

Experiment Card

Purpose. Turn a risky assumption into a single, testable hypothesis with a stop rule agreed in advance.

What it captures. A one-page record naming the hypothesis, the user and context, the method and sample, the primary and secondary metrics, the business stop rule and ethical harm signal, the threshold rationale and confidence (provisional, benchmarked, or evidence-based), the risks and mitigations, the instrumentation and data sources, and the decision date and owner.

When to use it. Before running any experiment or pilot, so the team agrees what will count as evidence and when to stop.

Common mistake. Setting or softening the stop rule after seeing the data. Project Swift set its stop rule, the defect-rate ceiling, the completion floor, and the override-usage signal before the test ran, so a weak result could not be reinterpreted as a success.

Decision rule. If a result crosses the stop rule, stop the test rather than extend it to rescue the hypothesis. A test that cannot fail is not an experiment.

Hypothesis *include a specific metric, a success threshold, and a time box*

User and Context

Method and Sample

Primary Metric

Secondary Metrics

Business Stop Rule *set before the test; do not soften after seeing data*

Ethical Harm Signal

Threshold Rationale and Confidence *how each threshold was set, and whether it is provisional, benchmarked, or evidence-based*

Threshold	Confidence	How it was set
Primary metric threshold		
Stop-rule thresholds		
Other thresholds		

A threshold set without a baseline is provisional, and the team revises it after the first cycle, before the next card binds it.

Risks and Mitigations

-
-
-

Instrumentation and Data Sources

Decision Date and Owner