1. Problem Statement:

1.1 Identifying the Exact Problem

On paper, YouTube is doing great. Shorts are getting over 200 billion views a day. Ad revenue crossed $36 billion in 2024. The platform beats Netflix in US TV watch time. So what exactly is the problem?

The problem is hiding inside those numbers. And if we look at the wrong thing, we end up building the wrong solution.This section breaks down what is actually broken, why it will not fix itself on its own, what it is costing the business, and how we are defining the problem for everything that follows.

1.2 Everyone Is Looking at the Wrong Thing

The way most people frame this problem is: users prefer Shorts now, so long-form videos are getting fewer views. That is true. But it is the kind of truth that leads you to the wrong answer.
If the problem is that users prefer Shorts, then the solution is better content or hoping creators figure it out. That is not a product problem, and it is not something a PM can solve.

Here is what is actually happening.

Most YouTube users do not only want Shorts. They want both formats. The same person who scrolls through Shorts on their morning commute would happily watch a 40-minute documentary on a Sunday evening. But the algorithm cannot tell when a user is in which mood. And when it is not sure, it defaults to Shorts, because Shorts produce a faster and cleaner engagement signal. That default is the real problem.
This is a product architecture problem, not a user preference problem.
The data makes this hard to ignore:

- Shorts views grew from 70 billion per day in early 2024 to over 200 billion per day by 2025. That is a 186% jump in about a year.
- The average length of YouTube's most popular videos dropped by 21% between December 2024 and May 2025, going from around 35 minutes to 28 minutes.
- After YouTube pushed an algorithm update in January 2025, ultra-long-form viewership dropped by 90% in just three months.

That last number is the one worth pausing on. Viewership of long videos did not gradually go down. It collapsed almost immediately after a platform-level algorithm change. Users did not suddenly stop wanting long content. YouTube stopped showing it to them.

One more thing worth noting. The Shorts algorithm and the long-form algorithm are built on completely different systems and use completely different signals. Shorts are ranked on swipe-through rate, replays, and early engagement in the first few seconds. Long-form videos are ranked on watch time, satisfaction, and session continuation. These two systems do not talk to each other in any reliable way. So a user who spends an hour watching Shorts about cooking may or may not get long-form cooking videos recommended, and when that cross-recommendation does happen, it is driven more by luck than by design. For users who genuinely want both formats, the platform currently has no consistent bridge between the two sides.

1.3 Why This Problem Is Getting Worse, Not Better

This problem is not going to balance out on its own. AI tools have made Shorts so cheap and fast to produce that a creator who once spent a week on one long video can now make 20 Shorts in the same time. Over 12 million Shorts are uploaded every day, and creator channels posting Shorts grew 50% year over year in 2024. The pool the algorithm picks from is getting flooded with short-form content, and long-form keeps losing ground.

This creates a loop that only goes in one direction. More Shorts get served, more Shorts get clicked, the algorithm learns that Shorts are what people want, and serves even more of them. Creators notice their long videos getting fewer views and start shifting to Shorts too. As that happens across thousands of creators, the long-form library gets thinner, which makes the problem harder to fix over time.

YouTube's own product decisions have pushed in the same direction: extending Shorts to 3 minutes in October 2024, launching AI tools that turn raw footage into polished Shorts, and closing the Trending page in July 2025 in favour of purely algorithmic charts.

The problem will not self-correct. Without a deliberate product intervention, the gap will keep widening.

1.4 What This Is Actually Costing YouTube

This is not just a user experience problem. It is a business problem showing up in three different places at the same time.The money gap is large and it affects everyone.
Shorts and long-form videos earn completely differently. Here is what the numbers look like side by side:

Metric	Long-form	Shorts
Creator earnings per 1,000 views (RPM)	$1 - $10+	$0.05 - $0.07
Earnings on 3 million views	~$3,810	~$180
YouTube's own cut	YouTube takes 45%	YouTube takes 55%
Ad slots per view	multiple mid-roll ads possible	pooled revenue, minimal slots
Real-world example	One major creator earned 21× more per view from long-form than from Shorts	

When a user watches Shorts instead of a long video, YouTube earns less, the creator earns less, and the advertiser gets less. The engagement numbers look fine but the economics do not. Shorts are high in volume and low in value.

YouTube's biggest advantage is its long-form library. Fifteen years of tutorials, documentaries, podcasts, and recordings that no competitor can replicate quickly. That library exists because creators could make a living on

YouTube through long-form ad revenue. If the algorithm stops surfacing long-form, views drop, revenue drops, and creators either shift to Shorts or move their best content to Spotify, Patreon, or wherever it pays better. Once they leave, the content goes with them. No algorithm change can surface videos that were never uploaded.

The subscription problem is simpler but just as damaging. When someone subscribes to a channel, they expect to see its videos. That is not happening. Long-form videos from subscribed channels are disappearing from home feeds while the Shorts feed fills up with content from channels the user never chose to follow. People end up searching for specific things rather than browsing. That kills discovery, which is how channels grow and how YouTube keeps users who are not sure what they want to watch. Lose that and YouTube becomes a search engine, not a platform people live in.

1.5 What We Are Solving and What We Are Not

A good problem statement has clear edges. Here is where we are drawing the line.

Out of scope	Why we're not touching it
AI creation tools for Shorts	Supply-side problem. Our job is discovery, not production.
Shorts monetisation policy	Platform economics decision - above product scope.
YouTube's TikTok marketing strategy	Brand and positioning. Different team, different PRD.
Content moderation within Shorts	Trust and safety domain - separate workstream.
YouTube Premium pricing	Commercial strategy. Not our problem to solve here.
Creator analytics dashboards	Creator-side tooling. Valuable, but not this PRD.

In Scope:
How the recommendation surface works between Shorts and long-form, and how to make it smarter for users who want both.

- How the home feed is currently failing users with mixed viewing intent, and what a better version looks like.
- Why users rarely make the jump from a Short into a related long-form video even when they would want to.
- Why subscribed-creator long-form content keeps losing to algorithmic Shorts in the feed.
- What it would take to give users more control over what shows up in their feed.

1.6 The Problem in Plain Language

Here is the one paragraph that should anchor everything else in this document:

YouTube's recommendation system picks Shorts by default when it does not know what a user wants. That default is keeping people stuck in Shorts sessions they did not consciously choose, hiding long-form videos from users who would genuinely watch them, and slowly pulling the platform away from the deep and high-value content that makes YouTube worth opening in the first place. The engagement numbers look fine on the surface. The economics and the creator ecosystem are quietly getting worse underneath.

And here is how we will know if we have actually solved it:

- Users who want long-form content are getting it recommended in their regular feed, not just when they search for it specifically.
- Users who watch Shorts are being shown a clear and visible path to related long-form content, and a meaningful number of them are taking it.
- Content from channels a user has subscribed to is consistently reaching them, whether it is a Short or a long video.
- Long-form creators can see that their content still has a fair shot at being discovered, so they keep investing in making it.

2. Market Context

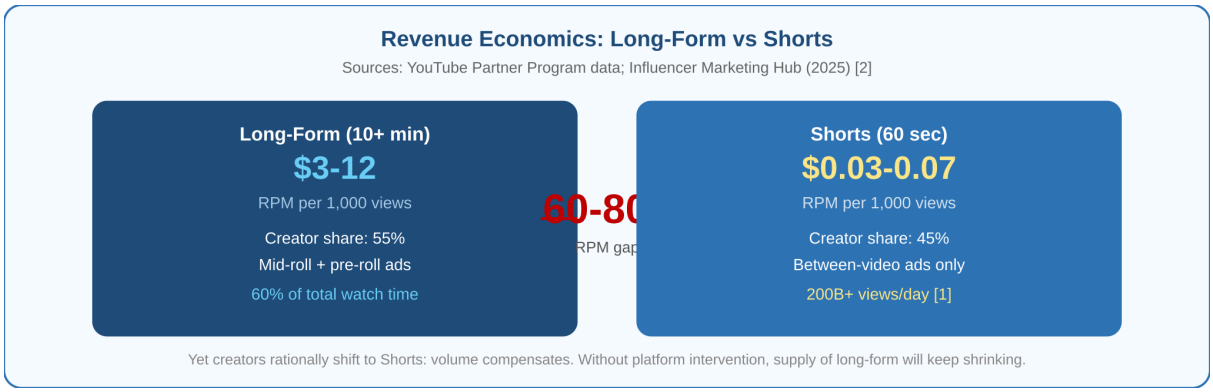
2.1 Market Size

The global video streaming market reached $544B in 2023 and is projected to exceed $1.9T by 2032. YouTube generated $60B in 2024 revenue [13], making it Alphabet's second-largest revenue driver after Google Search, with 2.7B monthly active users [5].

YouTube's dual position as both the dominant short-form and long-form platform is a strategic asset. It is also the source of the tension this PRD addresses.

2.2 Platform Economics: Long-Form Is the Revenue Engine

Sources: [2] Influencer Marketing Hub (2025); [13] Statista / Alphabet (2024);



> Despite Shorts generating 200B+ daily views [1], advertiser RPMs have not followed viewer attention. Long-form RPM is 60-80x higher per impression. The platform's premium inventory which is the asset behind $60B in revenue is long-form. So it's important to focus on long-form content to continue increasing revenue.

