

RESEARCH REPORT

What Keeps Viewers Watching Short-Form Video?

An Evidence-Graded Synthesis of Hooks, Structure,
Pacing and Endings

Senotel Research

October 2026

Building what comes next.



RESEARCH REPORT

What Keeps Viewers Watching Short-Form Video? An Evidence-Graded Synthesis of Hooks, Structure, Pacing and Endings

Senotel Research

October 2026

ABSTRACT

Short vertical videos are kept or swiped within seconds, yet almost no published research measures retention on 15 to 180 second vertical video directly; the platforms that hold such data publish guidance rather than results. We pose ten research questions that follow a viewer through a video, from the hook to the ending, and answer each by triangulating independent kinds of evidence: randomized headline experiments, zapping and eye-tracking studies of television and online video, viewing logs from millions of videos, psychophysiology and neuroimaging, narrative and film analysis, platform documentation and practitioner knowledge. We checked 76 sources against primary listings and graded each by how directly it measures video viewing; eight widely repeated claims, including the "eight-second attention span", were traced to their origins and did not survive. The answers converge on a value-of-continuing account of the stay-or-leave decision: a viewer keeps watching while the expected value of the next few seconds exceeds that of the next video in the feed, a value raised by open questions, moment-to-moment reward and visible progress, and lowered by processing cost and habituation. We translate the findings into six hook principles, eight retention principles, eleven genre templates and five length bands, and we label each by the strength of its support. We close with the open questions that only controlled experiments on short-form platforms can settle.

Keywords: short-form video; audience retention; hooks; curiosity; narrative structure; editing pace; evidence synthesis

1. Introduction

Short-form video is now one of the main ways people are entertained, informed and persuaded online. YouTube Shorts and Instagram Reels both accept videos of up to three minutes [1, 2], and the systems that decide who sees them reward time watched and videos finished [3–5]. For anyone who makes such videos, every second is a decision with a cost: viewers begin abandoning a slow-starting video within about two seconds [6], and every second after the payoff is a second in which they leave [7].

Advice about how to hold viewers is everywhere and mostly unsourced. The best-known figure, that human attention lasts eight seconds, less than a goldfish's, traces to no study at all (Appendix A). This report asks what the evidence actually supports. It does so for ten questions that follow a viewer from the first second to the last:

RQ1. How is an effective hook constructed?

RQ2. When do viewers decide to stay, and what does retention look like over time?

RQ3. What makes a person continue to watch?

RQ4. How is attention held from the hook to the end?

RQ5. How should different genres of video be structured?

RQ6. How should structure change with video length?

RQ7. How should pacing and visual change be used?

RQ8. How should sound be used?

RQ9. How should text and captions be used?

RQ10. How should a video end, and what do the platforms reward?

Our contributions are (i) a graded synthesis of 76 sources across these questions, with the strength of support stated for every answer; (ii) a short list of claims that fail on inspection; (iii) a value-of-continuing account that ties the separate findings together and makes testable predictions; and (iv) a set of operational principles, templates and length bands that creators, editors and automated editing systems can apply and test.

2. Method

2.1 Routes to an answer

Each question has more than one way in, and no single way is sufficient. Emotion, for example, cannot be observed directly, but its traces can: faces coded second by second while viewers decide whether to keep watching [8], skin conductance while the edit rate is varied [9], the words that hold readers on a page [10], and what people choose to share [11]. We therefore considered eight kinds of evidence for every question (Table 1) and report for each answer which routes were actually available.

Route	What it can establish	What it cannot
Cognitive mechanism	Why a feature works (curiosity, reward, limited capacity)	How large the effect is on video
Randomized field tests	Causal effects at scale, e.g. 22,743 headline A/B tests [12]	Effects on video; most test text
Zapping and eye tracking	Second-by-second staying or leaving on real video [8, 13, 14]	Short-form behaviour; the studies used television and online advertisements
Viewing logs at scale	Real behaviour across millions of videos [6, 15, 16]	Causation; the logs are correlational
Psychophysiology and imaging	Moment-level arousal and attention [9, 17, 18]	Population effects; samples are small
Narrative and film analysis	Structures common to successful stories and films [19–21]	That the structure causes the success
Platform documentation	What ranking systems say they measure [3, 4, 7]	Independence; not peer reviewed
Practitioner knowledge	Hypotheses from people who do the work daily [22]	Proof

Table 1. The kinds of evidence considered for each question.

2.2 Inclusion and verification

A source was included only after it was checked against a primary listing: the publisher's page, a DOI record, PubMed, conference proceedings, arXiv, or the platform's own page. Where only an abstract or a secondary account could be read, this is recorded and the claim is used only as far as that account supports it. Searches were run in English in October 2026. The final set comprises 62 peer-reviewed articles and preprints, 13 platform documents and 1 practitioner document; a further eight popular claims were traced to their origins and rejected (Appendix A). Appendix B lists every source with its design and grade.

2.3 Grading and triangulation

Each source was graded by how directly it answers the question "what keeps a person watching a video": **A**, peer-reviewed and measuring real viewing of video (field data, eye tracking, zapping, watch time, physiology); **B**, peer-reviewed but in an adjacent medium (headlines, text, advertisements, lectures) or about a mechanism; **C**, first-party platform or industry data, not peer reviewed and often self-interested; **D**, practitioner testimony. Figure 1 shows the distribution.



Figure 1. Sources by evidence grade. Sources graded between two levels (for example A/B) are counted at the stronger grade they were assigned for at least one finding.

Each answer is labelled **ESTABLISHED** when independent routes agree and at least one is grade A; **SUPPORTED** when the support is mechanistic or from adjacent media; and **HYPOTHESIS** when it rests on platform or practitioner sources alone or is our own construction.

A LIMIT THAT APPLIES THROUGHOUT

Almost no published research measures retention on 15 to 90 second vertical videos directly. The evidence therefore comes from neighbouring settings (television advertisements with a zap button, YouTube at scale, online lectures, headlines, film) and from mechanisms that do not depend on the screen on which they were measured. Where a number in this report is specific to short-form video, it comes from a platform and is graded C.

3. Findings

RQ1. How is an effective hook constructed? **ESTABLISHED IN PART**

What, in the first seconds, makes a viewer decide to keep watching instead of swiping? Routes taken: the cognitive mechanism of curiosity; story structure; randomized headline experiments; emotion measured second by second; visual salience; tolerance for delay; and the failure case, hooks that win attention and lose trust.

