and encourage sustained engagement. Empirical evidence also suggests that interfaces designed to minimize extraneous load not only enhance immediate performance on tasks but also support longer-term cognitive benefits, including better retention and increased attentional capacity^[7].

3. Research Methodology for Critical Analysis of Swipe-Up Interaction in Short-Form Video Platforms

Used heuristically I watch your data my screen numbers and i am entertained We got together 20 people between the ages of 18-30 to interact with pretend short-video platforms that had various swipe-up features. The subjective activity is the session Length,Swipe Speed and Task Duration that was received from the automatic System logs.

3.1 Experimental Setup

In terms of the research we looked at two completely different types of interfaces and saw what effect swiping up would

have on peoples level's engagement, cognitive load and overall satisfaction; A uses the common layout found in many short video apps which is a vertical swiping motion combined with an infinite scroll. It can let people continuously flip through an endless reel of media clips but also be able to respond instantly and smoothly switch contents. While such a framework is more interactive by lowering the effort required to look at the following clip, it can be too much as there are no natural breaks or indications indicating where one piece ends.

Variant B is smooth to swipe up, there's not much content on screen with some graphics that give hints as to where you should press. The media clip gets divided up into different segments based on concept or sequence consistency, there's also little bit of aesthetic stuff going on here like gradients or soft divider to indicate one set is ending and another starting. And also, we will reduce the unnecessary mind load by doing such changes giving structures like an organization and short breaks after every group, better deep thoughts rather than habit scrolling. In this paper I am looking at the two interfaces and trying to figure out what happens when a structure limits a signal and when we look for aesthetics in signals^[8].

3.2 Data Analysis

To gauge how well different interfaces worked for users, we made an overall score showing they did something useful with short videos: To be exact, interaction efficiency E stands for the ratio of active time spent on watching contents compared with all interaction times which can be expressed as :

Formula 2:

$$E = \frac{T_{interaction}}{T_{content}}$$

In this context, $T_{content}$ refers to the time during which users are actively engaged with the content, including watching videos attentively, reading textual information, or meaningfully interacting with interface elements. $T_{interaction}$ represents the total duration of all user interactions with the interface, such as swiping, clicking, pausing, and navigating through various components. A higher value of E reflects that users can consume a larger amount of content with fewer actions, indicating greater efficiency in interface design^[9].

Not going off of just session duration or swiping count as its measures, instead looking at how well people interact and engage cognitively. Look at E variant A beside E variant B purely and neutrally, see if different interfaces making aggregation things by visual distinction using separate visuals affect minds when made simple by less moving around. And then, when added with behavior record and statistic examination, it will find out trend about focus distribution and over-scrolling tending to give evidentiary advice for interface upgrading.

3.3 Heuristic Evaluation

To make a better judgement about each of the interface versions' ease and morality, we used an heuristic method which also includes ethical considerations along with accepted HCI guidelines. This was more particular to collection containing 10 Nielsen heuristic standard like having visible system status, aligning System/Everyday Concept, Autonomy and Choice, Consistency & Conventions, Prevent Errors, Recognised rather than Recalling, Flexible and Efficient, Design Aesthetically Pleasing and Supporting Users Identifying, Correct Mistakes And Having Manuals. Given the likelihood of ethical issues resulting from continuous content absorption and fixation-like behaviour, this contains ethical measures with regards to individual welfare, equitably dispersed focus, inducements via interface action^[10].

Heuristics and morality standards were evaluated on a five level likert scale, 1 being improper execution or no existence and 5 is fully realized. Sum up the separate ratings to form an overall Usability index (UI). It is a number representing both normal usability as well as ethical functioning. By using this comprehensive method it can spot exact interface good and bad points, then make suggestions for design places which may cause tired thoughts or wrong morals^[11]. And since the organized rating took place, those 2 interface differences could be compared by this particular study and see what kind of change in improvement is there for total happiness and a even amount of attraction, being honest with ourselves^[12].

4. Empirical Findings on Swipe-Up Interaction in Short-Form Video Platforms

4.1 Interaction Metrics

Table 1 gives the main indicators used to evaluate swipe up engagement success among the two interface configurations.

In variant A it could be seen that the mean session length was 35.2 minutes which had a mean of 212 swipe action and the engagement ratio E was 0.47. Contrarily, B (Finite Cluster Visuals and Small Visuals as Pause/Focus) was completed with a mean of 28.5 minutes by the conclusion of Stage 176 at an average rate of Engagement of 0.56. This kind of result means that Variant B let the person use material more easily, they needed less swipe for every part and didn't have to do extra thinking^[13].