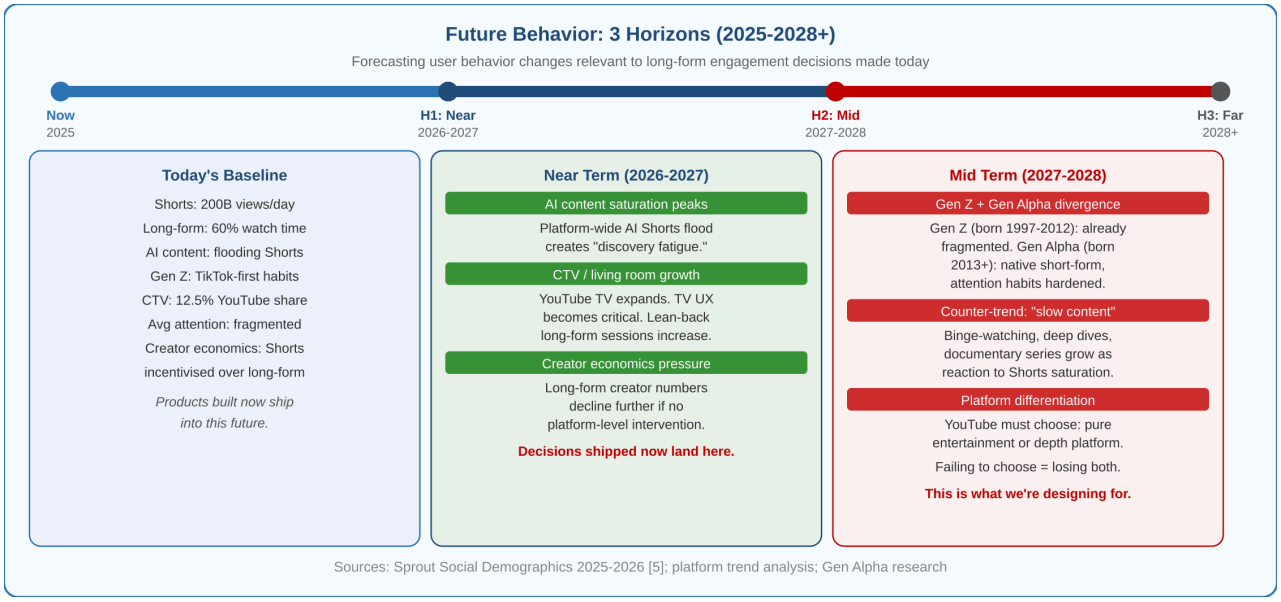
2.3 Trend Lines

Sources: [1] YouTube Official (2025); [7] Subscribr (2025); [12] Internal survey (n=67)

Metric	Shorts	Long-Form
Daily views	**~200 billion/day [1]**	~5B videos/day (~1B hours)
Avg session length	~14 min (12-18 Shorts/session)	~19 min per long-form session [12]
Avg content length	71% of Shorts are 15-30 seconds	Average YouTube video ~12 minutes
Discovery source	**~74% from non-subscribers [7]**	Majority from subscriptions/search
Watch time share	~40% and growing	**~60% of total platform watch time [3]**
Purchase intent	Brand awareness only	**3x higher than social average [5]**

A 2025 homepage redesign quietly made this worse: long-form slots per row were cut from 6 to 2 [6], halving discovery surface area for longer content, a self-inflicted wound on the platform's own revenue engine.

2.3.5 Future Behavior Analysis



Horizon 1 (2026-2027): CTV / Living Room Viewing Becomes the Long-Form Battleground

YouTube holds 12.5% of US CTV viewing time [5] - the largest share of any streaming service. TV sessions are structurally lean-back, longer in duration, and more tolerant of long-form content. But YouTube's TV UX in 2025 is substantially worse than its mobile Shorts UX:

- Remote navigation friction: Slow scroll, no hover preview, ad-trigger issues reported widely [12]
- No context-awareness: TV users get the same homepage as mobile Shorts scrollers
- No lean-back long-form surface: No equivalent of Netflix's 'Continue Watching' or episode queue for creator content

As CTV penetration grows - particularly in India, where CTV adoption is accelerating rapidly - this surface becomes the highest-leverage long-form opportunity on the platform. Users in the living room are already in the lean-forward state that long-form requires. YouTube just needs to meet them there.

Horizon 2: Gen Z vs Gen Alpha Divergence

Generation	Behavior Trajectory	Implication for Long-Form
Gen Z (born 1997-2012)	Attention habits fragmented but not fixed. Gen Z users show the highest 'guilt' after Shorts use [12]. Counter-trend: increasing appetite for long podcasts, documentaries, creator series. 91% are on YouTube [5].	YouTube has a window with Gen Z: they feel the Shorts trap, and some actively resist it. Long-form features that reduce commitment friction could convert this guilt into engagement.
Gen Alpha (born 2013+)	Native short-form. Attention patterns harder and earlier. Screen time begins at younger ages with shorter formats. TikTok habits start in childhood.	Gen Alpha is the most at-risk cohort for permanent long-form disengagement. YouTube has 2-3 years to establish long-form as a habitual format for this generation before habits harden.
Millennial + Older	Highest long-form consumption already. Binge-watching culture well-established. TV/desktop primary devices. 90% are on YouTube [5].	Retain and deepen. This cohort is the stable revenue base for long-form. Better post-watch recommendations and creator series architecture could significantly increase session length.

2.4 Who Loses If This Is Not Solved

Stakeholder	What They Lose	Severity
Creators	Earning potential from long-form; motivation to produce in-depth content	**P0: supply-side collapse if economics don't improve [2]**
Advertisers	Premium brand-safe, high-intent long-form inventory	High: brand advertisers follow watch time, not view counts
YouTube / Alphabet	Revenue moat vs Netflix, Spotify; differentiated value proposition	**High: Shorts alone cannot sustain $60B revenue base [13]**
Users	Access to educational, documentary, and depth-driven content	Medium-term: platform risks becoming a pure entertainment feed

3: Secondary Research

3.1 YouTube Platform Analysis

Current Discovery Surfaces

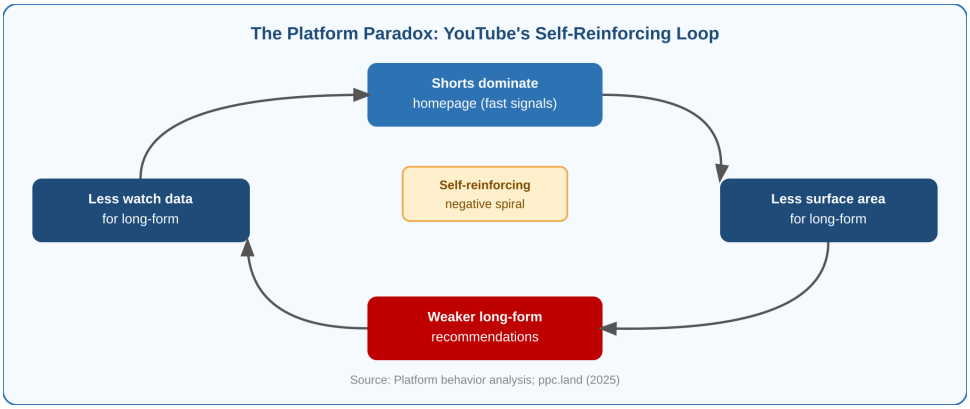
YouTube's key discovery surfaces in 2025: Homepage (algorithmic feed), Search, Shorts shelf, Subscriptions feed, Trending, and YouTube TV/CTV. The Shorts shelf appears before the long-form feed for most mobile users. **The 2025 redesign reduced long-form homepage slots from 6 to 2 per row** [6].

The Algorithm: Collaborative vs Content-Based Filtering

YouTube's recommendations rely primarily on collaborative filtering, 'users like you watched X.' This works well for Shorts: billions of daily views generate fast, high-confidence signals. Long-form is structurally disadvantaged: a 40-minute video accumulates signals slowly, completion data is sparse, and quality is harder to extract algorithmically.

Content-based filtering understanding what a video is about, who it's for, how credible the creator is is more critical for long-form discovery but harder to implement at scale. This is not a bug. It is an emergent consequence of optimising for engagement velocity over engagement depth.

The Platform Paradox: A Self-Reinforcing Loop



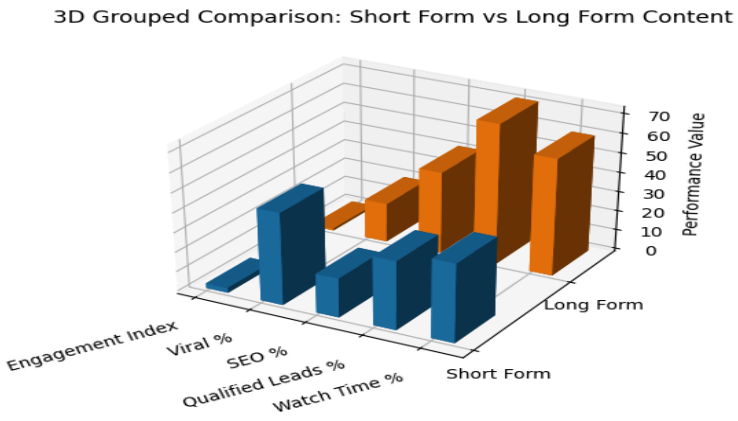
> This is not intentional product strategy : it is an emergent consequence of the different optimisation target. Engagement rate (clicks, short-term watch time) is being maximised at the expense of engagement depth (completion, return visits, subscriber loyalty).

Existing Long-Form Features

YouTube has built tools for long-form: Chapters, Key Moments, Clip sharing, Playlists, AI-generated summaries. However, all require creator opt-in and none are surfaced prominently enough to change viewer behaviour at scale. A user encountering a 25-minute video has no lightweight quality signal before committing.