Curiosity is a gap that feels closable. Curiosity arises when attention is drawn to a gap between what one knows and what one wants to know [23]. It requires a reference point (the viewer must know enough to see what is missing) and a gap small enough to close. It engages anticipated-reward circuitry, people spend scarce time and tokens to close it, and information met while curious is remembered better [24, 25].

Three orders produce three feelings. Structural-affect theory distinguishes *suspense* (the initiating event shown, the outcome withheld), *curiosity* (the outcome shown, how or why withheld) and *surprise* (a crucial fact withheld, then revealed) [26, 27]. Studies of spoilers show what opening with the result trades away: one set of experiments found spoiled stories enjoyed more, while a larger study found that spoilers reduced suspense, enjoyment and immersion [28]. Opening with the result therefore works only when it opens a new question that is interesting in itself.

Specific beats vague. The strongest evidence is randomized. Across 22,743 online news A/B tests, together with archival data and a preregistered experiment, titles framed as questions reduced engagement because readers judged them less informative [12]. An earlier field test found that question headlines drew more clicks, particularly those addressed to "you" [29], and most clickbait features raised clicks [30]. Informativeness reconciles the two: a question that already carries specific information differs from "Did you know this?". Readers also judge clickbait as less credible [31], and forward reference ("this", "here's why") is the most common lure in commercial headlines [32].

Surprise captures, joy retains. In the most direct study of the stay-or-leave decision on video, in which viewers' faces were coded while they were free to stop watching, the level of surprise concentrated attention and the rate at which joy rose retained viewers [8]. During television viewing, 72% of gaze shifts go to locations more surprising than average [33].

Faces, action and a single focus. Instagram photographs containing faces were 38% more likely to be liked [34]. Looming and sudden motion capture attention automatically [35]. When viewers' gaze scatters across the frame they are more likely to stop watching, and a logo shown long and centrally at the start drives them away [14]. Advertising guidance from the platforms says the same: start in the action, with people on screen [36], and show the key message within three seconds [37].

Seconds, not minutes. Viewers begin abandoning a video that takes more than two seconds to start, 5.8% more for each further second, and viewers of short content are the least patient [6].

A hook must be paid off. YouTube's own diagnosis of a weak opening is a mismatch between the promise and what follows [7], and practitioners treat the first minute as proof that the promise was honest [22].

Answer. A hook is the shortest statement of a specific, credible promise that opens exactly one gap the video will close. Its type should follow the material (Table 3). Six principles summarise the evidence:

Principle	Evidence
H1 The first spoken word arrives within about half a second; the first frame is a face or the action, never a logo, title card or greeting	[6, 14, 34, 36]
H2 The hook is the most specific, surprising or high-stakes moment available, about twelve words or fewer when spoken	[8, 12, 33]
H3 It names something concrete (a number, an object, a person, an outcome) and opens one question	[12, 23]
H4 It never answers that question itself; the answer comes later in the video, and a promise the video cannot keep is replaced	[7, 22, 31]
H5 Text on screen at the start is short (about eight words), restates the gap specifically, and is neither a vague question nor a summary	[12, 38]
H6 The hook type follows the material: result, stakes, contrarian claim, number, specific question, confession, visual, relevance	[26, 28, 29, 39]

Table 2. Hook principles. Established: informative beats vague; surprise and faces capture attention; delay loses viewers. Not established: the relative strength of hook types on short vertical video.

Type	What it does
result	show the outcome first, withhold how or why (curiosity)
stakes	show the problem or the risk, withhold what happens (suspense)
contrarian	a surprising or counterintuitive claim (surprise)
number	a specific number or a numbered promise (progress)
question	a specific question the video answers; never a vague one
confession	something the speaker admits or reveals
visual	the action or the reaction itself, before any words
relevance	a problem the viewer has, said to them directly

Table 3. Hook types and the effect each relies on.

RQ2. When do viewers decide, and what does retention look like over time? SUPPORTED

Routes taken: network logs of real abandonment [6]; how platforms define their metrics [7, 40]; per-second logs of lecture viewing [15]; YouTube at the scale of millions of videos [16]; short-video recommender research [41–43].

The first decision is made in seconds: abandonment rises with each second a video is slow to start [6], and YouTube Shorts reports whether viewers "stayed to watch" past the opening seconds without publishing the threshold [40]. The second decision is commitment: YouTube measures a video's introduction at 30 seconds and calls it good when more than half of viewers remain [7]. After that the retention curve slopes down, flat where viewers are satisfied, with spikes where they rewatch and dips where they leave [7]. In 862 lecture videos, about three in five interest peaks coincided with a visual transition [15]. Duration alone explains more than 58% of the variance in the share of a YouTube video watched, while total watch time still rises with length [16]. In a large crawl of YouTube viewing, how long people watched was associated with likes per view and the sentiment of comments [44].

Answer. There are two decisions with two deadlines (Figure 2): stay or swipe in the first one to three seconds, which the hook wins; and commitment, which the first real payoff wins. For a short video the first payoff should land within roughly the first quarter and not later than about fifteen seconds. These are design targets derived from the evidence, not published constants; Table 6 sets them per length.

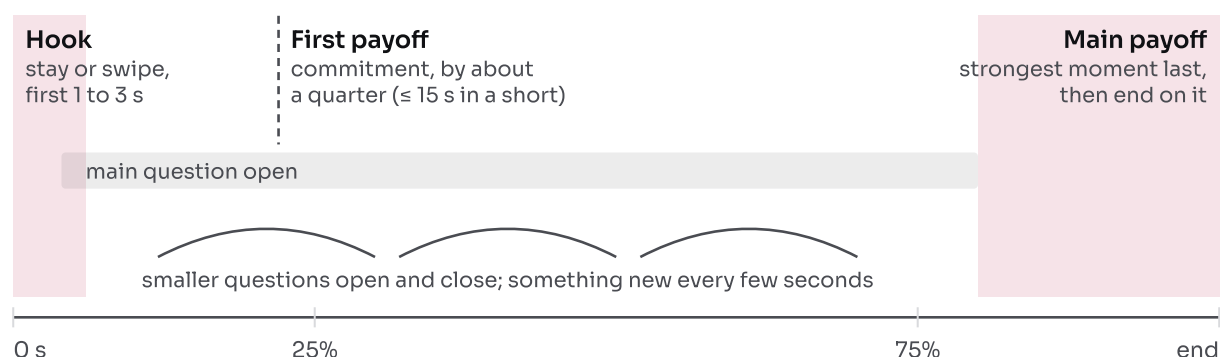


Figure 2. The structure of a short video implied by the evidence: two decisions, a main question that stays open until near the end, smaller questions that open and close along the way, and an ending on the payoff.

