

Model: Sonnet 5 High

Date: Aug 19, 2026

Prompt:

Before I build this out, the exposure design matters a lot for how the instrument is structured — getting it wrong would mean redoing most of the survey. Two quick questions:

Are the 30-second and 60-second spots two lengths of the same 6 creative concepts, or 12 fully separate ads (6 unique 30s + 6 unique 60s)? Same 6 concepts, cut in 2 lengths each
How should each respondent experience the ads? A few ads per respondent, rotated (sequential monadic)

I'm designing a sequential monadic test for 6 creative concepts across two cuts each, totaling 12 stimuli, and thinking through how to structure rotation so each respondent only evaluates a manageable subset rather than all 12 to avoid fatigue and order bias.

I'm leaning toward having each respondent see 3 different concepts (not repeating a concept across lengths, since that would bias recall), using a rotation design to balance order effects while still getting enough sample per concept for statistical power, with length treated as a separate analytical dimension.

For sample size, I'll flag that exact n depends on client budget and required precision, but typical ad testing aims for roughly 100-150 respondents per concept exposure cell. I'm leaning toward each respondent evaluating just 2 ads sequentially within their assigned length condition to minimize fatigue, using a balanced incomplete block design across the 6 concepts so every pair appears together evenly and position order is controlled.

I'll present this as an illustrative placeholder rather than a fixed number, noting final sample size depends on precision needs and client budget I don't have.

Moving to the actual instrument structure, I need sections for informed consent, screener criteria (age, category usage), but I realize the user never specified what product or brand these holiday ads are for — that's a critical gap I need to flag.

Rather than inventing a fictional brand or category, I'll build the template with clearly marked placeholders like [BRAND/CLIENT NAME] and [PRODUCT CATEGORY] so the structure stays adaptable without fabricating details that weren't given.

Now I'm sketching out the screener section — age quotas, category usage and purchase intent tied to a placeholder category, a holiday shopping intent question since this is a seasonal campaign, and device/screen requirements for watching the video ad properly.

I'm also adding quality checks like attention checks and bot detection, MR industry exclusions, and noting the 30s/60s length split is handled through randomization rather than screening.

Then I'm moving into the pre-exposure section, considering whether to include optional baseline brand awareness and category involvement measures before getting into the sequential monadic ad exposure design with two ads per respondent.

For the exposure block, I'm outlining the flow: embedding the video player for the assigned length condition, forcing full viewing before progression, then hitting respondents with the full measurement battery immediately after each ad before showing the next one. I'm drafting the standard ad-testing measures—overall appeal, uniqueness, relevance, and an open-end for message clarity.

I'm also adding sharing likelihood, length appropriateness given the 30/60 second test, and open-ended likes/dislikes questions, then moving into post-exposure comparative items like forced-choice preference ranking once respondents have seen both ads, plus diagnostic follow-ups and demographic/firmographic questions to close out the survey.

I should note the need for randomization instructions for the programmer—rotating ad order, balancing concept by position, splitting by length condition—and flag that sample size guidance is illustrative since actual requirements depend on confidence level and budget I don't have. I'll also sketch a brief analysis plan for how scores will be used to select the top concepts.

This is clearly a document deliverable given its length. I'm weighing whether the writing-style skill applies here—it's meant for content like articles, not necessarily a technical survey instrument with its own formal conventions around scale points and routing logic, so I'm leaning toward not invoking it.

For format, I'll default to a markdown artifact since there's no explicit signal she wants a Word doc, and offer docx conversion afterward. Given the length, I should build this properly as a structured artifact.

No markdown skill exists here, so I'll just create the file directly. This survey instrument is structured reference content she'll use elsewhere, which clearly qualifies as an artifact, so I'll write it as a markdown file in the outputs folder and present it in one pass rather than building it section by section.

I'm planning the structure now: placeholders for client name, product category, campaign claims, audience definition, incentive, and field dates, with a design of six concepts each cut at :30 and :60, tested between-subjects.

I'm deciding to keep length constant within each respondent rather than mixing durations, splitting the sample 50/50 into a :30 cell and a :60 cell, then using a balanced rotation within each cell so respondents see two of the six concepts with concepts, pairings, and exposure order all roughly balanced.

For sample size, since each respondent sees 2 of 6 concepts, exposures per concept equal a third of the cell size, so hitting roughly 100 exposures per concept-length cell would require about 300 respondents per length cell.