3.2 Format Analysis: The Intent Divide

Sources: [3] Amra & Elma (2025); [9] Still Waters Digital (2025); [12] Internal survey (n=67); [16] Consolidated Research



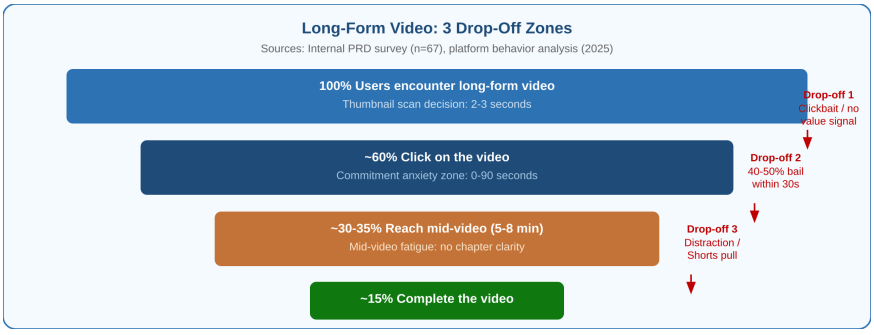
Key distinction

Short-form and long-form are not the same content at different lengths. They represent different user intents, psychological contracts, and viewing modes. Solving a lean-back experience with lean-forward tools or vice versa will fail.

Dimension	Short-Form	Long-Form
User posture	**Lean-back - 'kill time.'**	**Lean-forward - 'invest time.'**
Commitment	Zero: feed chooses for you	High: deliberate click, time risk
Completion rate	**20% higher for under-30s [3]**	**40-50% bail within 30 seconds [12]**
Watch time share	~40% and growing [7]	**~60% of total watch time [3]**
Post-watch emotion	Stimulated, slightly empty [12]	Informed, immersed, satisfied [12]

The Consumption Gap: 3 Drop-Off Zones

Sources: [12] Internal survey (n=67); [6] PPC Land (2025)



> The pattern across all three zones: perceived value must grow faster than perceived time cost. YouTube provides no mechanism to accomplish this. Shorts require zero commitment. Long-form requires faith. The platform has built no bridge between them.

3.3 AI: Cause and Potential Solution

How AI Accelerated the Problem

AI video creation tools have cut Shorts production costs dramatically. Creators can produce 50+ Shorts per day with minimal effort, flooding the platform with high-volume, low-quality content [7]. This has two compounding effects:

- **Volume effect**: More Shorts = more collaborative filtering data = stronger algorithmic confidence = more Shorts recommended
- **Signal dilution**: As Shorts dominate engagement signals, long-form receives proportionally fewer impressions regardless of quality

> AI-generated volume is not just adding noise it is reshaping the algorithm's preference signals in a way that structurally disadvantages crafted long-form content. This is the creator economics problem viewed from the algorithm's perspective.

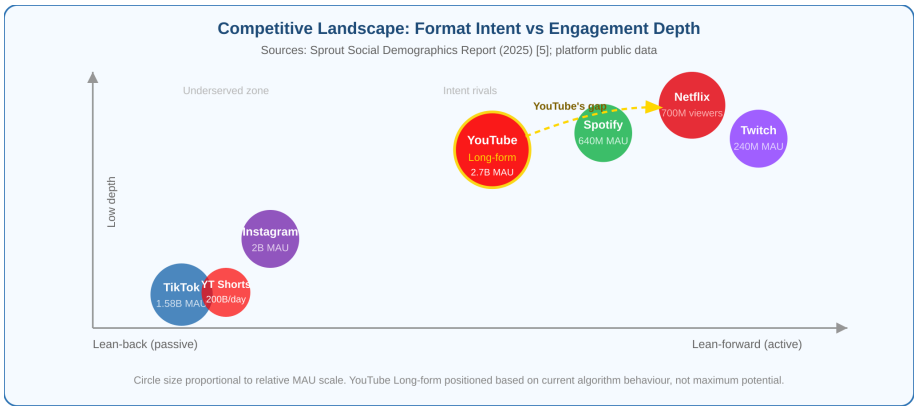
AI's Dual Role

The same technology that enables high-volume Shorts can solve long-form discovery if applied differently:

AI Application	How It Addresses the Long-Form Problem
Intent detection	Identify 'learn' vs 'entertain' sessions using search history, device, time of day, and session context
Quality scoring	Analyse video structure, pacing, chapter density, creator credibility - surface signals before user commits
Adaptive previews	Generate personalised 20-second previews tailored to what a specific user is most likely to find compelling
Cross-format bridging	Detect when a Shorts viewer's engagement pattern signals latent demand for the long-form version

3.4 Competitive Benchmarking

Sources: [5] Sprout Social (2025); [8] Netflix (2023); [11] Demand Sage (2025); [14] Netflix IR (2024); [15] Spotify IR (2025)



Direct Format Rivals

Platform	Scale	What They're Doing	Key Insight for YouTube

TikTok	1.58B MAU; 3.70% eng. rate [5]	Testing 10-30 min videos. Early data shows high abandonment even its dopamine-loop UX cannot easily convert passive scrollers to long-form	Even the best lean-back platform struggles with the lean-forward transition. The intent divide is structural, not YouTube-specific.
Instagram	2B MAU; Reels = 50% of all app time [5]	Deliberately balances Reels vs. feed. Made algorithm adjustments to prevent Reels cannibalising feed. IGTV was deprecated and long-form failed entirely.	Format coexistence requires intentional algorithm design. IGTV is a direct cautionary case: forcing long-form on a short-form audience without changing the product contract doesn't work.

Intent Rivals: Platforms That Solved Long-Form Retention

Platform	Scale	How They Keep Users in Long-Form	The Gap YouTube Must Close
Netflix	~700M viewers; 100B+ hrs watched [8]	Autoplay, episode continuity, '% viewers finished this,' series hooks, personalised trailers. Hooks users into series commitment, not just single-content quality.	Netflix's advantages are commitment architecture, series structure, cliffhangers, social proof, autoplay. A 20-min YouTube video competes against a 3-sec Short with none of this.
Spotify	~640M MAU; avg session extended from 3 min to 2+ hrs [15]	Transitioned users from 3-min songs to 2-hr podcasts through playlist behaviour, seamless format switching, passive consumption design.	Spotify normalised long-form without asking users to change habits. YouTube makes users make a conscious commitment that Shorts never require.
Prime Video	315M MAV; X-Ray Recap feature [5]	X-Ray Recaps let users preview key scenes and plot context before committing to a full watch. Reduces the time-risk of starting a long piece of content.	Prime's X-Ray is the closest analogue to what YouTube needs: a low-commitment 'try before you buy' mechanism that surfaces quality before the user commits.
Twitch	~240M MAU; 35M DAU; avg 3.1 hrs/session [11]	VOD + live combo retains community after live ends. Users stay for the creator relationship, not just the content. Deep interactive engagement via live chat.	YouTube has creator loyalty potential through Subscriptions but doesn't activate it subscribed channels get buried by an algorithm that prioritises Shorts over subscriber content.

> The Netflix Question: Netflix keeps users for 60+ minutes per session through commitment architecture (series, continuity) not content quality alone. YouTube has none of this for creator content. That is the gap.

3.5 Key Insights from Secondary Research

These are conclusions, not data points. Each insight maps directly to a pain point and informs a solution in Section 9.

#	Insight (What We Found)	Implication for YouTube
1	**Algorithm's collaborative filtering systematically disadvantages new long-form content discovery that is broken by design, not neglect.**	Fix requires explicit content-based filtering for long-form, not just tuning existing weights.
2	**Short-form and long-form represent different intents (lean-back vs. lean-forward). Current product surface is optimised for lean-back, crowding out lean-forward intent.**	Long-form needs a different product surface, not just different recommendation weights in the same feed.
3	**40-50% of users abandon long-form within 30 seconds. The problem is commitment anxiety and absence of quality signals, not content quality.**	try-before-commit mechanic (AI previews, chapter carousels, hover thumbnails) could recover the majority of abandoned sessions.
4	**Long-form RPM is 60-80x higher per impression than Shorts, yet creators rationally shift to Shorts because volume compensates.**	Creator incentive misalignment will not self-correct. Platform-level revenue or bonus programmes are required.
5	**Shorts reach 74% non-subscribers, a massive discovery pipeline. Yet 58% of Shorts viewers rarely or never convert to the source long-form video.**	The Shorts-to-long-form pipeline is an underused funnel. Higher ROI than acquiring new long-form viewers from scratch.
6	**AI-generated Shorts flood the platform with high-volume, low-quality content, diluting algorithmic signals and suppressing quality-based long-form recommendations.**	Quality scoring (completion rate, rewatch rate, expert signals) must outweigh raw engagement velocity in long-form recommendation logic.

4. Primary Research

4.1 Research Methodology

Method	Details	Sample
Quantitative Survey	16-question Google Form via WhatsApp	n=71 valid
User Interviews	Semi-structured calls/video, 20-30 min	n=3 (ages 18-24, 35-44, 25-34)

4.2 Survey Findings (n=71)

Respondent profile: 18-24 (18%), 25-34 (39%), 35-44 (28%), 45+ (14%) - spans all target cohorts aligned with YouTube India's core demographics.

