

Facilitator Guide

Workshop delivery and worked responses

Author	Jason Hreha
Version	1
Audience	Product and innovation leaders
Date	September 5 2026
License	Original content CC BY 4.0

How to use this document

Use this guide with the Student Casebook and Decision Workbook. It contains the proposed workshop, discussion rubric, source and reuse notes, worked responses, and measurement answers. Share the student packet before the worked analysis. A defensible response can differ from the historical outcome.

The three companion files are Decision Workbook, Student Casebook, and Facilitator Guide. Their PDFs and editable Word versions share those names. Use the student material before opening worked responses. Source URLs are reference links; the printed instructions and fictional records support offline practice.

Public facts, participant recollections, instructor analysis, proposed tests, and invented observations are labelled separately. No featured organization is claimed to have used DRIVE, BMF, or BFA. No participant study or independent validation of this teaching pack is reported.

Contents

1. Run a Behavioral Strategy workshop
 2. Teaching notes and reuse
 3. M PESA Worked Decision
 4. Instagram A Worked Decision
 5. Verify A Worked Decision
 6. Measurement Lab Answers
- References and online companions

Run a Behavioral Strategy workshop

Lead a session in which participants choose a target behavior, compare alternatives, and decide what to test. Use the M-PESA case for a first session. Instagram and GOV.UK Verify are alternatives for a group whose work is closer to consumer products or public services.

This is a proposed 90-minute teaching design. The timings below are planning estimates, not observed results from independent delivery. Prior knowledge of the framework is optional. A facilitator should read the case, its worked analysis, and the [teaching notes](#) [1] before the session.

Prepare the session

Download the [complete pack](#) [2], or use the [individual files](#) [3]. Give participants the student casebook and a blank decision workbook. Keep the facilitator guide separate until the discussion reaches the comparison stage. The slides are optional and include speaker notes.

Plan for individuals or small groups who can each produce one recommendation. Give each group paper or an editable workbook. Provide the case in advance if readers need more time or another format. Participants can work through the same material without a projector or an account.

Write this outcome where everyone can see it: **By the end, each group will recommend a bounded next step, explain its evidence, and name a finding that would change the decision.**

The 90 minute agenda

0-8

Activity: Frame the decision

Facilitator action: Explain outcome, behavior, and feature. State that the case does not document historical use of the framework.

Participant output: One sentence describing the decision

8-20

Activity: Read the case

Facilitator action: Open only the student packet. Ask readers to mark a fact, an inference, and an unknown.

Participant output: Three labelled notes

20-32

Activity: Generate alternatives

Facilitator action: Use workbook steps 1-2. Require an option that reduces the user's work, where plausible.

Participant output: Candidate comparison

32-45

Activity: Map the delivery chain

Facilitator action: Use steps 3-4. Ask whose action and support each option needs.

Participant output: Actor chain with evidence gaps

45-60

Activity: Design the test

Facilitator action: Use steps 5-6. Ask for the opportunity denominator, repeat window, acceptable support, and stopping rules.

Participant output: A test plan that someone could run

60-70

Activity: Commit to a recommendation

Facilitator action: Use steps 7-8. Ask each group to state its decision before the worked analysis opens.

Participant output: Recommendation and reversal condition

70-82

Activity: Compare the reasoning

Facilitator action: Read the worked analysis. For M-PESA, compare both responses. Use the discussion rubric below.

Participant output: One retained choice and one revision

82-90

Activity: Transfer the method

Facilitator action: Apply the last prompt to a different, nonconfidential setting.

Participant output: A new candidate and a measurement question

The table totals 90 minutes. Longer reading or discussion requires a longer session or advance reading; do not compress the test plan to preserve every presentation slide.

Prompts at the difficult points

When a group names a feature, ask: “Who does what, in which situation, and how does that action contribute to the outcome?” When it names an outcome, ask what observable action would have to happen first.

When a group chooses its favorite option immediately, ask for the strongest alternative and an option that removes a task from the user. Removing work can transfer work to an agent, operator, employer, or government team; put that actor on the chain.

When a group treats interest as evidence, ask what people actually did, how they were selected, and what help they received. When it treats repeated activity as automatic proof, ask whether the person had another opportunity and whether the action served the intended outcome.

When a group proposes a test, ask another group to explain how it would count success and failure. If the second group cannot tell, revise the record before discussing sample size.

Discuss the recommendation

Use the [six-part rubric](#) [4]. A defensible answer has a clear action and population, serious alternatives, an actor chain, traceable evidence, an executable test, and a proportionate next decision. Agreement with the historical outcome earns no special credit.

Invite each group to state one disagreement with the worked response and the evidence needed to resolve it. A well-supported decision to stop or collect missing evidence can be stronger than a confident proposal to expand.