I should be upfront that I don't actually know the client's required sample size or confidence level, so I'll flag this as a placeholder based on common ad-testing norms (~100-150 completes per cell) and offer to recalculate once they share their target margin of error. Now I'll move on to drafting the actual survey questions, scales, and skip logic.

I'm outlining the programmer/rotation notes covering length cell assignment, concept rotation, and order rotation, then laying out the five survey sections from screener through close. I'll write full question text with scale labels, response options, and bracketed programmer instructions before creating the file.

I'm estimating the survey length at 250-350 lines and working out LOI timing based on ad exposure loops, question batteries, screener, and demographics—landing around 12-15 minutes. I'll use bracket placeholders instead of inventing brand names or product claims, draft standard open-end wording for message takeaway, and pull together a typical emotion checklist for ad testing.

Now I'm planning the file structure, including a results interpretation section and a rotation matrix for pairing the six concepts—weighing whether to use a fixed round-robin pairing schedule across versions or simpler random assignment with quotas to balance exposure...

Given the sequential monadic design, here's how I'd structure it: 6 concepts × 2 lengths = 12 stimuli, with length as a between-subject split (each respondent sees only :30s or only :60s) and concept exposure rotated 2-per-respondent within their length cell. That keeps burden manageable and controls for order effects.

One thing I don't have: what's being advertised, who the client/brand is, or the target audience definition. I'm not going to invent a fake brand or category — I'll build the instrument with clearly marked placeholders you drop your specifics into. Same goes for sample size: I don't know your budget or required precision, so I'll show the math with a placeholder assumption rather than pretend I know a number.

Viewed a file, created a file, read a file

The design: length is a between-subject split (:30 vs. :60 cell), concepts rotate sequential monadic at 2-per-respondent to keep fatigue in check, and the whole thing is built to funnel into a scorecard for picking the top 2.

A few gaps I flagged with placeholders rather than guessing: brand/category name, the specific claims to test in the aided message-communication battery, target audience definition, and required sample size (that last one I can size properly the moment you have a target margin of error or a budget ceiling — just say the word). Want me to convert this to a Word doc for client sharing, or drop in real client/category details now?

[CLIENT/BRAND NAME] Holiday Campaign Ad Concept Test

Survey Instrument & Fielding Notes

1. Study Overview

Objective: Test six holiday ad concepts (each produced as a :30 and a :60) to identify the two concepts the client will move forward with for the holiday campaign.

What I don't know yet (placeholders below need your input):

- Brand/client name, product category, and the specific claims or key messages each ad is meant to land
- Target audience definition (beyond "holiday shoppers") — category buyers? existing customers? a specific demo?
- Required sample size / precision, and budget or field-length constraints
- The client's decision criteria for picking the top 2 (I've proposed a default scorecard below — flag if they already have one)

I haven't guessed at any of these; they're marked [PLACEHOLDER] throughout so you can drop in the real details before fielding.

Design Summary

Element	Approach
Stimuli	6 concepts (A–F), each in :30 and :60 cuts = 12 total executions
Length	Between-subject. Each respondent is randomly assigned to the :30 cell or the :60 cell and sees only that length throughout — this keeps the length comparison clean and mirrors how a person would actually be exposed in-market.
Concept exposure	Sequential monadic. Within their length cell, each respondent views and rates 2 of the 6 concepts , back to back, rating each fully

Element	Approach
	before moving to the next. Two ads is the ceiling I'd recommend for video ad testing — a third exposure measurably increases fatigue and flattens differentiation between concepts.
Rotation	Concepts are shown in rotation so each of the 6 gets roughly equal exposure within each length cell, and viewing order (1st vs. 2nd) is rotated to control for order/fatigue effects. See Section 2 for how to set this up in Qualtrics/Decipher.
Post-exposure	After both ads, a direct forced-choice comparison between the two the respondent saw, plus a diagnostic follow-up.

Sample Size — placeholder math, not a real number

I don't know your target precision or budget, so I can't tell you the "right" n. Here's the mechanic so you can plug in your own target:

- Each respondent contributes 2 concept exposures out of 6 possible concepts, so within a length cell, exposures per concept $\approx (2 \div 6) \times \text{cell size}$.
- If you wanted, say, ~100 completed exposures per concept (a common — not universal — floor for stable mean comparisons in ad testing), you'd need a cell size of ~300, so **~300 :30 respondents + ~300 :60 respondents $\approx$ 600 total.**
- If your precision needs are tighter (e.g., you need to detect small differences between close concepts) or you need reliable subgroup cuts, that number goes up. I'd size this properly once you know the required margin of error — happy to run that calculation when you have a target.