Insight 1: Time is the number 1 barrier, but it's a perception problem

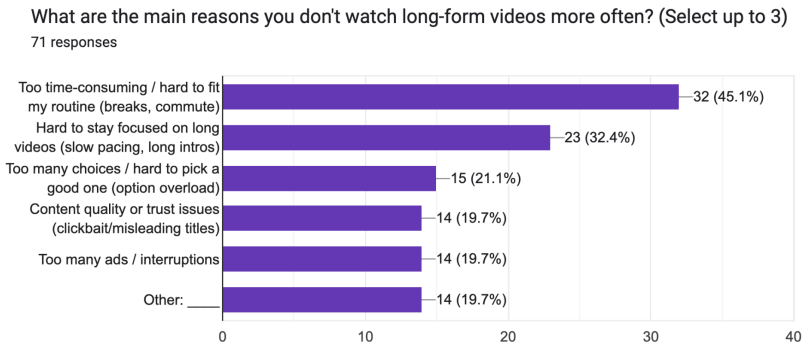
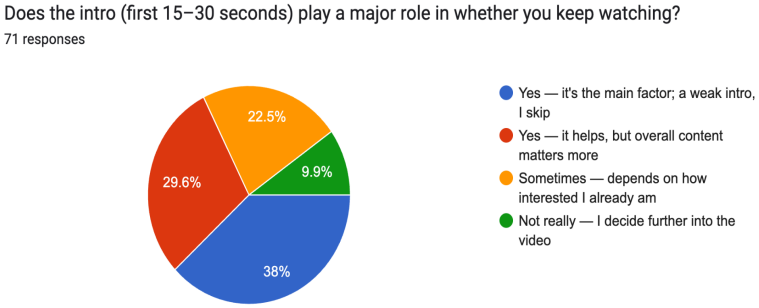


Fig 1: Barriers to long-form viewing (multi-select, n=71)

45% cite time constraints as the top barrier. Yet average users spend 95+ min/day on Shorts, equivalent to 5-6 long-form videos. Users have time; they lack confidence that a specific video will justify the investment.

Insight 2: The first 15-30 seconds make or break retention



34% say the intro is THE deciding factor for whether they keep watching; another 31% say it helps significantly. Combined, 65% of respondents make stay or,leave decisions within the opening seconds of a long-form video.

Insight 3: Users want to try before they commit

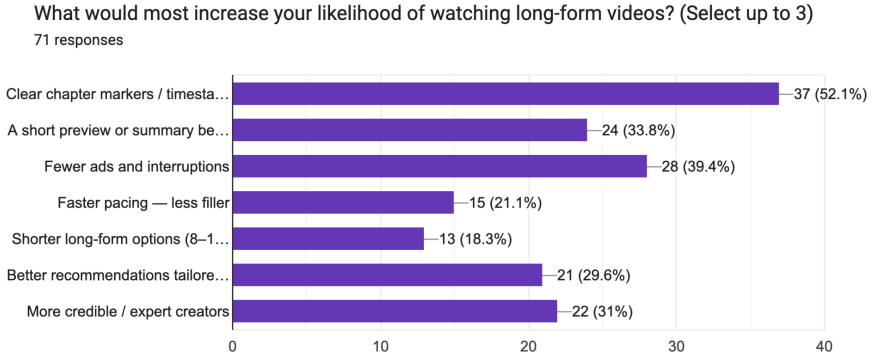
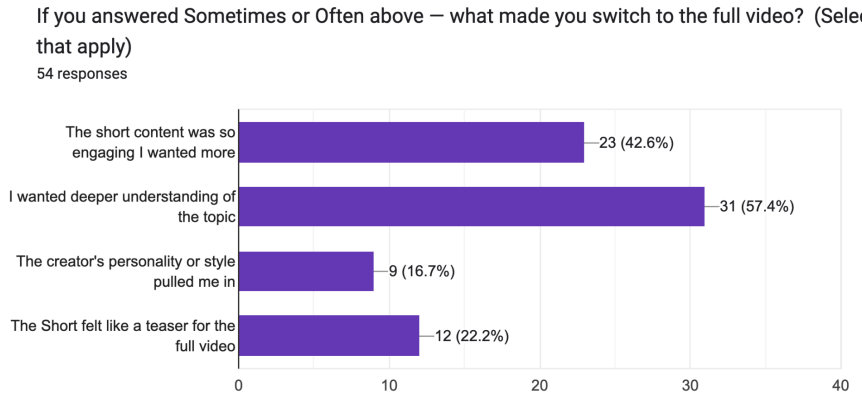


Fig 2: What would most increase long,form viewing (multi,select up to 3, n=71)

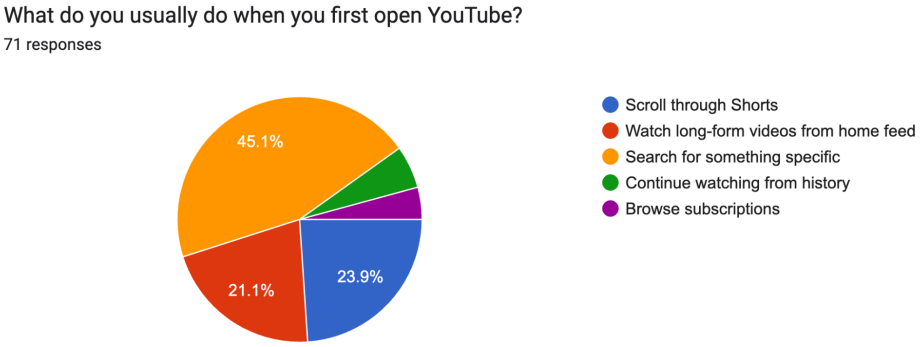
Chapter markers (52%), fewer ads (38%), and preview/summary before committing (27%) are the top feature requests. Users are explicitly asking for tools that de,risk the time investment.

Insight 4: Short→Long conversion has demand but the pipeline leaks



52% have gone from a Short to the full video at least sometimes. The number 1 reason: wanting deeper understanding (48% of converters). Yet 48% rarely or never convert. Latent demand exists but the current UX fails to surface it effectively.

Insight 5: Shorts is often an unintentional entry point



35% search for something specific as their first action, but 24% start by scrolling Shorts - the second most common entry behavior. Interviews suggest many opened the app with a different intent but were pulled into the Shorts feed.

4.3 User Interviews (n=3)

Semi-structured interviews across 3 distinct personas.

Interview 1: College Student, 18-24, Phone call, ~25 min

A heavy **YouTube Shorts** user who primarily turns to **long-form** for **education** and **music**. They described feeling **guilty** about time spent on Shorts and “trapped” in low-intent scrolling loops. When a Short triggers interest in a longer video, they noted there’s **no seamless one-tap path** to continue today; they typically have to **open the creator’s channel and manually search**, which adds friction and often breaks momentum. For choosing long-form, **thumbnails** and **like counts** act as meaningful quality cues, while generic **verification badges** don’t increase trust or likelihood to click.

> *“Adding some kind of tick doesn’t make sense. Quality signals matter more than a bare checkmark.”* - Student, 18-24

Interview 2: Homemaker, 35-44, Phone call, ~25 min

Uses YouTube shorts throughout the day, listens to **morning bhajans while cooking**, **Shorts in the afternoon**, and **kids' educational videos** later. Because the phone is shared with children, she raised concerns about **privacy** and a "corrupted" home feed (recommendations shifting away from her preferences). Retention is decided quickly: the **first 3-4 minutes** determine whether she continues a long-form video. When a Short sparks interest in the full version, she experiences the same friction as other users who have no **direct link**, so she must **go to the channel and search**, often losing momentum. She also finds the **home feed overloaded** and frequently **skips it entirely**, relying on **notifications** as her primary way to find what to watch.

> *"When I want the full video, there's no direct link from the Short - I have to visit the channel and search."*
> - Homemaker, 35- 44

Interview 3: IT Professional, 25-34, Google Meet, 22 min (recorded)

Uses YouTube **3-4 times per day,** primarily for **IT upskilling** and prefers **~40-minute** videos for depth. They view **Shorts as too superficial** for serious learning and **don't trust the home feed** for educational discovery; instead, they rely on **word-of-mouth recommendations** and then **search for creators directly**. Their top product is a **LinkedIn-style verified creator badge** that signals **accountability** (not just popularity). They also find **hover previews** unhelpful, "like a trailer that doesn't reveal the plot." On AI content, they're open to **AI-assisted** videos when **sources and attribution are transparent**, but they **distrust fully AI-generated** videos.

> *"I distrust 100% AI-generated content. Source transparency is essential for credibility."*
> - IT Professional (Deepak), 25-34

Patterns Across Interviews

Pattern	Student (18-24)	Homemaker (35-44)	IT Professional (25-34)
Short→Long bridge broken	No direct link; must search channel	Same, visits channel manually	Hover preview unhelpful; needs deeper signals
Generic badges don't build trust	"Tick doesn't make sense"	Judges by content, not badge	Wants LinkedIn-style verified badge
Feed doesn't match intent	Wants priority subscriptions	Bypasses feed via notifications	Relies on external recs, not feed

Output: 20+ raw pain points consolidated into 14 viewer, 4 creator, and 4 platform-level items, further prioritized through team voting into the final set in Section 4.

Synthesis

Three insights consistently surfaced across both research methods:

1. Users have the time for long-form but lack the confidence to invest it. The barrier is not availability - it's the absence of low-effort ways to assess whether a video is worth watching before committing (previews, chapters, trust signals).
2. The Short→Long conversion pipeline has real demand (52% have converted) but leaks at the UX level. Two interviewees independently described identical friction: no direct path from an engaging Short to its source video.

