

# Testing Protocol

## Design

This study will include two experiments that use both a within-subjects design and a between-subjects design. These experiments will offer ample qualitative data to implement into future improvements for a potential feature implementation stack inside of the Letterboxd product experience.

The tests are aiming to assess the psychological implications of the feature sets impact in accordance with the drafted goals in the design brief; first does the feature feel integrated into the total Letterboxd brand and product experience, second does the feature improve a willingness to engage in the app as a genuine 'social media alternative', third does this increase brand affection for the Letterboxd product, fourth does this feature lead to an increased feeling of belonging/community, fifth does the feature set increase the likelihood of initiating action into regular cinema-going behaviour, sixth will the feature set increase the willingness to engage deeper in the film world.

The foundation of the hypothesis for the design is that facilitating the frameworks in behaviour that create action and engagement as a second nature, frictionless experience will - as posited by psychologist Daryl J Bem - forge a stronger 'cinephilic' identity. The aim of the tests is to assess whether there is a potential opportunity for facilitating social engagement through the film industry, helping to initiate a sense of belonging through shared cultural artefacts and identity. Further the intention is to qualify the idea into a real-world product focused portfolio by proposing that these features have the potential to facilitate peripheral economic spheres and create a platform for more social settings in a hybrid social world. This could be something worth thinking about and designed.

## Participants

Leading these testing protocols whilst being mindful of a lean budget, the experiments will leverage Letterboxd's existing user base and low-cost digital tools. For **Experiment 1**, we would recruit two carefully defined cohorts drawn directly from our core Gen Z cinephile demographic—active Letterboxd users aged 18–25 who log at least two film interactions per week.

Specifically we'll invite 3 participants recruited via in-app metrics (those who log  $\geq 2$  films a week) and social media reach outs balanced by gender, geography, and Letterboxd tenure (new, intermediate, veteran). Screening questions will verify device comfort and baseline brand sentiment, ensuring a homogenous starting point. By using each person as their own control—exposing them to both "before" and "after" prototypes—we dramatically boost internal validity: individual differences (things like film taste or mood) are held constant, letting us isolate the feature's impact on perceived integration, engagement willingness, and brand affection.

By having each participant test both versions of the interface remotely—on their own devices (downloaded via TestFlight)—we eliminate facility rental costs and maximize internal validity through a within-subjects design.

For **Experiment 2**, we'll enlist another 6 similar users and randomly assign them to Control or Feature groups. Randomisation neutralises selection bias, while stratified sampling by region and social-media usage ensures that findings generalise beyond any one cohort—elevating external validity. We'll further screen for recent cinema attendance to capture a realistic baseline for the diary phase. Recruitment remains digitally native and by combining in-app recruitment, remote testing, digital surveys we achieve rigorous scientific control—counterbalancing order effects and minimising sampling bias—while staying well within a small-firm budget.

Anchoring both studies in rigorous screening, stratification, and random assignment ensures our sample reflects the real-world user base Letterboxd aims to delight—maximizing both the reliability and applicability of our insights. The limitation to this design however is a focus on iOS design as the prototyping and code snippets would only be suited for testing within the TestFlight download within the timeframe. Furthermore if allowed for more time and scale the intention would be to sample upwards of 40 users per group across the aforementioned demographic profile.

## Apparatus

### For experiment 1:

#### APPARATUS

- Prototypes: Two medium-fidelity interactive mock-ups (A = existing app; B = app + new feature downloaded via TestFlight on iOS App Store) on the same device (tablet or desktop).
- Brand Integration Scale (custom Likert 1–7: “This feature feels like part of Letterboxd”)
- Social-Media Alternative Scale (adapted from U&G theory)
- Brand Affection Scale (brand love/attachment items).
- Data capture: Screen-recording software + log of click-through rates.
- Environment: Quiet lab room or controlled online platform (e.g. Zoom with shared screen).

### For experiment 2:

#### APPARATUS

- Two cohorts ( $n \approx 3$  each), randomly assigned:
- Control group: Letterboxd mock-up without feature
- Treatment group: Mock-up with feature downloaded via TestFlight on iOS App Store
- Post-session survey:
- Sense of Community Index (SCI)
- Cinema Intention Scale (“How likely are you to buy a cinema ticket in the next month?”)
- Deeper Engagement Scale (e.g. intention to join film clubs, purchase merch, discuss films).
- Follow-up diary: 14-day online log of real-world behaviours (cinema visits, PVOD rentals, merchandise purchases and paraphernalia).

## Procedure

### Experiment 1: Within-Subjects “Feature Integration & Brand Affection”

**Goals:** 1) Brand integration 2) Engagement willingness 3) Brand affection

#### 1. Recruit & Consent ( $n \approx 3$ Gen Z cinephiles, balanced gender/age)

- **Screeners:** Deploy a short Typeform asking age (18–25), Letterboxd usage ( $\geq 2$  sessions/week), device type, and basic film fandom questions.
- **Quota sampling:** Ensure equal representation by gender and tenure on the platform (new: <6 months, mid: 6–18 months, veteran: >18 months).
- **Ethics & confidentiality:** Present an online consent form detailing purpose (“usability & brand study”), data usage, voluntary withdrawal, and anonymization procedures.

#### 2. Baseline Survey

- **Demographics:** Collect age, gender, geography, and device familiarity to control for extraneous variance.
- **Letterboxd familiarity:** Rate on a 1–7 scale (“How well do you know Letterboxd’s core features?”).
- **Cinema habits:** Record average monthly cinema visits and home-viewing spend to establish behavior baselines (ties to Goals 5–6).