Estimated Length of Interview (LOI)

Screeners (~2 min) + 2 ad exposures with full rating battery (~4–5 min each) + comparison (~1 min) + demos (~2 min) $\approx$ **12–15 minutes**. Two :60 ads pushes this toward the higher end — worth noting if incentive/budget is tied to LOI.

Suggested Decision Framework (flag if client already has one)

I don't know how the client wants to weight the metrics below, so here's a starting scorecard — swap in their actual priorities:

- **Overall Appeal** (Top-2-Box, 5-pt scale)
 - **Purchase/Consideration Intent** (Top-2-Box)
 - **Message Communication** — did the unaided open-end and aided battery show they got the intended message?
 - **Brand/Sponsor Linkage** — % correctly attributing the ad to [CLIENT/BRAND NAME] unaided
 - **Distinctiveness** vs. other holiday ads
 - Break ties with open-end diagnostics and the head-to-head comparison data
-

2. Programmer / Rotation Notes

1. **Length assignment:** Randomize respondents 50/50 into Length Cell A (:30) or Length Cell B (:60) at the start of the ad-exposure block. Do not let a respondent see both lengths.
 2. **Concept rotation:** Within each length cell, randomly select 2 of the 6 concepts per respondent using an "evenly present elements" / quota-balanced randomization so each concept accumulates roughly equal exposures across the cell. If your platform supports it, a balanced incomplete block design (each concept paired with each other concept roughly equally often) is the more rigorous option — quota-randomization is the practical default most platforms (Qualtrics, Decipher) handle natively.
 3. **Order rotation:** Randomize which of the respondent's 2 assigned concepts plays first vs. second.
 4. **Force full playthrough:** Disable the "Next" button until the video's runtime has elapsed (or use platform-native video-completion logic) so ratings reflect a full viewing.
 5. **Straightliner/speeder checks:** Include one attention-check item in the rating battery (e.g., "Please select 'Somewhat agree' for this item") and flag completes under a minimum plausible time-per-video.
 6. **Screen-out on device:** Ensure device supports video playback with sound; screen out respondents who can't play video/audio.
-

3. SECTION 1 — Screener

S1. Age What is your age? [Open numeric] [TERMINATE if under 18, or outside client's target age range – PLACEHOLDER: confirm target age range]

S2. Gender (for quota/demo purposes only) [Standard gender item – PLACEHOLDER: use client's standard demo question if they have one]

S3. Category/Brand Screen [PLACEHOLDER – e.g., "Have you purchased [PRODUCT CATEGORY] in the past 6 months?" or "Are you the primary/shared decision-maker for holiday gift purchases in your household?" Needs product category and target-audience definition from client.]

S4. Holiday Shopping Intent Do you plan to do any holiday shopping (gifts, decorations, food, travel, etc.) between now and the end of the year?

- Yes, definitely
- Yes, probably
- Not sure
- Probably not
- Definitely not

[TERMINATE "Probably not" / "Definitely not" – confirm with client]

S5. Industry Exclusion Do you or does anyone in your household work in any of the following industries?

- Market research or advertising
- [PLACEHOLDER: client's industry/category, to avoid competitor employees]
- None of the above

[TERMINATE if any box other than "None of the above" is checked]

S6. Device Check Are you currently using a device with sound available (headphones or speakers)? You'll be asked to watch a couple of short video ads with sound on.

- Yes
- No [TERMINATE]

4. SECTION 2 — Ad Exposure & Evaluation (*looped twice — once per assigned concept*)

[Programmer: repeat this full block for Concept Exposure 1 and Concept Exposure 2, per the respondent's rotation assignment. Video shown is the :30 or :60 cut matching the respondent's length cell.]

Instruction to respondent: "You're about to watch a short video ad. Please watch the entire thing with your sound on, just as you would if you came across it online or on TV. When it's finished, we'll ask you a few questions about it."

