

Knowledge-based Short Videos as Learner-centered Microlearning: Cognitive Processing, Self-regulation, and Social Presence

Can Zheng*

School of Fine Arts and Media, Jiangsu Second Normal University, Nanjing 210013, China

*Corresponding email: canzheng1228@jssnu.edu.cn

Abstract

Short videos based on knowledge are being used more and more as informal learning materials, but their educational value cannot be understood simply by looking at the number of people who access them or the quality of the content. The key issue is how learners process, regulate, and socially interpret these short audiovisual explanations. This paper presents a theoretical analysis of learning from knowledge-based short videos, with a focus on the learner, by combining cognitive load theory, multimedia learning, self-regulated learning, and social presence. Unlike accounts of platform ecosystems that concentrate on production, governance, and quality control, this paper looks at the mechanisms from the learner's point of view through which short videos become - or do not become - durable knowledge. It poses three questions: How does brevity affect cognitive load? In what way does algorithmic and fragmented viewing interact with self-regulated learning? How does the perception of instructor and peer presence influence persistence and meaning-making? The analysis maintains that short videos are educationally effective when they externalise structure, coordinate verbal and visual information, and prompt activity generated by the learner. In contrast, their pedagogical value is reduced when brevity removes necessary explanation, autoplay discourages deliberate pauses, and popularity cues displace judgments of credibility. A learner-centred framework is suggested in which instructional segmentation, metacognitive prompts, retrieval and transfer tasks, and socially meaningful feedback function as complementary conditions. The paper concludes that short-video learning should be designed and assessed as guided microlearning rather than as passive media consumption.

Keywords

Knowledge-based short videos, Microlearning, Cognitive load, Self-regulated learning, Social presence, Multimedia learning

Introduction

Knowledge-based short videos have now become a common means of accessing education. As people browse through platforms which were initially created for fast-paced audiovisual entertainment, they come across explanations covering subjects such as language, science, health, history, and professional skills. This situation gives rise to a theoretical issue which is different from the question of whether or not short-video platforms offer useful content. The appropriate question to consider is how a learner manages to turn a short, algorithmically supplied audiovisual experience into structured and retrievable knowledge.

The current body of research offers grounds for both optimism and caution. A systematic review of the use of

video in higher education showed that video enhances learning when it is combined with good teaching methods, and the effect is greater when the video is used to supplement rather than just replace existing teaching [1]. Similarly, studies into the design of instructional video identify signaling, segmenting, weeding, and matching modality as principles which assist learners in managing their limited working memory [2]. Nevertheless, the short-video environment introduces circumstances which conventional research on instructional video does not fully take into account. These include rapid switching between content, continuous recommendations, poor curricular sequencing, and variable social cues. They also include

the ability of learners to stop and skip material at their own discretion.

The initial paper in the project examined knowledge-based short videos from the point of view of a communication mechanism. It focused on content production, quality assurance, and the evaluation of learning effectiveness within the platform ecosystem. This paper, however, adopts a different approach. It takes the learner as its main area of analysis and investigates how cognitive processing, self-regulation, and social presence together influence learning from short videos. The aim is not to assert that a single theory can account for the whole phenomenon. Rather, the aim is to combine complementary theories into a mechanism model which indicates when brevity promotes learning and when it results in only shallow recognition without any durable understanding.

The analysis is based on three questions. The first of these is to examine the effect of the brevity associated with short videos on intrinsic, extraneous, and germane cognitive load. The second question is to look at the interaction between algorithmically fragmented viewing and learners' self-regulated learning processes. The third question concerns the influence of both instructor and peer presence on attention, motivation, elaboration, and trust. Instead of testing hypotheses empirically, the paper arrives at answers by means of theoretical synthesis. It describes short-video learning as guided microlearning in so far as the learner is provided with sufficient structure and activity so as to be able to integrate each fragment into a larger knowledge organisation.

Theoretical foundations

Cognitive load and multimedia processing

Cognitive load theory identifies three types of load. Intrinsic load comes from the complexity of the subject matter. Extraneous load results from a poor way of presenting the material. Germane processing involves the construction and automation of schemas. A shorter duration does not, by itself, reduce intrinsic cognitive load. A complicated causal model is still complicated whether it is explained in thirty seconds or thirty minutes. What differs is the amount of explanatory space for coordinating the various elements of the model.