3. The first 30 seconds and content-level trust signals are the two gates where long-form loses viewers before quality can speak for itself. 65% say intros determine retention; all three interviewees rejected generic verification in favor of quality-based credibility.

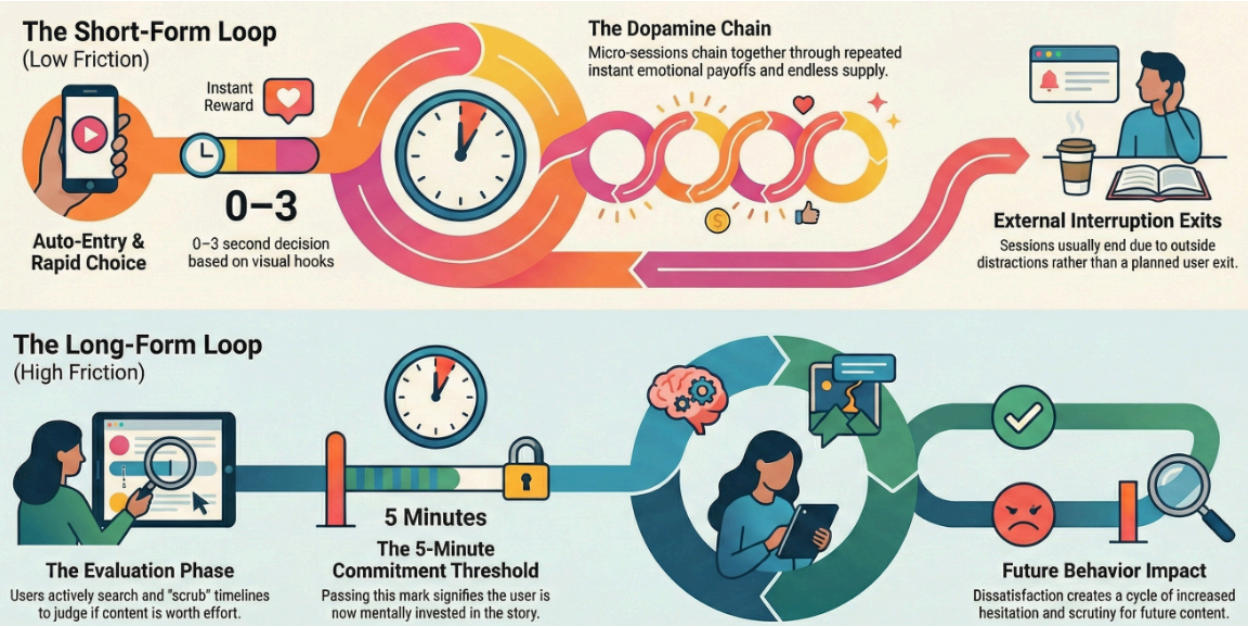
5. Behavioral Analysis

Focus: Drop-off points, content preferences, and engagement patterns for YouTube (Shorts + Long-Form) benchmarked against Netflix, Prime Video, TikTok, and Instagram Reels - analyzed through three user personas

5.1 Personas and their Behaviour:

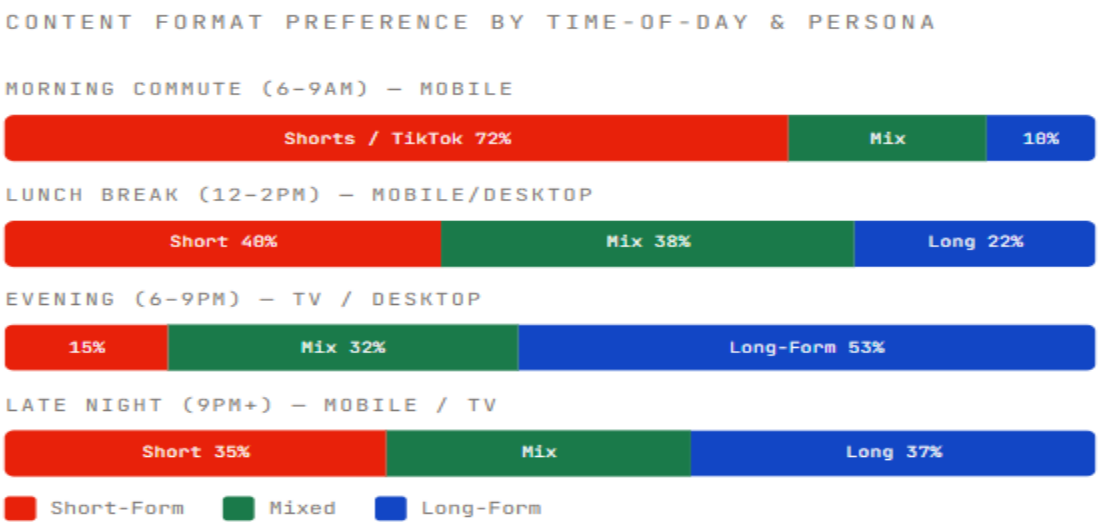
Grazers vs Intent Seekers
Grazers: Consume without agenda. Platform decides what's next. Passive mode. High swipe rate.
Primary persona: **Short Heavy** (Platform fit TikTok → YT Shorts → Reels)
Intent Seekers: Arrive with topic or creator in mind. Search-led. Deep engagement once committed.
Primary persona **Long Heavy** (Platform fit YT Long-Form → Netflix)
Context Switchers(Mixed): Switch between grazing and intent based on time-of-day, device, and emotional state. Most valuable audience - highest per-session monetizable potential if bridge mechanics exist.

5.1.1 Shorts Vs Long form User Journey & Behavior:



Shorts is a zero-friction swipe loop; long-form is a high-stakes commitment - and the gap between those two experiences is where most engagement dies.The long-form experience is dominated by intent-driven with friction at every stage. Entry requires active judgment, commitment.

5.3 Context & Situational Behaviour: Engagement behavior is deeply context-dependent. The same user may be a Short-Form Heavy persona at 8am on a commute and a Long-Form Heavy persona at 9pm on a TV screen.Device and time-of-day are more predictive of behavior than demographic alone.



5.4 Situational Triggers & Mode-Switching

Situational Trigger Map: Suits for all personas (Short/Mix/Long)
Mobile + Commute: Short-form default for all personas.
TV + Evening: Long-form default for Long Heavy and Mixed.
Desktop + Focused: Search-led long-form (tutorials, explainers).
Mobile + Pre-sleep: Short-form for Short Heavy; streaming for Long Heavy. Context predicts format choice more than stated preference.

Mode-Switching Barrier: <5%: Less than 5% of Shorts sessions lead to a long-form video watch in the same session.
The dopamine-primed state from short-form →makes the cognitive gear-shift to long-form
Mixed persona has highest switching potential (22% in our model). This <5% represents YouTube's single biggest untapped engagement lever

5.5 Core Findings: Three behavioral gap that exist in youtube's engagement. Lets understand how Competitors are solving this.

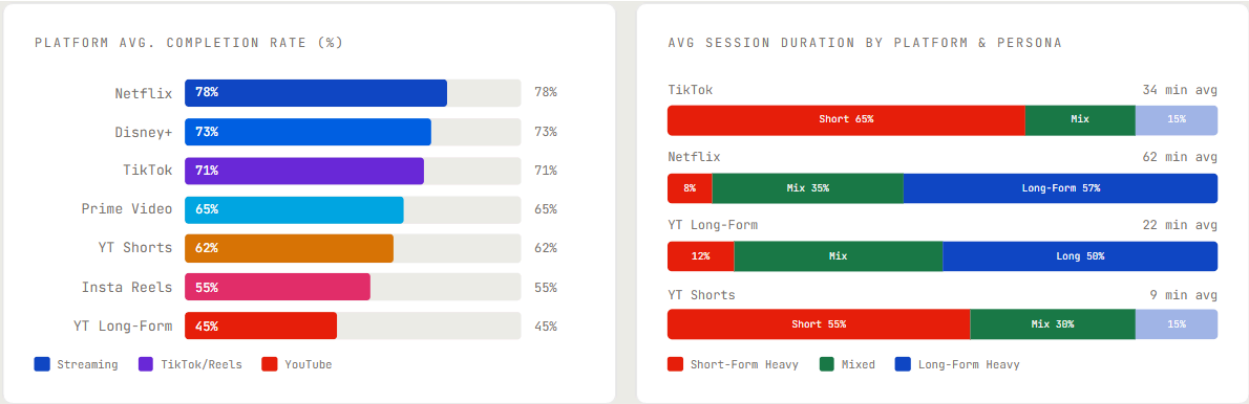
Gap 1:Sampling (Try before commit) Failure
See long video → Click → Unsure if worth time → Scrub / skip to check value → Value not clear fast → Exit early → Try another video → Repeat
No platform has solved pre-commitment uncertainty as well as Netflix's trailer mechanic.

Gap 2:Commitment Deficit
Watch short video → Instant reward → Brain expects fast entertainment → Long video feels like effort → Avoid long video → Watch more shorts → Habit strengthens → Long videos feel even harder → Repeat
hort-form habit rewires reward expectation - 15-second dopamine loops make 20-minute video commitments feel psychologically expensive.