Transfer exercise

Choose a routine process that can be described without private company or personal information: booking a meeting, returning equipment, or finding an internal document. State the outcome. Generate two candidate behaviors and one option that removes user work. Identify a dependent actor and one observation that could disqualify the preferred option. This is a planning exercise; it does not authorize collecting personal data or running a live study.

After the session

Let participants keep and revise their work. If you want to assess the teaching material, use the optional [host observation record](#) [5]. Record what needed explanation and what the group could do unaided. Do not infer a learning effect or method validation from satisfaction ratings.

Original teaching content: Jason Hreha, BehavioralStrategy.com, teaching pack v1.0 (2026). [CC BY 4.0](#) [6].

Teaching notes and reuse

These notes support the [workshop](#) [7] and [solo route](#) [8]. Read the student packet and worked analysis for your chosen case before leading a discussion. The goal is a defensible decision and a useful test, with uncertainty visible.

The package is a proposed teaching resource. No independent teaching trial or prospective validation study of this package is claimed. The optional record below is a template for future use, not a report of completed delivery.

Source boundaries

Material	What it supplies	What it does not supply
M-PESA packet [9]	Public participant accounts of an early service pilot and its delivery constraints	Complete transaction records, an unbiased comparison population, or a controlled test of Behavioral Strategy
Instagram packet [10]	Founders' retrospective explanation of an early product focus decision	Raw feature-level data, a causal estimate for removing features, or verified framework use
Verify packet [11]	A public audit of identity verification, service dependencies, costs, and targets	A randomized comparison or proof that one psychological constraint explains all noncompletion
Measurement lab [12]	Invented observations for arithmetic practice, plus two separate research examples	A real participant dataset, an effect estimate for the pack, or a universal success threshold

Evidence IDs BS-0079 through BS-0083 lead to the [evidence ledger](#) [13]. The individual pages link their primary sources and identify retrospective material. The printable guide and slide notes retain the relevant references. Learners should distinguish a reported event from the instructor's interpretation and proposed action.

Discussion rubric

This is an informal review aid, not a validated assessment scale. Review each row as **clear**, **needs revision**, or **unresolved because evidence is missing**. Do not add the categories into a numerical fit score or use a cutoff to award certification.

Element	Look for	Ask for a revision when
Outcome and action	A named population, observable action, context, and connection to the outcome	The answer gives only a feature, aspiration, or business metric
Alternatives	Plausible competing behaviors, including less user work where feasible	The comparison restates the favorite option in different words
Actor chain	Recipient, provider, operator, and other dependent actions where relevant	The plan counts only the first user's visible action

Element	Look for	Ask for a revision when
Evidence	Sources and observations separated from inference and unknowns	Interest, selection, or hindsight is treated as causal proof
Test and measurement	Eligible opportunities, repetitions, support costs, observation window, and prespecified rules	Someone else could count a different result from the same records
Next decision	A proportionate action with a stated reversal condition	Expansion follows automatically from activity or a high informal rating

A missing fact is not automatically a learner error. Credit a response that names the gap, limits its claim, and proposes a proportionate way to resolve it. Challenge a response that silently fills the gap with a favorable assumption.

Common discussion errors

A case outcome cannot establish why it happened. Ask what the source observed and which causal alternatives remain. A familiar behavior may still require new skills, trust, incentives, or operational support in a new setting.

Reducing user effort can be a good candidate, but it does not make work disappear. Ask who now performs the task, at what cost, and under which service conditions. A concierge test may help discover a viable service; its staffing and exception handling belong in the record.

A null result limits a claim about the tested intervention, population, and outcome. It does not prove that the underlying need is absent or that every possible intervention will fail. Conversely, an intermediate gain does not establish the final outcome.

Host observation record

Use the [editable text record](#) [14] after a session if useful. Keep observations at group level and omit identifying or confidential material. Participation in teaching does not by itself authorize research use or public quotation of learners' work.

Record the date, pack version, format, group size, prior familiarity in broad terms, actual reading and activity times, and any accessibility or language adaptation. Note which instructions required unplanned explanation. For each rubric row, summarize whether the group could complete it without new help and give an anonymous example of the difficulty.

Record whether the participants changed a decision after the comparison and why. Separate an observed change in a written response from a claimed improvement in real-world outcomes. Note facilitator experience and deviations from the agenda. A future evaluation would need a defined outcome, comparison, consent process where applicable, and a sampling plan before drawing broader conclusions.

There is no collection endpoint or requirement to send this record to the site. Use the [corrections page](#) [15] for a specific source or content correction.