---

**Session 1 (Prototype A, no feature)****1. Instructions & mindset priming**

- Read aloud: “Explore this version as you would naturally—there’s no right or wrong.” Framing reduces evaluation apprehension.

**2. 5 min free exploration**

- Capture screen recordings: navigation paths, click latency, and hesitation events (helps infer perceived integration and ease).

**3. Immediate Scales 1–3**

- Administer Brand Integration, Social-Media Alternative, and Brand Affection scales in randomized order to prevent response biases.
- Data logged automatically to a secure Google Sheet.

**4. 5-Minute Break (Neutral Video Clip)**

- **Stimulus:** A 2-minute nature clip (e.g., forests)—standardized across participants to reset mood and working memory.
- **Rationale:** Clears recency effects and reduces carryover of emotions from Session

---

**Session 2 (Prototype B, with feature; counterbalanced)****1. Counterbalancing**

- Half the cohort sees B first, half sees A first; order recorded to model any residual order effects statistically.

**2. 5 min free exploration**

- Same logging metrics collected—ensures direct within-subject comparison.

**3. Repeat Scales 1–3**

- Identical survey prompts allow paired t-tests on integration, engagement willingness, and affection.

**4. Debrief & Open-Ended Feedback**

- Prompt: “Describe in your own words how the new feature changed your experience.”
- Qualitative responses will be coded for emergent themes (e.g., “felt seamless,” “distracting,” “more social”).

---

**Experiment 2: Between-Subjects “Community, Cinema-Going & Deeper Engagement”**

**Goals:** 4) Community feeling 5) Cinema intent 6) Deeper engagement

**1. Recruit & Consent (n≈6 Gen Z Letterboxd users)**

- **Semi-random sampling:** Gather a randomized subset of active 18–25 year-olds; screen for ≥1 cinema visit in past month. Use in app-data to screen and validate relevance of subject to results.

**2. Random Assignment**

- **Mechanism:** Auto-generated via Apple Numbers RAND() function; participants are labeled “Control” or “Feature.”
- **Stratification:** Balanced by region (e.g., North America vs. Europe) and average weekly usage to uphold external validity.

**3. Exploration Phase (10 min)**

- **Task prompt:** “Use the app to plan your upcoming film activities—post a comment, join a collection.”
- **Observation:** While unrecorded, click-count logs and dwell times capture engagement depth.

#### 4. Immediate Survey

- **Sense of Community Index (SCI):** 8 items tapping belonging, mutual support, and shared identity.
- **Cinema Intention Scale:** Likelihood (1–7) to attend a theater screening or rent PVOD within 30 days.
- **Deeper Engagement Scale:** Intent to join film forums, purchase merch, or attend festivals.

#### 5. Diary Phase (Days 2–8)

- **Platform:** Daily Form link sent via SMS/reminder/email at 7 pm local time.
- **Entries:** Number of cinema/PVOD viewings, snack/merch purchases, social film discussions — timestamped to validate compliance. (Double checked abasing authentic in app behaviour to help increase data relevance)
- **Compliance checks:** Automated reminder if no entry by 10 pm, reducing missing data.

#### 6. Return Survey (Day 9)

- **Behavior vs. Intent:** Compare diary logs against initial intention scores—paired analysis tests predictive validity.
- **Final reflection:** Single-item open text: “How did this feature affect your real-world film habits and sense of community?”

### Linking back to the scientific method:

- Each step explicitly controls or measures confounds (e.g., mood resets, counterbalancing, stratified randomization).
- Within- and between-subject designs isolate feature effects on psychological constructs (integration, affection, belonging) and behavioral outcomes (attendance, spending).
- Quantitative scales plus qualitative coding ensure both internal rigor and ecological validity, directly testing whether your Letterboxd feature fosters deeper cinephilic identity, social engagement, and economic value for Gen Z.

## Appendix

### Variables and Experimental Design Elements

	Experiment 1	Experiment 2
<b>Design</b>	Within-subjects (repeated)	Between-subjects (indep.)
<b>IV(s)</b>	Prototype version (A vs B)	Group assignment (Control vs Feature)
<b>DV(s)</b>	Likert scores on Scales 1–3	SCI, cinema intention, engagement scores; diary outcomes
<b>Extraneous vars</b>	Prior brand love; device; mood	Recruitment channel; personal film habits; week’s schedule
<b>Order effects</b>	Counterbalanced order	N/A
<b>Demand characteristics</b>	Cover story (“usability study”)	Same cover story
<b>Social desirability</b>	Anonymous responses; filler items	Anonymous; include social-norm items
<b>Evaluation apprehension</b>	Emphasize “no right answers”	Emphasize confidentiality

<b>Experimenter bias</b>	Standardized script; blind admin	Automated online interface
<b>Internal validity</b>	High (each person = own control)	Moderate (randomization)
<b>Construct validity</b>	Validated scales; pilot test items	Same scales; pilot diary instructions
<b>External validity</b>	Lab/online sample → Gen Z cinephiles	Field-style diary → real behavior

## Keys to Scientific Rigor

- Counterbalancing in Exp 1 neutralizes order effects.
- Randomization in Exp 2 neutralizes selection bias.
- Cover story and anonymity minimize demand, social-desirability, and evaluation apprehension.
- Blind protocols (scripted instructions/automated interfaces) guard against experimenter bias.
- Validated instruments (or pilot-tested scales) ensure construct validity.
- Sample selection (targeting Gen Z Letterboxd users) and the week-long diary boost ecological and external validity.