RQ3. What makes a person continue to watch? ESTABLISHED MECHANISMS ; COMBINATION IS A HYPOTHESIS

Routes taken: nine mechanisms, each with its own evidence.

- **Entertainment keeps people; unentertaining information pushes them out.** In two zapping experiments, moment-to-moment entertainment lowered the probability of stopping, moment-to-moment selling information raised it, and the two interacted multiplicatively [13].
- **Anticipated reward.** Curiosity recruits reward circuitry and makes people pay time to close the gap [24, 25].
- **The pull to finish.** A 2025 meta-analysis found no memory advantage for unfinished tasks (the popular "Zeigarnik effect"), but a reliable urge to resume them [45]. An open question pulls people forward; it does not make them remember more.
- **Absorption.** Transportation into a story depends on identifiable characters, an imaginable plot and believability [46, 47].
- **Arousal and its direction.** Anxious, exciting and hopeful language holds readers longer and sad language loses them [10]; high-arousal content is shared more [11]; rising joy retains viewers [8].
- **Surprise** attracts the eye [33], but only inside a coherent sequence: scrambling the order of scenes lowers the shared brain response in regions tied to meaning [18].
- **Ease.** Simpler language holds readers longer [10]; messages that exceed processing capacity lose encoding [48–50].
- **Clarity.** Moments that make every viewer's brain respond alike predict what large audiences prefer [17]; tightly edited film produces that synchrony across more than 65% of cortex, an unedited shot under 5% [18].
- **Progress.** People accelerate toward a visible goal and quit less as it nears [39].

Answer. The mechanisms combine into a value-of-continuing account, developed in Section 4: a viewer keeps watching while the expected value of the next few seconds exceeds the value of the next video in the feed.

RQ4. How is attention held from the hook to the end? ESTABLISHED IN PART

Routes taken: open questions [23, 45]; attention responses to edits [9, 15, 48–50]; how stories pace themselves [19, 21, 51]; moment-to-moment value [13, 52, 53]; coherence [18, 54]; rhythm [20]; what to cut and what to keep [13, 55, 56]; long-form practice [22].

- **Questions in relay.** The pull to finish attaches to whatever is open [45]. One main question can stay open until near the end while smaller ones open and close along the way; each closure is a small reward, the rising joy that retained viewers in [8].
- **Change recaptures attention.** Within a scene, faster editing raised arousal and memory without measurable overload [9]; interest peaks sit on visual transitions [15]; the limited-capacity account behind both holds across 142 articles [50].
- **Overload is real.** Fast pacing combined with arousing or complex content lowers how much of the spoken message is encoded [48], and unrelated cuts cost more than related ones [49].
- **Escalate.** Narratives that move slowly while establishing and faster toward the end are evaluated more favourably [19]; cognitive tension peaks in the middle-to-late part of stories [21]; overall judgments favour an improving trend, the peak and the end [51].
- **Cut what has no moment-to-moment value,** because information without entertainment raises stopping [13]. Do not over-cut speech, however: a short silent pause before a word makes it better remembered [55]. One study found that filled pauses ("uh") even improved recall [56], but a 2021 replication found no benefit in English and the opposite in German [57]; the safe conclusion is only that ordinary fillers are not shown to harm comprehension, so the case for removing dead air is pace.
- **Cut where viewers do not notice.** A quarter to a third of cuts within a scene go unseen [58], and viewers perceive a new event when the action changes rather than when the camera does [54].

- **Vary the rhythm.** Shot lengths in successful films have grown more correlated with their neighbours, approaching a 1/f pattern, rather than holding a constant rate [20].

Answer. Eight principles (Table 4).

	Principle	Evidence
R1	Something new arrives every few seconds: information, a reveal, a laugh, a visual change. In a short video, no stretch of more than about eight seconds passes without one	[13, 15, 49]
R2	One main question stays open until the last 20 to 30% of the video; smaller questions open and close along the way	[8, 23, 45]
R3	Beats escalate: weaker material early, the strongest payoff last	[19, 21, 51]
R4	Cut repetition, preamble, tangents, greetings, sign-offs and dead air; keep a short breath before a key word	[13, 55–57]
R5	Cuts within a thought fall on sentence boundaries; cuts intended to re-engage change the content, not merely the angle	[54, 58]
R6	Never stack several new elements in the same second (text, cutaway, sound, zoom); fast cutting only over simple content	[48, 49]
R7	Pace in waves: quicker runs, then a held beat before a payoff	[20, 55]
R8	Visual emphasis is meaningful: a punch-in on the stressed line, a cutaway only when it shows what is being said	[35, 59, 60]

Table 4. Retention principles.

RQ5. How should different genres be structured? SUPPORTED LOGIC PROPORTIONS ARE DEFAULTS

Routes taken: narrative science [19, 21, 26, 46, 47, 61]; instructional design [15, 38, 39, 59, 60, 62]; humour [53, 63]; persuasion and controversy [10, 11, 64]; demonstration and credibility [31, 52]; platform guidance [36, 37].

Three findings carry most of the weight. Viewers of tutorials drop out more than viewers of lectures and jump to the part they need [15], so steps must be easy to find (signalling improves retention of key information [38]) and progress must be visible [39]. Comedy depends on the performer's timing: punchlines in conversation are not preceded by significant pauses [63], so an editor should never add dead air before one, and never separate a setup from its punchline. Opinion content draws more discussion as controversy rises, but only up to a moderate level, beyond which discomfort wins [64]. Across tens of thousands of films, television episodes and other texts, the speed, volume and circuitousness of a story's path through its subject matter predict its success, with different drivers in different forms [65].

Answer. Eleven templates (Table 5), each listing beats in order with the share of the video each usually occupies. The evidence supports the logic of each template; the proportions are design defaults, not measured optima.