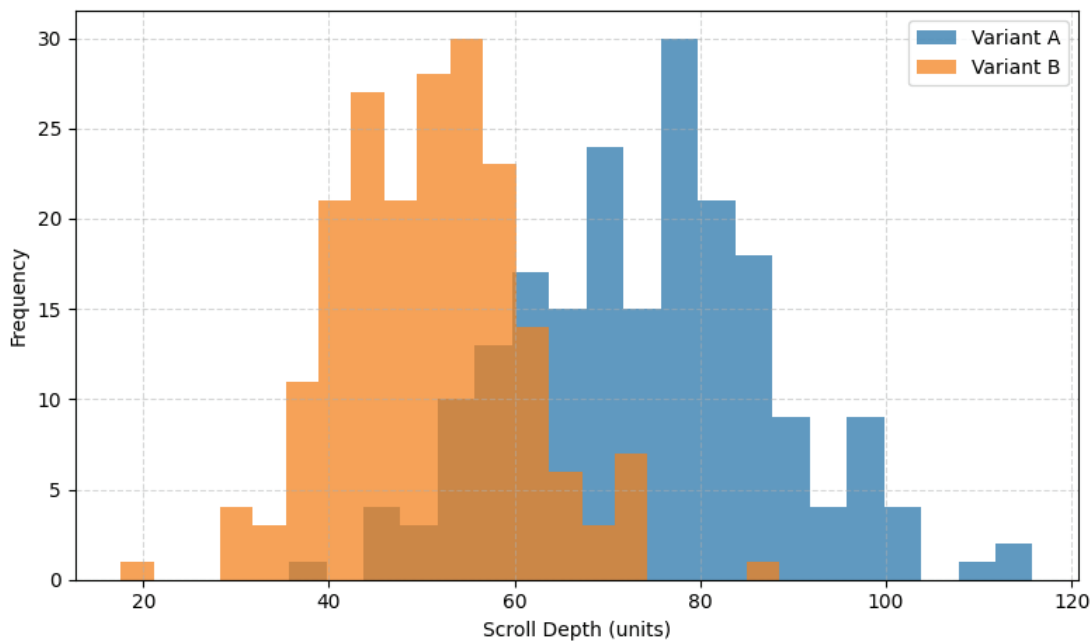
Additionally, it is shown in figure 1 that the scroll heights change between both iterations and also shows how often each interface scrolled. Variant A has more spread out as well due to the usual, unbroken scrolling. On the other hand Variant B has a tighter cluster closer to the mean scroll level meaning that by restricting the scrolling the participants stopped and looked at the content in between the clusters. It is in line with the numerical productivity findings: A better participation and endurance of mind could be made by changing designs^[14].

In summary, if we add browsing time and swipe speed as well as looking at scrolling distance then carefully placing the screen may increase engagement by limited info and making it obvious a warning about habitual scroll. And those are useful pieces of information for customers who care about keeping their minds sharp and healthy.

Table 1. Swipe-Up Efficiency Metrics

Interface Variant	Mean Session Duration (min)	Mean Swipe Frequency	Engagement Efficiency (E)
Variant A	35.2	212	0.47
Variant B	28.5	176	0.56

Figure 1. Scroll depth distribution by interface variant.



4.2 Usability and Ethical Assessment

Interface configuration's practicality and morality has been decided via a complete heuristic and moral model like it is shown within table 2: As for Variant A, it represented the classic infinite scrolling layout which scored decently on regular usability criteria like uniformity(4.1) and reaction time(3.6) but failed when it came to mental strain(2.8)and ethical impact(2.5). It means such findings suggest even if it is still recognizably and coherently designed, it puts more work for the person, and brings about moral anxiety around compulsion and being preoccupied.

Design B had a small amount of grouped content, with some graphics that stopped and it was the best in all aspects: Uniformity was at 4.4, all the answers lining right up against a 4.2 for those wanting things to be free-wheeling. And categorically Mental Workload and Moral Consequence Metrics went up greatly to 4.0 and 4.1, meaning the display removed

all remaining nonessential load on cognition as well as resolved any moral questionings over consumer health.

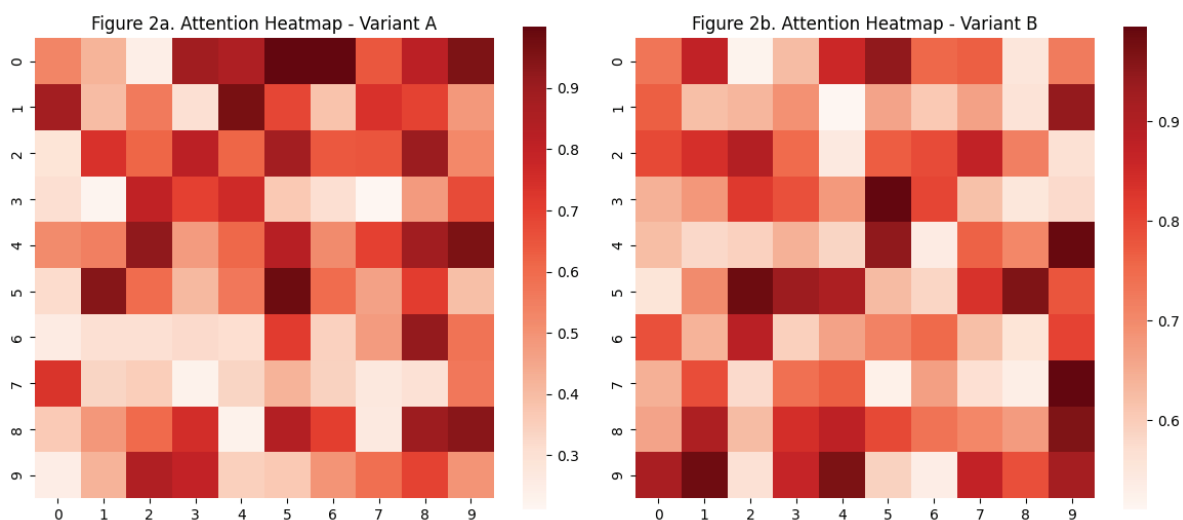
Additionally we have figure 2 which is the viewer focus heat maps over swipe up engagements. Option A shows dense center fixation and almost no peripheral feeling which means the people were too focused on uninterrupted media stream, it will make them tired and less think. But with option B we have a much more even spread of interest onscreen so this means that it must be true that data partitioning and adding in the graphical signal caused people to go for less dumb interactions.