Multimedia learning theory also offers a complementary

explanation. People process verbal information and visual material through partly separate channels. Each channel has a limited capacity. Meaningful learning requires selecting relevant material, organizing it into verbal and pictorial models, and integrating these models with existing knowledge. Research on teacher video coaching suggests that watching and discussing recordings of one's own teaching practice with a coach can support changes in pedagogical behavior and promote reflection on teaching practice [3]. Further evidence from research on captioned videos connects repeated exposure and the use of audiovisual presentations to vocabulary comprehension and working-memory processes [4]. While research into vocabulary learning via TikTok shows that learners find short-video language practice to be useful and easy to access [5]. In short videos, brevity benefits learning only with strong structural cues. A brief explanation with one main point, a clear causal sequence, and a simple example reduces extraneous processing. Conversely, a short explanation packing unfamiliar concepts into rapid narration may increase extraneous load despite its small length.

Video research also shows that visual complexity and audiovisual features have a non-linear effect on engagement [6]. From a learner-centered standpoint, this result should not be taken as an indication that visual stimulation should be maximized. Although novel transitions, dense captions, and fast edits may attract attention, they can compete with the verbal explanation for working-memory resources. The key theoretical distinction is thus between attention capture and attention guidance: The first can start viewing, while the second decides whether cognitive resources are directed towards the intended learning object.

Self-regulated learning in algorithmic environments

Self-regulated learning is made up of learners repeatedly employing forethought, carrying out the learning process, and then engaging in self-reflection. In this process, learners establish goals, choose strategies, keep an eye on their understanding, and modify their behaviour according to the feedback they receive. Short-video platforms make all the stages of this cycle more difficult. The recommendation systems provide content to users even though they have not yet decided on a learning goal.

Autoplay decreases the need for learners to make a deliberate decision to switch over. The number of likes, shares, and completions gives rise to social and behavioural feedback which is simple to observe but only poorly indicates that understanding has taken place. Furthermore, comparative studies into the use of TikTok and YouTube also indicate that perceived usefulness, ease of use, and attitude influence the way in which learners use video-sharing platforms [7]. So platform acceptance becomes a prerequisite for access rather than a sign of having learned.

Research into the gamification of digital platforms indicates that the features of such platforms can have an effect on learners' ability to regulate their own learning, particularly by enhancing motivation and encouraging repeated participation [8]. Yet repeated participation does not mean effective self-regulation; a learner may build a steady habit of watching educational videos without being able to recall, explain, or apply what they have seen. The key issue is not whether the platform promotes activity, but whether that activity is linked with metacognitive information. A worthwhile learning environment should help learners ask: What am I trying to learn? What do I already understand? What is uncertain? And can I apply this idea in a new situation? Reviews of microlearning methods point out advantages in terms of flexibility, motivation, and learner control, but they also caution that separate units may hinder deep conceptual understanding if they are not linked to wider learning sequences [9]. A more recent critical review, which concentrated on TikTok and YouTube Shorts, arrives at the same theoretical conclusion. Although short-video microlearning can help to boost motivation and facilitate initial understanding, complex knowledge must be integrated with more extensive learning arrangements [10]. Therefore, self-regulated learning through short videos requires both learner autonomy and external structure. The learner needs to have control over the pace and the amount of repetition, as well as cues that make stopping, retrieving, and making connections count as legitimate learning activities.

Social presence, trust, and meaning-making

The idea of social presence involves how communication partners are seen as real, available, and

socially significant in a mediated interaction. In the case of knowledge-based short videos, presence is achieved via the use of voice, gaze, gesture, direct address, narrative style, comments, and the perception that one is a member of a learning community. Social cues can lead to greater involvement since learners regard the speaker as talking directly to them rather than simply delivering impersonal information.

Video-based learning provides evidence for the value of social and generative activity. In a study involving university students who made single-take explanatory videos, the process of explaining was found to enhance both immediate memory and transfer when compared with simply writing an explanation [11]. This result is significant not only in the context of learners producing their own video content; it shows that learning is improved when learners take on an active communicative role. Video systems that encourage viewers to predict, explain, respond, or come up with a counterexample can turn social presence from a merely decorative element into a cognitive resource.

Social presence also involves a risk; it is possible for a confident speaker, an attractive visual style, and a high level of comments to create the impression of credibility without in any way improving epistemic reliability. Studies on misinformation indicate that accuracy prompts can direct attention towards truth and reduce the intention to share false content [12]. This has an implication for learning in that social warmth and cognitive trust should be distinguished. A speaker may seem to be present and supportive even if they are giving an incomplete or inaccurate explanation. Learners therefore need social cues which encourage them to examine the material - such as source links, uncertainty indicators, correction practices, and prompts to compare explanations - rather than cues that simply strengthen their identification with the creator.