Genre	Hook types	Beats in order (share of the video)	Evidence
Story / personal experience	stakes, result, confession	hook (8%): the stakes or the result, in the speaker's words context (10%): who and where, one line rising (40%): complications, each raising the stakes turn (20%): the climax or the reveal resolution (15%): what changed, or the lesson end (7%): end on the payoff or a callback to the hook	suspense/curiosity order [26]; identifiable person + imaginable plot [47]; tension rises late [21]; endings weigh most [51]

Genre	Hook types	Beats in order (share of the video)	Evidence
Tutorial / how-to	result, number, relevance	hook (8%): show or state the finished result and the promise ("3 steps") why (7%): why it matters, one line steps (60%): the steps in order, each one signposted mistake (10%): the common mistake or the trick (a surprise) reveal (15%): the final result, then stop	tutorial viewers drop out and jump to what they need [15]; signal each step [38]; visible progress [39]; key message first [37]
Explainer / educational / Q&A	contrarian, question, number	hook (8%): a counterintuitive fact or a specific question with stakes intuition (25%): the idea in plain words or an analogy example (30%): one concrete example or evidence twist (20%): the misconception or the surprising part takeaway (17%): one line the viewer keeps	curiosity [23–25]; processing ease [10]; signal the key idea [38]; no seductive details [59, 60]
Opinion / commentary / hot take	contrarian, relevance	hook (8%): the claim, bold but short of outrage reasons (45%): reasons in rising strength counter (15%): the best objection, acknowledged strongest (20%): the strongest reason verdict (12%): the verdict, and a real question to the viewer	talk rises with controversy only to a moderate level [64]; arousal holds and spreads [10, 11]; sends and comments rank [5]
Comedy / stand-up / sketch	visual, contrarian, confession	hook (10%): the premise, or the funniest line teased setup (25%): the setup punchlines (50%): punchlines, escalating; never cut between a setup and its punchline button (15%): a callback or the biggest laugh, then stop	humour raises attention [53]; laughs are moment-to-moment value [13]; performers' timing has no dead air before punchlines [63]
List / tips	number, relevance	hook (8%): the number and the specific benefit items (82%): the items, rising in strength, the best last recap (10%): a one-line recap	progress pull [39]; peak and end [51]; signposting [38]
Interview / podcast clip	contrarian, result, stakes	hook (10%): the guest's strongest line who (7%): who they are, one line answer (65%): the answer or the story payoff (18%): the payoff line, then stop	it must stand alone: a gap needs a reference point [23]
Review / product demo	result, contrarian	hook (8%): the verdict or a surprising result demo (40%): the product working, on screen specifics (25%): two or three specifics downside (12%): one honest downside verdict (15%): the verdict	demonstrate before entertaining [52]; credibility [31]
Reaction	visual	hook (10%): the peak reaction, a face trigger (25%): what is being reacted to reactions (50%): the reactions, escalating final (15%): the final reaction	faces engage [34]; rising joy retains [8]
Vlog / day in the life	visual, result, stakes	hook (8%): the most interesting moment of the day journey (75%): in order, time compressed, one new thing per segment end (17%): the peak or a reflection	one new thing per segment [13]; end on the peak [51]
Talk / motivational / speech	contrarian, confession, relevance	hook (10%): the most quotable line story (40%): a story or an example principle (30%): the principle call (20%): the call to act, rising	rising emotion retains and spreads [8, 11]

Table 5. Genre templates. Each beat's share of the video is a design default.

RQ6. How should structure change with length? ESTABLISHED DIRECTION THRESHOLDS ARE DEFAULTS

Routes taken: completion and watch time against length [15, 16, 41, 62]; memory for experiences against duration [51, 66]; what the platforms count [1–3, 42, 43]; story pacing [19]; progress perception [67]; practitioner experience [22].

Completion falls with length while total watch time rises [16, 41]. Platforms that rank on watch time correct for duration [42, 43], so padding a video buys nothing, while TikTok names finishing a longer video as a strong signal of interest [3]. Remembered evaluations neglect duration and weigh peaks and endings [51, 66]: a longer video is not judged better for its length, only for its moments. Engagement with educational videos falls for longer videos [62].

Answer. Choose the shortest length that still holds the hook, the context the payoff needs, and the payoff; every additional beat must earn its place with something new. Table 6 gives the shape of each length band.

Band	Length	First payoff by	Longest stretch without something new	Re-hook midway	Shape
micro	up to 15s	6 s	3 s	no	one idea, one payoff; the hook IS the premise; payoff by about 70%; end on the payoff so the loop restarts cleanly
short	15–45s	10 s	5 s	no	hook within 3s, context within 5s, one or two developments, payoff at 70–90%, end within 2s of it
standard	45–90s	15 s	7 s	no	hook, context, 2–3 beats with a small payoff every 10–15s, the main payoff in the last 20%, a short close
long	90–180s	15 s	8 s	yes	as standard, plus a new question at 40–60% to re-hook, signposts between sections, a small payoff every 15–20s
mid	over 3 minutes	30 s	10 s	yes	the first 30s confirm the promise; chapters with signposts; a re-engagement point every 1–2 minutes; a payoff kept for the end

Table 6. Length bands. The first-payoff and lull targets are design targets derived from RQ2 and RQ4, not published constants.

RQ7. How should pacing and visual change be used? SUPPORTED RATE IS A HYPOTHESIS

Routes taken: edit rate and arousal [9]; pacing with arousing content [48]; limited capacity and its meta-analysis [49, 50]; rhythm in film [20]; unnoticed cuts [58]; event boundaries [54]; interest peaks [15]; looming [35]; irrelevant visuals [59, 60]; platform guidance [36].

Answer. Visual change is a lever on attention whose cost depends on relevance and on how demanding the content already is. Change visuals often over simple content and less over dense explanation [48]. Make cuts within a thought invisible and make attention-grabbing changes meaningful [54, 58]. A cutaway that does not show what is being said is a cost rather than decoration: irrelevant additions reduce what people take away, with a pooled effect of $g = -0.33$ across 58 studies [59, 60]. Use zooms for emphasis, not constantly, or they stop being a signal [35, 49]. Vary the interval between changes [20]. No universal cut rate is supported; the common practice of a visual change every four to six seconds on a talking head remains a hypothesis to be tested.

RQ8. How should sound be used? SUPPORTED

Routes taken: rising intensity and orienting [68, 69]; musical tempo and mode [70]; pauses and fillers [55, 56]; limited capacity [49]; platform survey data [71].

Answer. Speech is the content, and nothing should mask a word that matters [49]. Musical tempo sets energy and mode sets mood [70], so music is best chosen by the story's emotion and pace rather than its topic. Rising sound intensity is perceived as a larger change and raises alertness [68, 69]: a riser into a reveal is grounded in physiology, while a riser into nothing teaches viewers to ignore it. Sound effects should mark moments the eye also sees, sparingly, and count against the budget of R6. Users report that sound matters on TikTok [71], a self-interested survey.