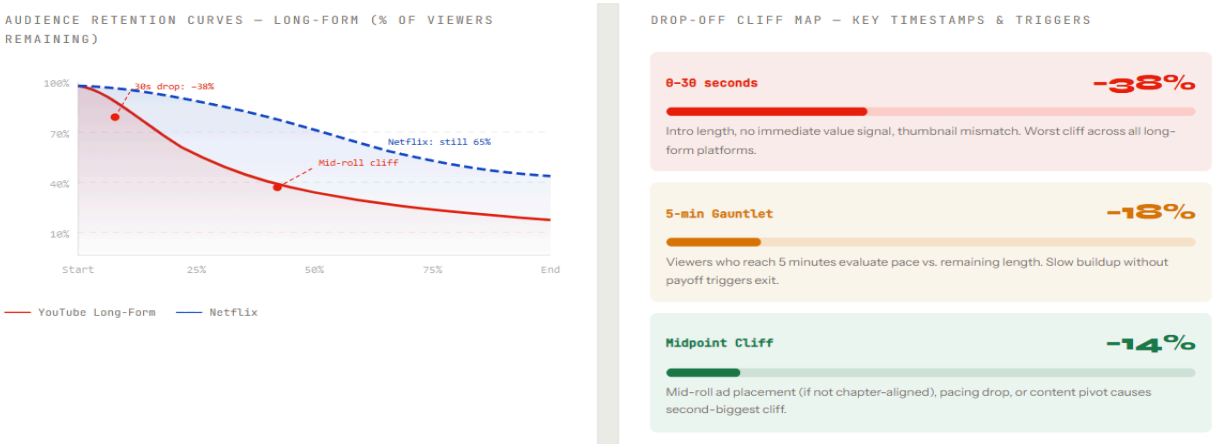
Gap 3: Habit Lock-in Imbalance
Open app → Algorithm shows engaging content → Scroll continuously → Variable rewards → Strong habit → Open again automatically → Repeat
YouTube's long-form viewers rely on intent (search), while short-form users rely on algorithm - creating two disconnected engagement loops.

5.6 Metrics, Drop-Off Curves

Benchmarking against competitor equivalents: Watch Time & view Distribution



Drop-Off Curves - Timing & Cliff Analysis:



5.7 Behavior Patterns

Streaming platforms (Netflix, Prime Video): Streaming platforms achieve 65-78% completion rates vs. YouTube's 45% by solving three problems YouTube hasn't: pre-commitment uncertainty, mid-content friction, and session extension loops. Long-Form Heavy persona is most susceptible to streaming migration.

Short-Form Competitors (TikTok & Instagram Reels Ecosystem): TikTok's FYP algorithm creates the most potent engagement loop in media history, capturing 40% of mobile short-form viewing time. Short-Form Heavy personas show 60-70% TikTok session overlap with YouTube Shorts usage.

Dopamine Burst Architecture

Micro-rewards every 15-60s create a slot-machine reinforcement schedule. Satisfaction comes from the act of consuming, not the content itself.

Zero Friction Model
No title evaluation, no length commitment, no scrubbing. The content presents itself - cognitive cost approaches zero.

Viral Chaining
Sound trends, duets, stitches, and remix culture create content chains - a single viral sound surfaces 10,000+ related clips, extending sessions organically.

5.8 Cross-Platform Behavioral Matrix
Mapping entry → loop → exit across all platforms reveals structural advantages competitors hold. The "Engagement Depth Index" scores each platform on shallow (passive) vs. deep (intentional) engagement modes.

Journey Comparison - Entry → Loop → Exit

PLATFORM	ENTRY TRIGGER	ENGAGEMENT LOOP	EXIT PATTERN	RECOVERY RATE	DEPTH MODE
YT Long-Form	Search or algo recommendation	Sequential video, playlist (rare), subscription feed	Expectation fail, ad friction, algorithm rabbit hole	~35% return same session	Deep
YT Shorts	Homepage shelf or direct	Infinite vertical scroll, algorithm-fed	External interrupt, format fatigue, deliberate pivot	~65% open again <1hr	Shallow
TikTok	FYP (no intent needed)	Variable reward FYP loop, sound trends, duets	External only — no natural endpoint	~78% return <30min	Shallow
Netflix	Hover trailer reduces pre-commit	Episode → auto-next → binge architecture	Sleep timer, episode end, "Are you still watching?"	~62% continue same day	Deep

5.9 Situational Triggers & Mode-Switching

Mobile + Commute: Short-form default for all personas. **TV + Evening:** Long-form default for Long Heavy and Mixed. **Desktop + Focused:** Search-led long-form (tutorials, explainers).

Mobile + Pre-sleep: Short-form for Short Heavy; streaming for Long Heavy. Context predicts format choice more than stated preference.

Mode-Switching Barrier: <5%: Less than 5% of Shorts sessions lead to a long-form video watch in the same session. The dopamine-primed state from short-form makes the cognitive gear-shift to long-form neurologically difficult. Mixed persona has highest switching potential (22% in our model). This <5% represents YouTube's single biggest untapped engagement lever

Behavioral Gaps
Try before commit (Critical): 40-50% exit in 30s. Cause: no pre-commitment mechanism, long intros, thumbnail deception. Impact: £xM in potential watch-time revenue lost per day.

Format Commitment Deficit (Critical): <5% Shorts → Long crossover. Short-form habit degrades long-form attention threshold progressively. Short Heavy persona almost never crosses over without a deliberate mechanic.

Mid-Content Ad Friction(High): Mid-roll ads placed without chapter-awareness cause 25-40% mid-video abandonment. Netflix's ad-free model is the behavioral benchmark viewers are being trained toward.

Habit Lock-in Imbalance(High): TikTok's variable-reward FYP creates stronger behavioral compulsion than YouTube's subscription model. Predictable creators = weaker excitement loop

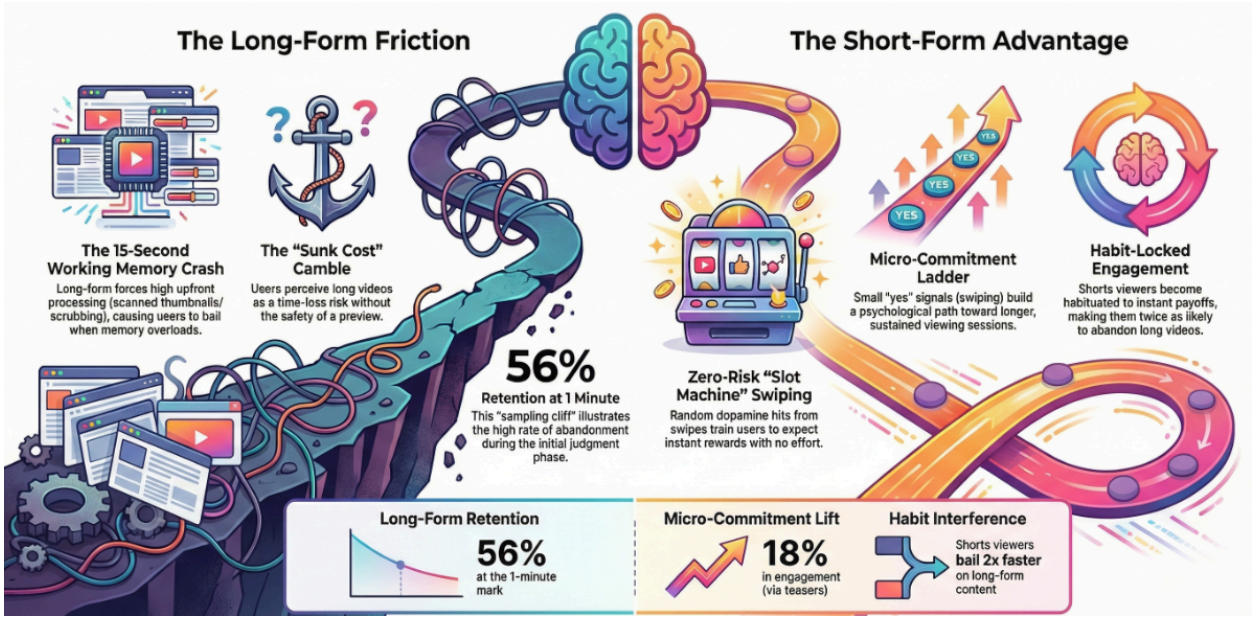
Subscriber Relationship Weakness (Medium): Subscriptions should function like Disney franchise trust - guaranteed click. Instead, notification failure means <20% of subscribers see new videos within 24 hours.

No Binge Architecture(Medium): YouTube has no equivalent of Netflix's auto-next episode mechanic for related long-form content. Playlist usage is <8% of sessions. Session extension opportunity is structural.