How short videos shape cognitive processing

Brevity as segmentation, not simplification

The main theoretical assertion of this paper is that brevity should be regarded as segmentation not as simplification. While segmentation breaks the learning process down into manageable sections, simplification involves omitting information or relationships. A short video promotes learning to the extent that it isolates a

single meaningful unit while keeping intact the relationship which makes that unit understandable. It hinders learning when it gets rid of the relationship and is left with only a memorable slogan.

It results in three requirements for design. The first is that each video must include a bounded learning activity - such as defining, distinguishing, predicting, calculating, or applying. The second is that the video should show the relationship between the elements by using signals, spatial organisation, or a concise worked example. The third is that the unit must go beyond itself by means of a question, a prerequisite, or a following step. If this third condition is missing, then segmentation turns into fragmentation. Although evidence-based principles back the first two conditions, microlearning research supports the importance of sequence-level integration.

The extent to which a learner's prior knowledge affects these effects is that learners who are familiar with the subject matter can make up the missing links by means of their existing schemas, while beginners treat the omissions as ambiguity. Since this is the case, video designs should not assume that all viewers require the same level of compression. The use of captions, definitions, and optional expansions can offer a variety of entry points without forcing each learner to go through the same explanation.

Generative processing and retrieval

To watch something is not the same as to learn since the act of recognising it while watching can produce the false impression of having mastered it. In order to achieve generative processing, learners have to choose, organise and integrate information by means of activities such as explaining, predicting, comparing and retrieving. The study using one-take videos shows the value of explanation as an active learning activity. In the case of short videos, similar activities can be included in the interface or associated with the video such as pause-and-predict prompts, one-sentence explanations, low-stakes quizzes and transfer questions.

The order is important: If the prompt comes before the answer, it provides an opportunity for retrieval. But if it comes after a well-explained answer, it might just lead to agreement.

In a similar way, a comment thread can encourage deeper and more meaningful further elaboration when learners are required to justify their answers, but it can become a distraction if it only rewards quick reactions. The core design principle that puts the learner at the centre is therefore to turn passive exposure into a small but clear learning cycle which includes prediction, explanation, feedback and reapplication. Table 1 summarizes the proposed cognitive-processing model.

Table 1. Learner-side cognitive mechanisms in knowledge-based short-video learning.

Learning condition	Primary learner process	Likely benefit	Main failure mode	Design response
One bounded concept	Selective attention and schema activation	Lower extraneous load	Conceptual oversimplification	State the learning action and preserve key relations
Signaled verbal-visual coordination	Dual-channel organization and integration	Better comprehension	Caption or animation competition	Synchronize, signal, and remove decorative detail
Pause, prediction, or retrieval	Generative processing	Stronger retention and transfer	Passive recognition	Place prompts before answers and require explanation
Series-level sequencing	Integration with prior and later knowledge	Cumulative understanding	Fragmented viewing	Provide prerequisites, links, and progression markers
Direct address and feedback	Socially supported persistence	Greater involvement and elaboration	Popularity mistaken for truth	Pair presence with sources, uncertainty, and correction

Self-regulation under algorithmic fragmentation

From feed consumption to goal-directed learning

A platform's feed arranges the amount of exposure according to predicted interest rather than necessarily following the learner's curriculum. This creates a tension between algorithmic convenience and pedagogical coherence. As a result, it reduces the cost of coming across relevant content but shifts the responsibility of deciding what to show to the recommendation system. A learner might start with a real question and then move on to other clips that have been chosen for their novelty. In this situation, the first self-regulatory challenge becomes goal stabilization - that is, the learner has to keep their question going for long enough to avoid being drawn into irrelevant but appealing transitions.

The next difficulty is to keep an eye on the learning process. Watch time and video completion provide readily available signals, but neither shows whether learners can explain the concept without referring to the video. These metrics measure engagement rather than understanding. A self-regulated learner should have the

chance to pause and check their understanding. The third difficulty concerns reflection: After having viewed the material, the learner has to decide whether the idea can be applied to a new problem, linked to another source, or can be used safely in practice.

The fact that these difficulties arise means that interface features should be assessed in terms of their impact on the self-regulation cycle rather than in terms of their ability to extend session length. This reframing shifts attention from engagement metrics to learning processes. Instead of autoplay, a "save for review" function, a clearly visible learning playlist, a reminder to retrieve the material at a later time, or a prompt asking learners to state the question they have not resolved can have a more direct effect on regulation. Although gamification may lead to greater persistence, persistence only becomes educationally significant when it is connected to monitoring and adaptation.

A learner-regulation framework

Table 2 translates the self-regulated learning cycle into short-video design and use.

Table 2. Self-regulated learning cycle for knowledge-based short videos.