RQ9. How should text and captions be used? ESTABLISHED

Routes taken: a review of more than 100 caption studies [72]; a signalling meta-analysis [38]; irrelevant detail [59, 60]; sound-off viewing surveys and platform tests [37, 73, 74].

Answer. Caption the speech: captions improve comprehension of, attention to and memory for video for most viewers [72], and many people watch in public with the sound off (69% in one survey [74]). Use short phrases synchronised with speech and emphasise one key word [38]; show the hook as text in the first second; avoid decorative text that is not part of the message [59, 60]; and keep text readable, within the five to ten words per second that platform guidance suggests as a ceiling [37].

RQ10. How should a video end, and what do the platforms reward? ESTABLISHED

LOOPS ARE A HYPOTHESIS

Routes taken: peak-end effects and duration neglect [51, 66, 75]; platform ranking documentation [3–5, 7]; sharing [11]; payoff and credibility [31].

What people remember and rate is dominated by the peak and the end, length adds little, and a better ending improves the memory of the whole [51, 66, 75]. YouTube optimises "valued watch time", that is, watch time viewers rate highly in surveys [4]; Instagram's head named watch time, likes per reach and sends per reach as the main signals [5]; TikTok weights finishing [3]. Dips at the end appear where viewers got what they came for [7], and high-arousal moments are shared [11].

Answer. End on the payoff or immediately after it, and cut outros and anything after the punchline. The last line should be the strongest still available, or a callback that resolves the hook. In opinion content, a genuine question to the viewer serves better than a request to follow. A loop-friendly ending, in which the last line leads back into the first, is a hypothesis: rewatching counts on all three platforms, but none publishes how.

4. A value-of-continuing account

The findings of RQ1 to RQ10 are not independent rules; they describe one decision made repeatedly. At every moment a viewer of a feed can stay or move on, and the next video is always one gesture away. We propose that the viewer stays while the expected value of the next few seconds exceeds the value of that alternative, and that this value has five components, each separately supported:

$\text{value}(t) = \text{anticipated reward of open questions [23–25, 45]} + \text{value of the present moment: news, emotion, laughter, surprise [8, 13, 33, 53]} + \text{pull of visible progress [39, 67]} - \text{processing cost [10, 48–50]} - \text{habituation since anything changed [15, 49]}$

The account explains why the principles take the form they do. Hooks matter disproportionately because the prior hazard of leaving is highest in the first seconds [6] and only an open question can raise value before any content has been delivered (H1 to H6). The first payoff matters because anticipated reward decays unless it is confirmed (RQ2). Cutting dead air, greetings and repetition removes moments of near-zero value (R4). Escalation and relayed questions keep anticipated reward from decaying (R2, R3). Overload principles bound the cost term (R6). And endings matter beyond their share of the running time because memory weights the end (RQ10).

The account is a hypothesis about how the components combine, and it makes testable predictions. It predicts that (i) retention curves on short-form platforms should drop most steeply in the first one to three seconds and again after the main payoff; (ii) moving the first payoff earlier should flatten the curve after it, holding content constant; (iii) adding a new open question at the midpoint of a long short-form video should reduce mid-video loss; and (iv) cutaways unrelated to the speech should increase loss in the seconds that follow, relative to related cutaways. A companion report describes a computational version of this account and its use in automated editing [76].

5. Discussion

What the evidence settles. Three things about openings are as well supported as anything in this field: a specific promise outperforms a vague one, a slow start loses viewers by the second, and surprise and faces capture the eye. Two things about the middle are well supported: entertainment keeps people while unentertaining information pushes them away, and coherence is what allows a change of picture to recapture attention rather than cost it. About endings, the peak and the final moment dominate what people remember.

What it does not settle. Much common advice, including universal cut rates, rankings of hook types, loop endings and the exact proportions of genre templates, is plausible and unmeasured on short vertical video. We have labelled these as hypotheses rather than dropping them, because they are useful defaults; but they should be treated as defaults to be tested.

Where the evidence disagrees. Question headlines won clicks in one field test [29] and lost engagement across thousands of randomized tests [12]; spoilers raised enjoyment in one set of experiments and lowered it in a larger study [28]; filled pauses aided recall in one study and not in its replication [56, 57]. In each case we report both and adopt the reading that survives both: informative questions, result-first openings that open a new question, and no claim that fillers help.

A research agenda. The open questions can be settled only with controlled experiments on short-form platforms: randomized variants of the same video differing in one element (hook type, first-payoff timing, cut rate, cutaway relevance, ending), with retention curves as the outcome. The predictions of Section 4 give the first such experiments.

6. Limitations

This is a structured evidence review, not a registered systematic review: the search was not exhaustive, it was limited to English-language sources, and grading was done by a single team. Most evidence comes from settings adjacent to short-form video, so effect sizes may not transfer. Platform documents are self-interested and can change without notice. Several findings rest on single studies or small samples, which we have marked. Finally, the value-of-continuing account is a synthesis of our own; its components are supported, its combination is not yet tested.

7. Conclusion

What keeps viewers watching is not a secret formula but a small set of well-supported mechanisms: curiosity that feels answerable, a reward that arrives early and keeps arriving, a story clear enough to follow, change that means something, and an ending that lands. The evidence is strongest where it has been tested at scale and weakest exactly where short-form platforms hold the data. We offer the principles here as a starting point and the predictions as an invitation to test them.

Disclosure. Senotel develops automated video-editing tools, and this review informed that work. No external funding was received. The literature search, synthesis and drafting were carried out with the assistance of an AI system (Claude, by Anthropic) under the direction of Senotel; every source was checked against a primary listing, and no reference was taken from the system's memory alone.