5.10 Key Psychological theories that explain the Behavioral Gaps

Users face a cognitive threshold to commit to long-form (40-50% bail at 30s vs Shorts' 62% completion)-rooted in friction, risk, and mismatched habits. key theories explain this

Theory	Why It Creates Friction	YouTube Impact	Quant Gap
Cognitive Load Theory	Long: thumbnail→scrub→judge (high processing). Shorts: swipe (zero load). Users bail at **working memory overload** (first 15s).	56% retention at 1min explains sampling cliff.	**40-50% 30s drop**
Loss Aversion (Prospect Theory)	"10min video?" = **perceived time loss risk**. Shorts = zero-risk; long = sunk cost gamble post-click.	Needs **micro-proof** (hook/preview) before commit.	**Runtime aversion**



6. User Persona

Attribute	Shorts Heavy	Long-form Heavy	Mix
Entry behavior	Opens app → scrolls Shorts → checks Notifications → Subscriptions feed → taps video	Searches directly or follows Notifications → Subscriptions feed → opens video	Opens app → browses feed → taps Notification → video; mixes both entry paths
Primary goal	Entertainment, bite-sized news, mood uplift	Skill building, devotional, podcasts, entertainment, music, long-form news	Skill building, entertainment, music, podcasts, news, devotional - all of it
Typical session	Multiple short bursts throughout the day; rarely one long sitting	One or two intentional deep-dive sessions; search is the primary entry	Varies by day and mood; can do both a Shorts binge and a 30-min watch in the same day
Key pain points	No clear bridge to related long-form; feed overrun by unsubscribed creators; long videos feel like a risky time commitment	First 3-4 min decides watch-through; Home feed is noisy with irrelevant organic content that breaks the intent-driven session	Long videos vanish from the homepage even when actively watching; algorithm forces Shorts; Shorts-to-long nudge is weak

Success metrics (examples)	Resume rate, completion within 24-72 hrs, fewer abandonments	Completion rate, session length, repeat topic engagement	Continuous watch time, lower early exits, higher satisfaction prompts

6.1 Persona - Shorts Heavy

Just one more, okay maybe ten more

The phone comes out and Shorts are already playing. Vertical, instant, no commitment. There is no decision to watch. It just happens. Shorts fill every gap in the day. The session ends not by choice but by distraction or a dead battery. Then comes the familiar guilt that the last 45 minutes could have been spent differently.

"I keep seeing Shorts from channels I have never heard of. The people I actually subscribed to? I have to go hunt for them."

6.2 Persona - Long-form Heavy

If the first four minutes do not earn it, I am out.

This user comes with a plan. Search is the entry point. They know what they want and go find it. Quality is judged in the first 30 seconds. By minute three the decision is made. They have no guilt about leaving. They have a queue waiting. The home feed is noise, not opportunity.

"The first few minutes are everything. If it does not earn my attention, I close it and move on."

6.3 Persona 3 - Mix (Shorts + Long-form)

"I want both, but the app keeps choosing for me."

A Day in the Life:

In rushed moments, this user scrolls through Shorts. When time opens up, they shift into long-form: a documentary, a podcast, a cooking video. Intent changes with context, not preference. Both formats have a real place in their day. The problem is that the app does not see this. It sees the Shorts session and doubles down, pushing more Shorts and burying the long-form content this user also genuinely wants.

"I watched 40 minutes of Shorts and could not find the documentary I had saved. The app just pushed me toward Shorts and did not let go."

---

7. Challenges in AI Implementation:-

Technical Challenges:-

Challenge	Description	Impact
Data Bias & Feedback Loops	AI learns from past behavior. If users mostly watch Shorts, the system keeps recommending	Hard to rebalance the ecosystem. Reinforces short attention behavior

	Shorts and suppresses long-form content.	
Cold Start Problem	New creators or new long videos lack watch history and engagement signals, making it difficult for AI to rank them accurately.	New long videos get low visibility. Slower content discovery.
Real-Time Personalization Complexity	AI must process watch time, drop-off points, device type, time of day, and session intent to optimize recommendations. Scaling this for billions of users is computationally expensive.	High infrastructure cost. Latency issues.
Ad Revenue vs User Experience Conflict	The recommendation system must balance showing more ads without hurting retention. Over-optimizing ad frequency may reduce watch time.	Revenue volatility. Creator dissatisfaction.

User Experience Challenges:-

Challenge	Description	Impact
Perceived Manipulation	If recommendations feel too aggressive, users may feel the algorithm is pushing content or controlling what they watch.	Trust issues. Reduced platform loyalty.
Over-Personalization Bubble	AI repeatedly shows similar content, limiting variety.	Echo chambers. Content fatigue. User boredom.
Attention Overload	Too many recommendations through autoplay, next suggestions, Shorts feed, and notifications overwhelm users.	Mental fatigue. Shorter sessions.

8. Pain Points

8.1 Viewer Side

Rank	Pain Point	Category	Severity	Why Prioritized

1	Users don’t have a low-effort, “try-before-you-commit” way to sample long-form videos as they do with Shorts, which makes it harder to discover, engage with, and watch more relevant long-form content	Viewer	P0	Highest-cited barrier; explains why 73% watch-time share is eroding to Shorts
2	Videos often feel too long compared with the limited time available, which affects about 45% of users	Viewer	P0	25% of respondents' first action is "scroll Shorts," many with a different original intent.
3	Shorts as a compulsive feature instead of an option: cannot switch off shorts, leading to distraction when launching the app or website	Viewer	P1	69% of survey respondents make stay-or-leave decisions in the first 30 seconds; no platform-level Skip Intro exists, unlike Netflix (136M daily uses)
4	Weak intros kill retention within 15-30 seconds because creators fail the hook, and the platform offers no Skip Intro mechanism	Viewer	P2	17% cite clickbait/trust issues as a direct barrier. "Trusted creator" is the #2 click trigger — users rely on familiarity because the platform provides no quality indicators (completion rates, chaptered badges, expert/educator tags, topic-specific credibility).
5	No content-level trust signals for long-form videos — users face clickbait thumbnails, AI-generated content, misleading titles, and no way to assess quality before committing, forcing them to default to a small set of known creators.	Viewer	P2	Only appears when creators manually link it, and operates on a separate recommendation system. 58% of respondents (n=64) rarely or never convert, yet the #1 reason among those who do is "deeper understanding" — latent demand exists, but the current UX fails to surface it at the right moment.
6	Shorts-to-long-form conversion is weak despite an existing CTA — the link to the source video suffers from poor discoverability (small, non-persistent, buried in UI)	Viewer	P2	Only appears when creators manually link it, and operates on a separate recommendation system. 58% of respondents (n=64) rarely or never convert, yet the #1 reason among those who do is "deeper understanding" — latent demand exists, but the current UX fails to surface it at the right moment.
7	Many educational video “courses” are incomplete, pushing users to rely on dedicated learning platforms instead.	Viewer	P3	71% of Indian YouTube users use the platform for skill acquisition, but unstructured content can't compete with course-format platforms.

8.2 Creator Side

#	Pain Point	Severity	Description	Evidence	Implication
1	Revenue Imbalance Makes Long-Form Economically Risky	P0	Long-form RPMs 60-80x higher per impression but Shorts volume compensates. Twitch: 50-70% sub rev share; Kick: 95/5. Creator spends more time — what if they won't get views? Feels demotivated.'	Long-form RPMs 60-80x higher per impression but Shorts volume compensates. Twitch: 50-70% sub rev share; Kick: 95/5. Creator spends more time — what if they won't get views? Feels demotivated.'	Higher long-form rev share, bonus programs, or minimum RPM guarantees needed to shift creator incentives toward quality depth.

2	New Creators Can Only Get Discovered Through Shorts	P1	Discovery algorithm overwhelmingly favors Shorts for new creators. Long-form from unknown creators receives minimal impressions. Forces 'Shorts first, long-form later' strategy — shaping creator habits away from the format YouTube needs to protect.	Reaching is difficult for new creators — relying on Shorts to go viral.' Shorts = hundreds of billions monthly views vs long-form's asymmetric discovery.	Dedicated long-form discovery pathway — 'Rising Long-Form' section, algorithm boost for first-time long-form publishers.
3	AI-Generated Content Noise Drowns Quality Creators	P3	AI tools enable 50+ low-quality Shorts/day, flooding recommendations. Algorithm optimizes for engagement velocity (clicks, watch-through), favoring mass-produced content over crafted long-form depth.	More creators using AI → high-volume, low-quality Shorts becoming noise → low viewership of deep, high-quality content.'	Content quality scoring (completion rates, rewatch rates, satisfaction) should outweigh raw engagement metrics in long-form recommendations.

9. Solution Space

9.1 The Problem We Are Solving

Imagine you are scrolling through YouTube. A short clip grabs your attention instantly. You watch it. You move on. Now imagine a 45-minute documentary thumbnail sitting right next to it. You have no idea if it is worth your time. You might click. You might not. Most people do not.

This is the core tension YouTube has not yet resolved. Shorts are frictionless to sample. Long-form videos are a blind commitment. Users are not lazy. They are just rational. And right now, the rational choice is to skip.

Netflix already solved this for movies. Prime Video built X-Ray Recaps for exactly this reason. YouTube, despite sitting on the world's largest video catalog, has neither. That gap is what we are designing for.

Pain Point: Users lack a low-effort, try-before-you-commit way to sample long-form content the same way they can with Shorts.

WHY NOW	YouTube already captures per-second engagement data for every video the audience

retention graph visible in YouTube Studio. This data has never been productised as a viewer-facing discovery feature. The infrastructure exists; the feature does not.