To summarize, combining numbers heuristic metrics and visual attention patterns is pointing out that we should put morality and minds in our user interface development. B has better results showing that the purposeful designs make for much stronger bonds and ethical relationships with your customer^[15].

Table 2. Heuristic Evaluation Scores

Heuristic Criterion	Variant A	Variant B
Consistency	4.1	4.4
Feedback	3.6	4.2
Cognitive Load	2.8	4.0
Ethical Impact	2.5	4.1

Figure 2. Heatmap of Attention Concentration during Swipe-Up Interactions



5. Implications and Conclusions on Swipe-Up Interaction in Short-Form Video Platforms

Within a short video eco system I looked at it closely through the HCI and Ethics theories: The results indicate things like limited content groups and subtle focus tips can cut extra thought down a lot and lessen chances of bad stuff due to feeling forced by looking, though it's still doable and fun. B's architecture had the highest interaction speeds, with lowest cognitions load decrease and balanced attention sharing that proved purposeful interface modifications were able to make people use digital spaces responsibly. Conclusion points out that it's important for reflective design ideas to be combined with interface engineering, so that user health comes first compared to engaging them. For future works, it's about the lasting behavior changes we create when developing these moral, cultural interaction preferences among others for a sustainable short video experience.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

- [1] Chen, M. X., Hu, H., Yao, R., et al. (2024). A survey on the design of virtual reality interaction interfaces. *Sensors*, 24(19), 6204.
- [2] Chen, Y. R., Wang, C. M., & Hsieh, H. L. (2014). Application of interactive design on web interfaces into long-distance health-care for elderly: Example of the website “iShow-Nantou Smart Town”. *Gerontechnology*, 13(2).
- [3] Cheng, X., Su, X., Yang, B., et al. (2023). Understanding users’ negative emotions and continuous usage intention in short video platforms. *Electronic Commerce Research and Applications*, 58, 101244.
- [4] Faudzi, M. A., Che Cob, Z., Ghazali, M., et al. (2024). User interface design in mobile learning applications: Developing and evaluating a questionnaire for measuring learners’ extraneous cognitive load. *Heliyon*, 10(18), e37494.
- [5] Huang, C., & Lee, H. (2023). Interface ethics and digital well-being: Investigating design impacts on user welfare. *Human-Computer Interaction*, 38(2), 125-142.
- [6] Kim, H., Park, S., & Lee, J. (2023). Cognitive load in mobile app navigation. *Computers & Education*, 168, 104202.
- [7] Li, J., & Zhou, R. (2022). Measuring cognitive load in HCI research. *ACM Computing Surveys*, 54(6), 1-37.
- [8] Liu, T., & Li, J. (2023). Research on short video app interface design based on scene interaction theory: Taking TikTok as an example. *Design*, 8(4), 3825-3832.
- [9] Luo, W., Zhao, X., Jiang, B., et al. (2025). Swiping disrupts switching: Preliminary evidence for reduced cue-based preparation following short-form video exposure. *Behavioral Sciences*, 15(8), 1070.
- [10] Montag, C., Sindermann, C., & Elhai, J. D. (2022). Understanding the psychology of TikTok use. *Computers in Human Behavior*, 132, 107258.
- [11] O’Brien, H. L., Toms, E. G., & Johnson, D. (2021). User engagement in digital media: A review. *Information Processing & Management*, 58(3), 102531.
- [12] Sui, M., & Hawkins, I. (in press). Excessive scrolling but attentive news reading? How short video use and awareness of its negative impact shape news engagement across devices. *Media Psychology*.
- [13] Wu, X., Chen, C., & Quan, L. (2024). Visual analysis and interactive interface design of students’ abnormal behavior introducing clustering algorithm. *Technology and Health Care*, 32(6), 4947-4963.
- [14] Xiao, L., Li, X., & Mou, J. (2026). Exploring user engagement behavior with shortform video advertising on short-form video platforms: A visual-audio perspective. *Internet Research*, 36(1), 154-188.
- [15] Yang, L., & Wang, F. (2022). Attention allocation in short video apps. *User Modeling and User-Adapted Interaction*, 31, 123-141.