References

1. YouTube. (2024, October). Shorts of up to three minutes. YouTube Official Blog announcement.
2. Instagram. (2025, January 20). Reels of up to three minutes. Instagram announcement.
3. TikTok. (2020, June 18). How TikTok recommends videos #ForYou. TikTok Newsroom. newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you
4. Goodrow, C. (2021, September 15). On YouTube's recommendation system. YouTube Official Blog. blog.youtube/inside-youtube/on-youtubes-recommendation-system/
5. Mosseri, A. (2025, January). Statements on Instagram's ranking signals, as reported by *Social Media Today*.
6. Krishnan, S. S., & Sitaraman, R. K. (2012). Video stream quality impacts viewer behavior: Inferring causality using quasi-experimental designs. In *Proceedings of the ACM Internet Measurement Conference (IMC '12)*. ACM.
7. YouTube. (n.d.). Measure key moments for audience retention. YouTube Help. support.google.com/youtube/answer/9314415
8. Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, 49(2), 144–159. <https://doi.org/10.1509/jmr.10.0207>
9. Lang, A., Zhou, S., Schwartz, N., Bolls, P. D., & Potter, R. F. (2000). The effects of edits on arousal, attention, and memory for television messages: When an edit is an edit can an edit be too much? *Journal of Broadcasting & Electronic Media*, 44(1), 94–109.
10. Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. *Journal of Marketing*, 87(5), 793–809. <https://doi.org/10.1177/00222429231152880>
11. Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192–205. <https://doi.org/10.1509/jmr.10.0353>
12. Fang, D., & Wheeler, S. C. (2026). Titles framed as questions reduce reader engagement. *Journal of Consumer Psychology*. Advance online publication. <https://doi.org/10.1002/jcpy.70031>
13. Woltman Elpers, J. L. C. M., Wedel, M., & Pieters, R. G. M. (2003). Why do consumers stop viewing television commercials? Two experiments on the influence of moment-to-moment entertainment and information value. *Journal of Marketing Research*, 40(4), 437–453.
14. Teixeira, T., Wedel, M., & Pieters, R. (2010). Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Science*, 29(5), 783–804. <https://doi.org/10.1287/mksc.1100.0567>
15. Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the First ACM Conference on Learning @ Scale* (pp. 31–40). ACM.
16. Wu, S., Rizioiu, M.-A., & Xie, L. (2018). Beyond views: Measuring and predicting engagement in online videos. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 12(1). arXiv:1709.02541
17. Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., & Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, 5, 4567. <https://doi.org/10.1038/ncomms5567>
18. Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2(1), 1–26.

19. Laurino Dos Santos, H., & Berger, J. (2022). The speed of stories: Semantic progression and narrative success. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001171>
20. Cutting, J. E., DeLong, J. E., & Nothelfer, C. E. (2010). Attention and the evolution of Hollywood film. *Psychological Science*, 21, 440–447. <https://doi.org/10.1177/0956797610361679>
21. Boyd, R. L., Blackburn, K. G., & Pennebaker, J. W. (2020). The narrative arc: Revealing core narrative structures through text analysis. *Science Advances*, 6(32), eaba2196. <https://doi.org/10.1126/sciadv.aba2196>
22. *How to succeed at MrBeast production*. (c. 2022). [Internal production handbook, made public in September 2024; its authenticity was confirmed to the press by former producers, not by the company].
23. Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1), 75–98. <https://doi.org/10.1037/0033-2909.116.1.75>
24. Kang, M. J., Hsu, M., Krajovich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, 20(8), 963–973. <https://doi.org/10.1111/j.1467-9280.2009.02402.x>
25. Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, 84(2), 486–496. <https://doi.org/10.1016/j.neuron.2014.08.060>
26. Brewer, W. F., & Lichtenstein, E. H. (1982). Stories are to entertain: A structural-affect theory of stories. *Journal of Pragmatics*, 6, 473–486.
27. Hoeken, H., & van Vliet, M. (2000). Suspense, curiosity, and surprise: How discourse structure influences the affective and cognitive processing of a story. *Poetics*, 27(4), 277–286. [https://doi.org/10.1016/S0304-422X\(99\)00021-2](https://doi.org/10.1016/S0304-422X(99)00021-2)
28. Leavitt, J. D., & Christenfeld, N. J. S. (2011). Story spoilers don't spoil stories. *Psychological Science*, 22(9), 1152–1154. <https://doi.org/10.1177/0956797611417007>
- Contested by Johnson, B. K., & Rosenbaum, J. E. (2015). Spoiler alert: Consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation. *Communication Research*, 42(8), 1068–1088.
29. Lai, L., & Farbroth, A. (2014). What makes you click? The effect of question headlines on readership in computer-mediated communication. *Social Influence*, 9(4), 289–299. <https://doi.org/10.1080/15534510.2013.847859>
30. Kuiken, J., Schuth, A., Spitters, M., & Marx, M. (2017). Effective headlines of newspaper articles in a digital environment. *Digital Journalism*, 5(10), 1300–1314. <https://doi.org/10.1080/21670811.2017.1279978>
31. Molyneux, L., & Coddington, M. (2020). Aggregation, clickbait and their effect on perceptions of journalistic credibility and quality. *Journalism Practice*, 14(4), 429–446. <https://doi.org/10.1080/17512786.2019.1628658>
32. Blom, J. N., & Hansen, K. R. (2015). Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, 76, 87–100. <https://doi.org/10.1016/j.pragma.2014.11.010>
33. Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, 49(10), 1295–1306.
34. Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces engage us: Photos with faces attract more likes and comments on Instagram. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM.
35. Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999–1010.
36. Google, & Kantar. (2021). The short and the long of ABCDs effectiveness. Think with Google.
37. TikTok. (n.d.). Auction ads creative tips. TikTok Ads Help Center.
38. Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, 23, 1–24.
39. Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusory goal progress, and customer retention. *Journal of Marketing Research*, 43(1), 39–58.
40. YouTube. (n.d.). Shorts analytics: Viewed versus swiped away. YouTube Studio Help.
41. Quan, Y., Ding, J., Gao, C., Li, N., Yi, L., Jin, D., & Li, Y. (2023). Alleviating video-length effect for micro-video recommendation. *ACM Transactions on Information Systems*. arXiv:2308.14276
42. Zhan, R., Pei, C., Su, Q., Wen, J., Wang, X., Mu, G., Zheng, D., & Jiang, P. (2022). Deconfounding duration bias in watch-time prediction for video recommendation. In *Proceedings of KDD 2022*. arXiv:2206.06003
43. Zhao, H., Cai, G., Zhu, J., Dong, Z., Xu, J., & Wen, J.-R. (2024). Counteracting duration bias in video recommendation via counterfactual watch time. In *Proceedings of KDD 2024*. arXiv:2406.07932
44. Park, M., Naaman, M., & Berger, J. (2016). A data-driven study of view duration on YouTube. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 10(1), 651–654. <https://doi.org/10.1609/icwsml.v10i1.14781>