[VIDEO PLAYER: Concept [X], :[30/60] cut]

Q1. Overall Appeal Overall, how would you rate this ad? *5-pt scale: Excellent / Very good / Good / Fair / Poor*

Q2. Unaided Message Takeaway (*open end — ask before any aided questions*) In your own words, what was this ad trying to tell you or get you to do? [*Open text*]

Q3. Relevance How relevant is this ad to you and your holiday season? *5-pt scale: Extremely relevant / Very relevant / Somewhat relevant / Not very relevant / Not at all relevant*

Q4. Distinctiveness How different or similar is this ad compared to other holiday ads you've seen? *5-pt scale: Extremely different / Very different / Somewhat different / Not very different / Not at all different (very similar to others)*

Q5. Believability How believable is this ad? *5-pt scale: Extremely believable / Very believable / Somewhat believable / Not very believable / Not at all believable*

Q6. Emotional Response (*multi-select — adjust word list to fit the creative*) Which of the following, if any, describe how this ad made you feel? Select all that apply.

- Happy
- Excited
- Warm / sentimental
- Nostalgic
- Amused
- Inspired
- Indifferent
- Annoyed
- Confused
- Stressed
- None of these

[PLACEHOLDER: swap in emotion words that match the creative's intended tone once client shares the concepts]

Q7. Aided Message Communication (*agree/disagree battery — needs client's actual claims*)

[PLACEHOLDER – e.g., "Please indicate how much you agree or disagree that this ad communicated the following: '[Specific claim/offer #1]' /

'[Specific claim/offer #2]'" – 5-pt agree/disagree scale. Cannot write this without the client's actual campaign messaging.]

Q8. Brand/Sponsor Linkage — Unaided Who do you think this ad is for? *[Open text]*

Q9. Brand/Sponsor Linkage — Aided Was this ad for **[CLIENT/BRAND NAME]**?

- Yes
- No
- Not sure

Q10. Purchase/Consideration Intent After watching this ad, how likely are you to consider **[CLIENT/BRAND NAME / PRODUCT CATEGORY]** for your holiday shopping this year? *5-pt scale: Extremely likely / Very likely / Somewhat likely / Not very likely / Not at all likely*

Q11. Likelihood to Share How likely would you be to share this ad with friends or family (e.g., forward, post, mention it to someone)? *5-pt scale: Extremely likely / Very likely / Somewhat likely / Not very likely / Not at all likely*

Q12. Length Perception Did the length of this ad feel:

- Too short
- About right
- Too long

Q13. Open-End Likes What, if anything, did you like most about this ad? *[Open text]*

Q14. Open-End Dislikes / Suggestions What, if anything, did you dislike about this ad, or what would you change? *[Open text]*

[Attention check — place once, in either the first or second loop only, not both:] **QC. Attention Check** This question is to make sure you're reading carefully. Please select "Somewhat agree" for this item. *5-pt scale: Strongly agree / Somewhat agree / Neither / Somewhat disagree / Strongly disagree*

5. SECTION 3 — Post-Exposure Comparison

[Shown once, after both ad exposures are complete. Reference both concepts by whatever labels were used, e.g., "Ad 1" and "Ad 2" — don't reveal concept names/numbers to the respondent.]

C1. Overall Preference Now that you've seen both ads, which did you prefer overall?

- Ad 1
- Ad 2
- No preference / liked them equally

C2. Preference Reason (*open end*) What made you prefer that ad over the other? [*Open text — skip if "no preference"*]

C3. Which Felt More For You Which ad felt more relevant to your holiday plans?

- Ad 1
 - Ad 2
 - No preference
-

6. SECTION 4 — Demographics

[Standard demo battery – using placeholders since target audience isn't defined yet]

- Household income [*PLACEHOLDER: client's standard bands*]
 - Region/state
 - Household composition (presence and ages of children in household — relevant given holiday/family themes)
 - Race/ethnicity [*if needed for client's reporting standards*]
 - Holiday shopping behavior: typical budget range, primary channels (online vs. in-store), timing (early bird vs. last-minute)
-

7. SECTION 5 — Close

Thank you for your feedback! Your responses will help [CLIENT/BRAND NAME] create better holiday advertising.

[Incentive/redemption info – PLACEHOLDER]

Next Steps Before Fielding

1. Fill in all [PLACEHOLDER] items — brand name, category, target audience, specific claims, incentive.
2. Confirm target sample size once precision/budget requirements are set.
3. Confirm the decision-framework weighting with the client, or provide their existing one.
4. Have the 12 video files (6 concepts × 2 lengths) hosted and ready for the survey platform before programming.
5. Soft-launch with ~30–50 completes to check LOI, video playback, and open-end quality before full field.