Phase	Learner question	Short-video affordance	Evidence of learning
Forethought	What do I need to understand and why?	Learning objective, prerequisite label, searchable series	Goal statement or prediction
Performance	Am I following the explanation?	Segmentation, pause, captions, worked example	Retrieval response during viewing
Monitoring	Can I explain it without looking?	Hide-answer prompt, self-test, compare-and-contrast question	Accurate explanation or distinction
Reflection	Can I use it elsewhere?	Transfer task, follow-up link, correction invitation	Application to a novel case
Adaptation	What should I watch, repeat, or stop?	Progress map, uncertainty tag, optional remedial clip	Revised plan based on performance

The way the framework alters the concept of learner control is that control does not merely mean being able to skip ahead. It means being able to make informed decisions regarding pace, repetition, difficulty, and what the next steps should be. A platform which offers a great deal of choice without providing diagnostic information may appear to promote autonomy. In fact, it leads to dependence, since learners have to rely on popularity signals in order to decide what should be given their attention.

Social presence as a double-edged learning resource

Presence and engagement

Social presence can make short explanations feel more approachable: Direct address reduces psychological distance, a familiar speaker provides continuity across a series, and comments help learners feel comfortable asking questions. This is particularly useful in informal learning situations in which the learner might not have access to a teacher who could answer any follow-up questions. The use of TikTok for educational purposes in

the field of health education shows how short videos can be used to summarise content and aid in the application of complex ideas when they are part of a familiar instructional relationship [13].

Presence ought to be linked with cognitive activity. Although a friendly tone which leads to longer viewing without the need for prediction, retrieval, or application may be motivationally effective, it is still pedagogically incomplete. The main theoretical distinction lies between affective presence and epistemic presence. Affective presence refers to whether the learner feels that they are being accompanied and encouraged. Epistemic presence refers to whether the interaction enables the learner to consider reasons, evidence, boundaries, and alternative explanations. Short-video learning requires both of these aspects, but the second ensures that the first does not amount to nothing more than persuasion.

Community interaction and credibility

By providing an opportunity for clarification, disagreement, and explanation among peers, comment spaces can lengthen the learning episode. Research into learning in social media environments refers to such settings as micro-ecosystems which can boost participation and collaboration but also lead to problems relating to distraction, privacy, and digital literacy [14]. The governance issue is made worse since algorithmic selection is part of broader socio-technical systems whose negative effects cannot be blamed solely on the interface design [15]. These mixed effects are in line with the current framework in that interaction within a community helps self-regulation when it offers informative feedback, but diminishes it when it replaces evidence with popularity.

A learning community that is to be regarded as credible therefore needs norms which make reasoning visible. Learners should be encouraged to cite sources and identify any assumptions they make. They should also be encouraged to tell the difference between uncertainty and error and to explain why an alternative answer does not work. Although accuracy prompts provide a basic form of platform intervention, the more important design aim is epistemic participation. This participation requires deliberate instructional support and practice. This refers to the state in which learners become co-regulators of understanding rather than merely consumers of the creator's reputation.

Integrated model and implications

The analysis results in an integrated model that is learner-centred and based on four related conditions: cognitive manageability, generative activity, self-regulated sequencing, and epistemic social presence. Cognitive manageability makes it possible for the learner to direct their limited working-memory resources towards the concept in question. Generative activity transforms passive viewing into explanation, prediction, retrieval, or transfer. Self-regulated sequencing links the separate clips together into a sequence that is aimed at a specific goal. Epistemic social presence provides both motivation and feedback while at the same time keeping the credibility of the information open to verification.

The conditions are not interchangeable but complementary. A video that is clearly signalled but which does not include retrieval might be understandable yet be quickly forgotten. A retrieval prompt that lacks earlier conceptual clarity might indicate a state of confusion rather than one of learning. A creator who is motivating but who does not make the source transparent might increase confidence without improving accuracy. A well-organized playlist without learner reflection might result in completion without leading to understanding. The model thus predicts that educational value is greatest when the four conditions support one another.

For creators, the model suggests that each short video should be based on a learning action rather than a topic label, employing a single main relationship, including a pre-answer prompt, and offering a source or a link for the next step. For educators, it advises putting the short videos into a sequence of retrieval and transfer activities rather than giving them as separate exposure tasks. For platforms, it recommends that they measure and provide support for behaviours such as pausing, self-testing, explaining, and checking the source, as well as for views and completion. As for learners, it provides a practical rule which is to, after having watched the video, close it and then explain the idea, state its boundaries, and apply it to a new example.