45. Ghibellini, R., & Meier, B. (2025). Interruption, recall and resumption: A meta-analysis of the Zeigarnik and Ovsiankina effects. *Humanities and Social Sciences Communications*, 12. <https://doi.org/10.1057/s41599-025-05000-w>
46. Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
47. van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, 40(5), 797–817. <https://doi.org/10.1086/673383>
48. 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. *Journal of Broadcasting & Electronic Media*, 43(4), 451–475.
49. Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
50. Huskey, R., Wilcox, S., Clayton, R. B., & Keene, J. R. (2020). The limited capacity model of motivated mediated message processing: Meta-analytically summarizing two decades of research. *Annals of the International Communication Association*, 44(4), 322–349.
51. Baumgartner, H., Sujan, M., & Padgett, D. (1997). Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research*, 34(2), 219–232.
52. Teixeira, T., Picard, R., & el Kaliouby, R. (2014). Why, when, and how much to entertain consumers in advertisements? A web-based facial tracking field study. *Marketing Science*, 33(6), 809–827. <https://doi.org/10.1287/mksc.2014.0854>
53. Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science*, 37(2), 191–203. <https://doi.org/10.1007/s11747-008-0096-y>
54. Magliano, J. P., & Zacks, J. M. (2011). The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*, 35(8), 1489–1517. <https://doi.org/10.1111/j.1551-6709.2011.01202.x>
55. MacGregor, L. J., Corley, M., & Donaldson, D. I. (2010). Listening to the sound of silence: Disfluent silent pauses in speech have consequences for listeners. *Neuropsychologia*, 48(14), 3982–3992. <https://doi.org/10.1016/j.neuropsychologia.2010.09.024>
56. Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*, 65(2), 161–175. <https://doi.org/10.1016/j.jml.2011.03.004>
57. Muhlack, B., Elmers, M., Drenhaus, H., Trouvain, J., van Os, M., Werner, R., Ryzhova, M., & Möbius, B. (2021). Revisiting recall effects of filler particles in German and English. In *Proceedings of Interspeech 2021* (pp. 3979–3983). <https://doi.org/10.21437/Interspeech.2021-1056>
58. Smith, T. J., & Henderson, J. M. (2008). Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *Journal of Eye Movement Research*, 2(2), 1–17. <https://doi.org/10.16910/jemr.2.2.6>
59. Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, 7(3), 216–237. <https://doi.org/10.1016/j.edurev.2012.05.003>
60. Sundararajan, N. K., & Adesope, O. (2020). Keep it coherent: A meta-analysis of the seductive details effect. *Educational Psychology Review*, 32(3), 707–734. <https://doi.org/10.1007/s10648-020-09522-4>
61. Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5, 31. <https://doi.org/10.1140/epjds/s13688-016-0093-1>
62. Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning @ Scale* (pp. 41–50). ACM. <https://doi.org/10.1145/2556325.2566239>
63. Attardo, S., Pickering, L., & Baker, A. (2011). Prosodic and multimodal markers of humor in conversation. *Pragmatics & Cognition*, 19(2), 224–247. <https://doi.org/10.1075/pc.19.2.03att>
64. Chen, Z., & Berger, J. (2013). When, why, and how controversy causes conversation. *Journal of Consumer Research*, 40(3), 580–593. <https://doi.org/10.1086/671465>
65. Toubia, O., Berger, J., & Eliashberg, J. (2021). How quantifying the shape of stories predicts their success. *Proceedings of the National Academy of Sciences*, 118(26), e2011695118. <https://doi.org/10.1073/pnas.2011695118>
66. Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, 65(1), 45–55. <https://doi.org/10.1037/0022-3514.65.1.45>
67. Conrad, F. G., Couper, M. P., Tourangeau, R., & Peytchev, A. (2010). The impact of progress indicators on task completion. *Interacting with Computers*, 22(5), 417–427. <https://doi.org/10.1016/j.intcom.2010.03.001>
68. Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395, 123–124.

69. Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., Herdener, M., Scheffler, K., & Seifritz, E. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cerebral Cortex*, 18(1), 145–150. <https://doi.org/10.1093/cercor/bhm040>
70. Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, 20(2), 151–171.
71. TikTok, & Kantar. (2021). Evolution of sound. TikTok for Business.
72. Gernsbacher, M. A. (2015). Video captions benefit everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 195–202. <https://doi.org/10.1177/2372732215602130>
73. Facebook. (2016, February). Internal tests of captioned video ads, as reported by *Adweek*.
74. Verizon Media, & Publicis Media. (2019). Survey of 5,616 U.S. adults on viewing video with sound off.
75. Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, 4(6), 401–405.
76. Senotel Research. (2026). *From evidence to edits: Rule-guided language-model video editing and a small on-device student model*. Senotel research report.
77. Microsoft Canada. (2015). *Attention spans* [Consumer insights report].
78. BBC Radio 4. (2017). *More or Less*: the goldfish attention-span claim.
79. Facebook IQ, & Nielsen. (2016). Brand effect analysis of 173 video campaigns. Reported on Facebook's newsroom.
80. Mark, G. (2023). *Attention span: A groundbreaking way to restore balance, happiness and productivity*. Hanover Square Press.

Appendix A. Claims that did not survive checking

Claim	What the record shows
Human attention lasts eight seconds, less than a goldfish's	Traced to an unsourced figure reproduced in a 2015 consumer report [77] and examined by BBC Radio 4 [78]; no study measures a single "attention span" in this way
47% of a video's value lies in its first three seconds	The figure is the share of advertising-recall lift among people who watched for under three seconds in brand-lift studies [79], not value and not retention
Attention on screens lasts 47 seconds	An average time on one screen before switching during office work [80], not a limit on attention to a video
Open loops work because unfinished things are remembered (the Zeigarnik effect)	No memory advantage in a 2025 meta-analysis; the reliable effect is the urge to resume [45]
Always add a dramatic pause before a punchline	Punchlines in conversation are not preceded by significant pauses [63]
Question hooks always win	Mixed: they won clicks in one field test [29] and reduced engagement across 22,743 A/B tests [12]
Captions increase completion by 80%	80% of survey respondents <i>said</i> captions make them more likely to finish [74]; this is not a measured effect
Fillers such as "um" make speech more memorable	One positive study [56] was not replicated [57]

Appendix B. Evidence table

Ref.	Source	Design	Grade
[1]	YouTube (2024)	Platform or industry document	C
[2]	Instagram (2025)	Platform or industry document	C
[3]	TikTok (2020)	Platform or industry document	C
[4]	Goodrow (2021)	Platform or industry document	C
[5]	Mosseri (2025)	Platform or industry document	C (secondary report of a primary statement)
[6]	Krishnan & Sitaraman (2012)	23 million views, 6.7 million viewers (Akamai), quasi-experimental	A
[7]	YouTube (n.d.)	Platform or industry document	C
[8]	Teixeira et al. (2012)	Lab: automated facial-expression coding + eye tracking + the option to zap (stop watching), 58 adults, 28 ads	A
[9]	Lang et al. (2000)	Lab, within-subject, edit rate varied slow to very fast (edits = camera changes within one scene)	A
[10]	Berger et al. (2023)	NLP over 600,000+ reading sessions of 35,000+ articles on nine sites, plus experiments	B (text, not video)
[11]	Berger & Milkman (2012)	Three months of New York Times articles + experiments	B
[12]	Fang & Wheeler (2026)	Archival Reddit (N = 53,030) and academic-article (N = 3,078,791) data, 22,743 online news A/B tests, a preregistered experiment (N = 400)	A (very large, includes randomized A/B tests)