Reuse and adaptation

Original teaching content in teaching pack v1.0 is licensed under [CC BY 4.0](#) [16]. This includes this page, the new decision packets and worked analyses linked from the Education hub, the decision workbook, measurement lab and answers, workshop, solo route, and the original text and data in /downloads/teaching-pack/v1/.

You may share, teach, translate, and adapt that original material, including commercially, with attribution and an indication of changes. Outside publications, quotations, trademarks, fonts, and any third-party template elements retain their own rights; linking a source does not relicense it. The two existing methodology guides and older case notes retain their existing site license unless they carry another asset-level notice. See the [license map](#) [17].

Suggested attribution: “Adapted from Jason Hreha, BehavioralStrategy.com, Public teaching pack v1.0, 2026, CC BY 4.0, <https://behavioralstrategy.com/education/>. Changes: [describe your changes].” For an unchanged copy, replace “Adapted from” with “By” and omit the change description.

Keep the source notes, fictional-data labels, evidence limits, and version with an adaptation. Update the decision, actor chain, denominators, and worked analysis when changing the setting. Remove organization-specific facts that no longer apply. Use the names Behavioral Strategy and DRIVE accurately; the license does not imply endorsement, certification, or affiliation.

The [manifest](#) [18] lists the pack files and their checksums. Version 1.0 was prepared on 5 September 2026. Later revisions should carry their own version and change note.

M PESA Worked Decision

Read after completing the [decision packet](#) [9]. A strong answer connects evidence to a bounded next investment. Agreement with the eventual historical choice is not the marking rule.

This page contains original instructor reasoning and a proposed [Behavior Market Fit](#) [19] test. These are not recovered M-PESA plans. No evidence reviewed for this lesson establishes that the team used DRIVE, BMF or BFA, or that this teaching method improves real decisions.

Response A Test a narrow transfer route

Decision. Allocate the next learning budget to a bounded person-to-person transfer test. Start with people who have an actual transfer need and a reachable intended recipient. Define the target as completing the transfer through to usable money at the other end. Do not commit to a national launch on this evidence.

Why this is reasonable. The alternative-use observations in the packet make a transfer service worth investigating. A test can ask whether the recipient gets value without adding a lending institution's loan-record requirements to every transaction. That removes one class of dependency from the proposed service. It does not remove agents, financial controls or customer support.

Strongest objection. The pilot does not supply a paid, representative transfer funnel. The earlier source's caution about use beyond repayment should reduce confidence in extrapolation. A handful of resourceful participants might account for the most memorable examples. Recruiting only similar enthusiasts would repeat this problem.

Treatment of the other options. Repayment remains plausible because it has a recurring, identifiable occasion and a checkable institutional endpoint. Before expanding it, request evidence that the institution can credit and reconcile payments within an agreed operating budget. An open-ended pilot extension is less attractive than a test with a defined target and reversal rule.

Fit judgment. A desire to support a relative makes the transfer goal plausible for a defined segment. It does not show that entering a number, managing a PIN or finding cash is easy. A [Behavior Fit Assessment](#) [20] should carry those unknowns forward. Assigning a high score to each dimension would conceal the test's purpose.

What changes this decision. Stop expansion if ordinary support cannot deliver recipient access reliably, or if people with a new genuine need consistently choose the existing route at realistic fees. Reconsider repayment if its complete workflow has materially better repeat completion and support economics for a clearly defined population. The choice is conditional on evidence, not on the attractiveness of the remittance story.

Response B Establish delivery capacity first

Decision. Hold consumer expansion. Use the next bounded work period to audit complete payment episodes within the existing pilot population and test a sustainable operating routine. Decide on a new target population only after the service can distinguish delivered value from incomplete or rescued transactions.

Why this is reasonable. The packet leaves the numerator, opportunity denominator and support cost unresolved. Buying a larger transfer test now could produce another ambiguous result if neither endpoint

can be verified. This response gives more weight than Response A to the risk of measuring an operation the team cannot yet support.

The work to fund. A service operations owner would match customer, agent, platform and endpoint records for the same episodes. That owner would classify the reason for each incomplete result and identify a normal support routine that can be staffed and costed. For a loan episode, this includes the correct borrower credit. For a cash remittance, it includes recipient access. Keep the populations and endpoints separate in the report.

A candidate with less user work. Before adding another borrower step, examine whether record matching can be handled inside the service and institution. Compare total customer and staff work with the current route. This is a proposed design question, not a claim that the historical team had this option ready or that automation would necessarily help.

Strongest objection. This approach postpones learning about new transfer customers. An operations review can become an excuse to retain the original product or perfect every edge case. Cap it around the specific missing evidence; do not redesign the whole financial system.