Conclusion

The paper examines knowledge-based short-video learning from a theoretical point of view that is different

from the ones focusing on platform production and quality governance. Its main claim is that the brevity of short videos is not necessarily advantageous and is not inherently superficial. Rather, their educational impact depends on whether brevity serves as structured segmentation or results in information loss.

Theories involving cognitive load and multimedia help to show that it is important to carry out signaling, coordination, and the elimination of irrelevant details. Self-regulated learning demonstrates that information available in an algorithmic form does not automatically lead to goal-directed study. Learners need prompts and affordances for monitoring, retrieval, reflection, and adaptation. These specific supports are absolutely essential for turning algorithmic exposure into purposeful learning. Social-presence theory explains why direct address and interaction within a community can keep people involved, although the difference between affective and epistemic presence stops social warmth from being taken for trustworthy knowledge in most practical settings.

The framework that results views knowledge-based short videos as a form of guided microlearning and shifts the way in which they are assessed from simply counting views and completions to focusing on comprehension, retrieval, transfer, and informed self-regulation. This shift carries significant implications for future research. Future research needs to test the model at different levels of learner expertise. It should also compare autoplay with deliberate stopping and retrieval, and examine whether source-checking and peer explanation improve both calibration and learning. The most effective short-video environment won't be the one which keeps learners watching for the longest period of time, but rather the one that enables them to know when to watch, when to stop, and what to do with what they have seen.

Funding

This work was supported by the General Project of Philosophy and Social Science Research in Colleges and Universities of Jiangsu Province, China (Grant No: 2025SJYB0355).

Acknowledgements

The author would like to show sincere thanks to those techniques who have contributed to this research.

Conflicts of Interest

The author declares no conflict of interest.

References

- [1] Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., Lonsdale, C. (2021) Video improves learning in higher education: a systematic review. *Review of Educational Research*, 91(2), 204-236.
- [2] Mayer, R. E. (2021) Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition*, 10(2), 229.
- [3] van der Linden, S., van der Meij, J., McKenney, S. (2022) Teacher video coaching, from design features to student impacts: A systematic literature review. *Review of Educational Research*, 92(1), 114-165.
- [4] Teng, M. F. (2025) Incidental vocabulary learning from captioned video genres: vocabulary knowledge, comprehension, repetition, and working memory. *Computer Assisted Language Learning*, 38(5-6), 1301-1340.
- [5] Alshreef, N. R., Khadawardi, H. A. (2023) Using TikTok as a tool for English vocabulary learning in the EFL context. *English Language Teaching*, 16(10), 125.
- [6] Xiao, L., Li, X., Mou, J. (2026) Exploring user engagement behavior with short-form video advertising on short-form video platforms: a visual-audio perspective. *Internet Research*, 36(1), 154-188.
- [7] Sun, J., Lee, S. K., Wong, N. (2026) Technology acceptance model in video-sharing platforms: comparisons between TikTok and YouTube. *Frontiers in Computer Science*, 8, 1883472.
- [8] Alhalafawy, W., Zaki, M. T. (2022) How has gamification within digital platforms affected self-regulated learning skills during the COVID-19 pandemic? Mixed-methods research. *International Journal of Emerging Technologies in Learning (IJET)*, 17(6), 123-151.
- [9] Alias, N. F., Razak, R. A. (2025) Revolutionizing learning in the digital age: a systematic literature review of microlearning strategies. *Interactive*

Learning Environments, 33(1), 1-21.

- [10] Kholilah, D., Istiqhfaroh, D. I., Safitri, R. A., Wahidah, I. N., Razaqi, R. S. (2025) Effectiveness of microlearning-based learning media tiktok and youtube shorts: review critical towards knowledge transfer and depth understanding draft. *Journal on Research and Review of Educational Innovation*, 3(3), 18-28.
- [11] Börger, J., Spilles, M., Krull, J., Hagen, T., Hennemann, T. (2023) Fostering university students' learning performance using the one-take video approach. *Active Learning in Higher Education*, 24(3), 305-319.
- [12] Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., Rand, D. G. (2021) Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590-595.
- [13] Holt, N. (2026) Using Tiktok as an educational tool for health students. *Compass: Journal of Learning and Teaching in Higher Education*, 18(2), 23-28.
- [14] Agustin, N., Rahayu, S. (2025) The role of social media as a micro-ecosystem in supporting community-based e-learning platforms: a systematic literature review. *Journal of Technology Informatics and Engineering*, 4(3), 369-390.
- [15] Saurwein, F., Spencer-Smith, C. (2021) Automated trouble: the role of algorithmic selection in harms on social media platforms. *Media and Communication*, 9(4), 222-233.