Ref.	Source	Design	Grade
[13]	Woltman Elpers et al. (2003)	Two experiments, moment-to-moment ratings and zapping	A
[14]	Teixeira et al. (2010)	Eye tracking + zapping, 31 commercials, ~2,000 viewers; model + controlled experiment	A
[15]	Kim et al. (2014)	Per-second logs, 862 edX videos	A
[16]	Wu et al. (2018)	5,331,204 YouTube videos published Jul–Aug 2016	A
[17]	Dmochowski et al. (2014)	EEG inter-subject correlation in a small sample, compared with large-audience ratings and social-media activity	A
[18]	Hasson et al. (2008)	fMRI inter-subject correlation	A (small samples)
[19]	Laurino Dos Santos & Berger (2022)	40,000+ movie scripts; replicated on 10,000+ TV episodes	B
[20]	Cutting et al. (2010)	150 films 1935–2005	A (descriptive)
[21]	Boyd et al. (2020)	~40,000 narratives, ~20,000+ factual texts	B
[22]	Practitioner handbook (c. 2022)	Practitioner document	D
[23]	Loewenstein (1994)	Theory review	B
[24]	Kang et al. (2009)	fMRI + behavioural (trivia)	B
[25]	Gruber et al. (2014)	fMRI	B
[26]	Brewer & Lichtenstein (1982)	Theory + experiments	B
[27]	Hoeken & van Vliet (2000)	Experiment	B
[28]	Leavitt & Christenfeld (2011)	Three experiments, 12 short stories	B, mixed
[29]	Lai & Farbroth (2014)	Field tests on Twitter and a shopping site	B
[30]	Kuiken et al. (2017)	1,638 rewritten headlines with click-through rates (Blendle)	B
[31]	Molyneux & Coddington (2020)	Two online experiments	B
[32]	Blom & Hansen (2015)	Corpus analysis, 100,000 headlines from 10 Danish news sites	B
[33]	Itti & Baldi (2009)	Eye tracking during TV viewing	A
[34]	Bakhshi et al. (2014)	~1.1 million Instagram photos	A (correlational, platform scale)
[35]	Franconeri & Simons (2003)	Visual search	B
[36]	Google & Kantar (2021)	Platform or industry document	C (modelled association, self-interested)
[37]	TikTok (n.d.)	Platform or industry document	C (correlational, ads)
[38]	Schneider et al. (2018)	Meta-analysis, 103 studies, N = 12,201	A (learning)
[39]	Kivetz et al. (2006)	Field data and experiments	B
[40]	YouTube (n.d.)	Platform or industry document	C
[41]	Quan et al. (2023)	Micro-video logs	A (platform logs)
[42]	Zhan et al. (2022)	Industrial recommender research	C/A (company-authored)

Ref.	Source	Design	Grade
[43]	Zhao et al. (2024)	Industrial recommender research	C/A (company-authored)
[44]	Park et al. (2016)	Random YouTube crawl + Chrome-extension viewing logs	A
[45]	Ghibellini & Meier (2025)	Meta-analysis	A
[46]	Green & Brock (2000)	Four experiments (N = 97, 69, 274, 258)	B
[47]	van Laer et al. (2014)	Meta-analysis, 132 effect sizes, 76 articles	A (meta-analysis; mostly text/ads)
[48]	Lang et al. (1999)	Lab	A
[49]	Lang (2000)	Theory	B (theory)
[50]	Huskey et al. (2020)	Meta-analysis, 142 articles, 683 effects	A
[51]	Baumgartner et al. (1997)	Experiments with moment-to-moment affect traces of ads	A
[52]	Teixeira et al. (2014)	Web-based facial tracking field study, 82 ads	A
[53]	Eisend (2009)	Meta-analysis (38 studies after screening)	A (ads)
[54]	Magliano & Zacks (2011)	Behavioural segmentation + fMRI	A
[55]	MacGregor et al. (2010)	ERP + surprise recognition test	B
[56]	Fraundorf & Watson (2011)	Story retelling with natural speech	B
[57]	Muhlack et al. (2021)	Replication in German and English	B
[58]	Smith & Henderson (2008)	Eye tracking while detecting cuts in feature films	A
[59]	Rey (2012)	Review and meta-analysis	A (learning)
[60]	Sundararajan & Adesope (2020)	Meta-analysis, 58 studies	A (learning)
[61]	Reagan et al. (2016)	Sentiment trajectories of 1,327 Project Gutenberg books	B
[62]	Guo et al. (2014)	6.9 million viewing sessions, 4 edX courses	A
[63]	Attardo et al. (2011)	Acoustic analysis	B
[64]	Chen & Berger (2013)	Field data + four experiments	B
[65]	Toubia et al. (2021)	Tens of thousands of movies, TV shows, papers; semantic trajectories	B
[66]	Fredrickson & Kahneman (1993)	Film clips with real-time ratings (N = 32; N = 96)	A (film clips)
[67]	Conrad et al. (2010)	Web-survey experiments	B
[68]	Neuhoff (1998)	Psychophysics; fMRI + autonomic measures	B
[69]	Bach et al. (2008)	Psychophysics; fMRI + autonomic measures	B
[70]	Husain et al. (2002)	Experiment, one Mozart sonata in four versions	B
[71]	TikTok & Kantar (2021)	Platform or industry document	C (survey, self-interested)
[72]	Gernsbacher (2015)	Review of 100+ studies	A (review)
[73]	Facebook (2016)	Platform or industry document	C

Ref.	Source	Design	Grade
[74]	Verizon Media & Publicis Media (2019)	Platform or industry document	C (self-report)
[75]	Kahneman et al. (1993)	Cold-pressor experiment + colonoscopy field data	B
[79]	Facebook IQ & Nielsen (2016)	Platform or industry document	C, self-interested

Table B1. Every source cited, with its design and grade. Numbers match the reference list.