9.2 Solution Overview - AI-Powered Preview Engine

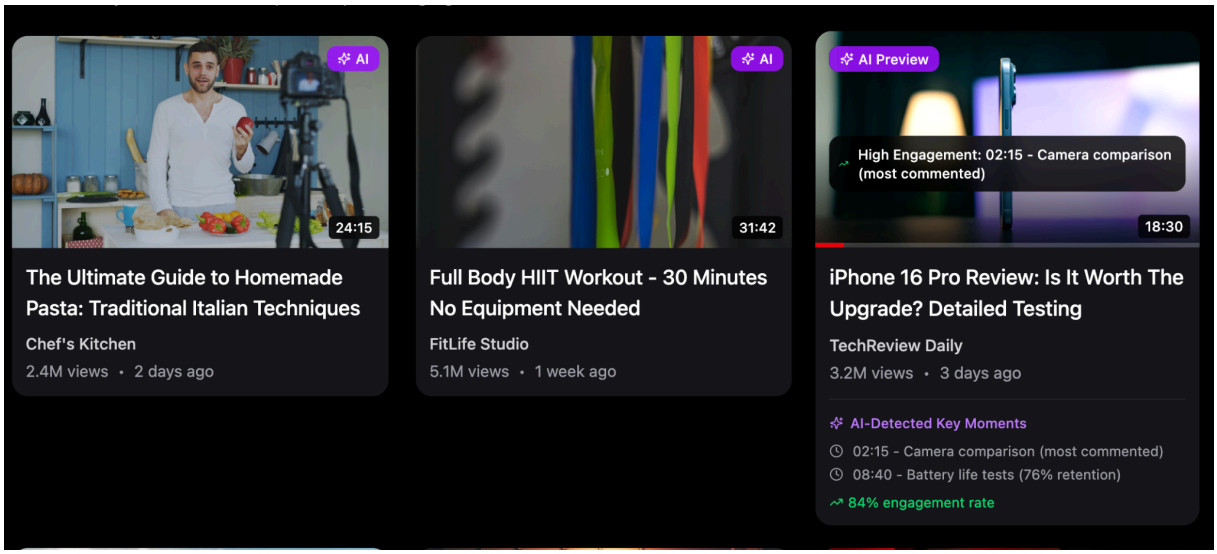
The proposed solution is an AI-generated, data-driven preview clip that surfaces on the video thumbnail during browse giving users a 30 - 60 second 'try before you commit' experience before they open any video.

The engine works in layers. At the lightest touch, a silent 15 to 30 second autoplay moment appears when a user hovers over a thumbnail, selecting only the highest-impact frame automatically. If the user wants more, they can unlock a user-controlled 30 to 60 second preview that understands what kind of viewer they are and shows them the most relevant cut. And underneath all of this sits a personalization model that learns watching behavior over time to make every preview smarter.

Layer	What It Does
Layer 1: Smart Hover Preview	15 to 30 sec silent autoplay on thumbnail hover. AI picks the single highest-impact moment.
Layer 2: Dynamic 30 to 60 Sec Preview	User-controlled extended preview with three selectable modes: Quick Gist, Insight Mode, and Drama Mode.
Layer 3: Interactive Timeline Map	AI-detected chapter markers with clickable entry points to the most replayed, most commented, and most controversial moments.
Layer 4: Personalization Engine	Preview style adapts dynamically per user based on rewatch behavior, drop-off patterns, and content preferences.

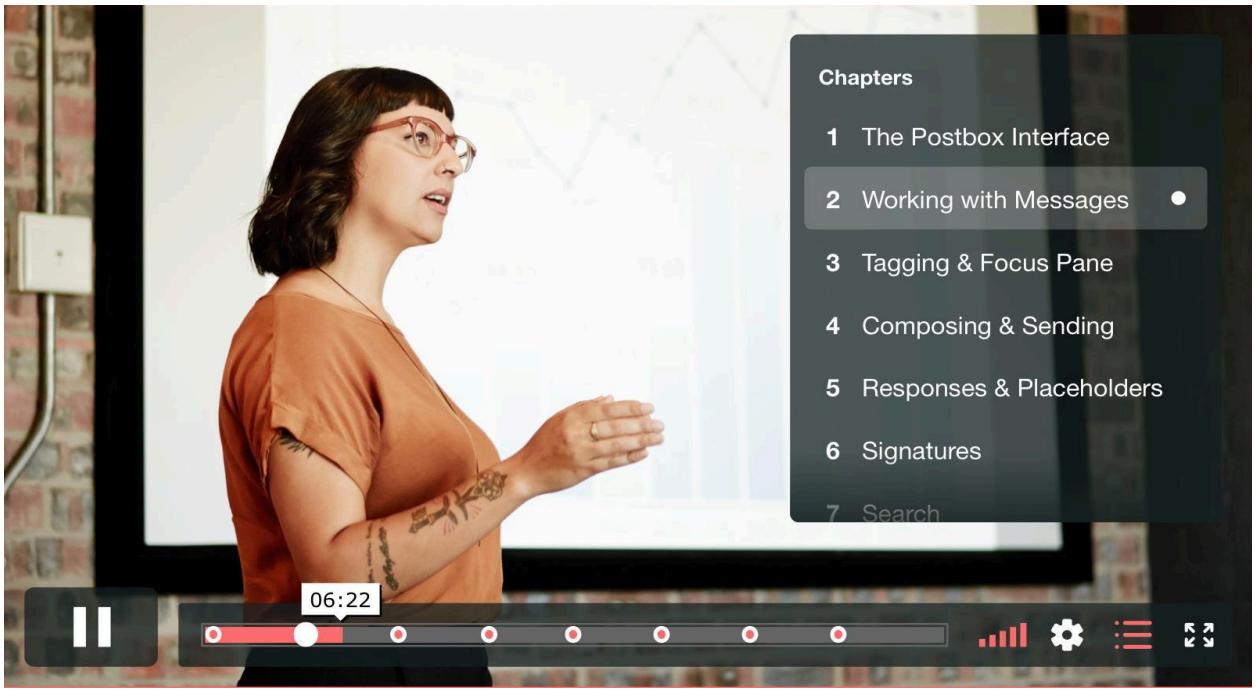
Layer 1- Smart Hover Preview

15 to 30 sec silent autoplay on thumbnail hover. AI picks the single highest-impact moment based on **most replayed segment** and high **comment activity timestamps**..



Layer 2- Dynamic 30 to 60 Sec Intro
User-controlled extended intro section with three selectable modes: Quick Gist (Summary), Insight Mode (Can be used by learners) and Drama Mode (Eg. Best scenes/moments for entertainment)

Layer 3- Interactive Timeline Map
AI-detected chapter markers with clickable entry points to key chapters for learners, the most replayed, most commented, and most controversial moments.



9.3 How the AI Decides What to Show

The model ranks candidate segments across four signals, weighted as follows:

Signal	Weight	Source	Why It Matters
Audience retention peak	40%	YouTube Studio graph	Predicts where existing viewers stay engaged
Most-replayed segment	30%	Replay heatmap data	Indicates high-value, re-watchable moments
Comment activity timestamp	20%	Comments API	Mark's moments that provoked a reaction
Visual/audio energy	10%	Video analysis	Avoids static, low-stimulus clips in preview

9.4 Clip construction logic

The number of segments in the preview is adaptive based on video length:

Video Length	Preview Duration	Segments Selected	Selection Rationale
< 10 minutes	20-30 seconds	1 segment	Single highest-retention peak
10-30 minutes	30-45 seconds	1-2 segments	Peak + highest comment-density timestamp
30+ minutes	45-60 seconds	2-3 segments	Peak, most-replayed, and structural climax

9.5 Smaller Solutions That Complete the Picture

The Preview Engine is the centerpiece, but discovery is not a single moment. These smaller, actionable solutions address the full journey from browsing to finishing a video.

- **Smart Chapter Navigation** AI auto-detects structure inside long-form videos and surfaces chapters as clickable entry points. Users no longer face a 40-minute wall. They can jump directly to the hook, the controversy, the key argument, or the conclusion.
- **Adaptive Algorithm Weighting** If a user repeatedly skips Shorts or exits them quickly, the recommendation engine automatically reduces Shorts weight and surfaces more long-form content. The feed learns without the user having to configure anything.
- **LLM-Powered Search** Users can search for a concept, a moment, or a specific argument within a video, not just a title. The search understands intent and returns timestamped entry points directly to the relevant section.

9.6 One Moonshot Idea

MOOD-AWARE PREVIEW SELECTION

What if YouTube knew how you were feeling before you even started scrolling? Phone sensors already passively capture micro-signals, including time of day, movement patterns, ambient light level, and how fast you are swiping through content. Combined with recent watch history and engagement speed, the system could infer whether you are in a focused, exploratory, or passive mode right now.

In focused mode, the preview leads with the core argument and framework. In entertainment mode, it opens with the tension or the punchline. In passive mode, it picks the most visually stimulating and emotionally resonant 10 seconds. The same video shows up differently to the same person depending on when they encounter it.

> ***Moonshot: Preview content adapts not just to who you are but to what state you are in right now. One video. Infinite entry points. Personalized to the moment, not just the profile.***

10.Success Metrics & KPIs

10.1 Primary Metrics

Metric	Definition	Target	Measurement Window
Hover-to-click conversion rate	% of preview plays that result in full video opens	+25% vs static thumbnail baseline	90 days post-launch
Long-form average watch time	Mean watch duration across long-form sessions	+15% from pre-launch baseline	90 days post-launch
Short → Long conversion rate	% of Shorts sessions that result in ≥1 long-form watch	+20% vs pre-launch	90 days post-launch

10.2 Secondary Metrics

Metric	Definition	Proxy For
Early drop-off rate (0-30 sec)	% of long-form opens that exit within the first 30 sec	Improved expectation-setting from the preview
Preview engagement rate	% of eligible thumbnail views that trigger a preview	Feature adoption and UX friction
Creator Preview Engagement Score	Avg watch% of long-form opens originated from preview	Content-preview quality alignment

10.3 Guardrail Metrics

The following metrics must not regress as a result of this feature:

- Shorts daily active users - preview should convert, not cannibalise
- Creator upload frequency - if creators perceive AI previews as a threat to their content framing, upload rates may decline
- Thumbnail click-through rate on non-preview surfaces - confirms preview is additive, not substitutive

The goal is simple: make long-form video feel less like a gamble and more like a confident choice. The Preview Engine is how we get there.