Exit and reversal. In this teaching example, allow two ordinary payment cycles to test the defined routine. That duration is an instructor's planning assumption. If records remain irreconcilable or delivery requires exceptional rescue, stop expansion and revise the route. If ordinary staff can reconcile the full chain within a predeclared budget, proceed to the transfer test below. If too few genuine opportunities occur, report insufficient evidence and choose whether one targeted extension is worth its cost.

These responses have the **same business objective**. Response A values faster learning about a new use. Response B values resolving the delivery uncertainty first. Neither assumes an obligation to keep the loan partnership or advance the remittance service. A learner can choose either while making a clear, testable judgment about which uncertainty is most costly.

A complete proposed transfer test

The following protocol makes Response A concrete. Response B could use it after the operations gate. All amounts of time, counts and thresholds below are **illustrative teaching choices**, not historical M-PESA measurements, universal cutoffs or a study ready to run. No participant testing has been conducted for this lesson.

Population opportunity and endpoint

Recruit consenting sender-recipient pairs in one defined transfer route where an ordinary agent service can be provided at both ends. Require a current, independently described reason to send money. Record the existing route and expected need timing before offering the test. Include variation in phone familiarity; report whom the access requirements exclude.

The unit is a **pair with a genuine transfer opportunity**, not an account, a message or a transaction attempt. A new opportunity requires a new underlying need. Retrying the same failed payment does not create another opportunity.

For this initial cash-remittance test, completion requires all of the following:

1. The sender chooses the service after the fee and expected delivery time are explained.

2. The intended amount is correctly transferred and confirmed by the service record.
3. The intended recipient obtains the cash within the promised period, with matched endpoint evidence.
4. No unresolved amount, recipient or balance discrepancy remains for that episode.

The transfer promise is **within one day** for this teaching example. Retaining funds electronically could be a valid user choice in another product definition; record it here as a different endpoint. Do not silently count it as completed cash delivery.

Support that belongs in the offer

The planned service includes an initial explanation, normal agent assistance, replenishment of cash and electronic balances, a staffed help route and an accountable reconciliation owner. Estimate the cost of that support before the test. Maintain it if it is part of the intended commercial service.

Separate normal assistance from exceptional rescue. A project leader making a special trip, personally funding an agent or handling a customer's credentials would fall outside the planned routine. Log the action, cost and result. An eventual completion can still be recorded, but it does not pass the ordinary-support gate. Never deny necessary help merely to improve the experiment's appearance.

Use the intended fee. If the test must subsidize transactions or devices, disclose the subsidy and limit the claim to subsidized conditions. A reward for completing a transfer changes what the test measures. Compensation for research time should be recorded separately from transaction use.

The packet does not supply a defensible price or support-cost budget. The operator must cost and approve those inputs before running this protocol. Leave commercial viability unresolved until then. A precise completion threshold cannot replace a missing operating budget.

Observation and denominators

Plan an **eight-week observation period**, with a minimum of **40 pairs with a first verified opportunity** and **20 of those pairs with a later verified opportunity**. These small illustrative counts support an exploratory decision; they do not establish population precision or isolate a causal effect. If the expected need occurs less often, change the window before enrollment and explain why.

Measure	Numerator	Denominator and interpretation
First-opportunity completion	Pairs meeting the full cash-delivery endpoint on their first opportunity	All enrolled pairs with a verified first opportunity, including those who choose another route or never attempt the service
Repeat-opportunity completion	Pairs meeting the endpoint on their next genuine opportunity	All enrolled pairs with a verified next opportunity, including those whose first episode failed; report previous completers separately
Ordinary-support completion	Completed episodes within the declared support routine	All verified opportunities; report first and later opportunities separately
Support burden	Total ordinary support time and cost; exceptional rescue reported separately	Show totals and cost per completed episode; avoid making failure costs disappear from the total

Measure	Numerator	Denominator and interpretation
Access and harm	Failed visits, delayed access, errors, loss and unresolved balances	Show counts against the appropriate opportunities and affected pairs, plus each serious incident

Distinguish five states: completed; known not completed; opportunity verified but outcome unresolved; no opportunity occurred, verified; and opportunity status unknown. Unknown outcomes remain in a verified-opportunity denominator. People whose need status is unknown must be reported separately from the verified no-opportunity group.

Report a conservative completion rate that counts unresolved outcomes as unconfirmed, and an upper bound that assumes they completed. If that interval crosses a decision threshold, the result is insufficient for that gate. For pairs whose later need is unknown, also show the result if each had an uncompleted opportunity. That is a sensitivity check, not a measured rate. Record both nonuse and failures, and compare the time and total cost with the participant's existing route without claiming a randomized comparison.

Rules agreed before results

For the exercise, use these provisional rules and challenge their business rationale. A real team would set fees, service promises and resource limits from the intended operation before recruitment.

Decision	Proposed rule
Continue to a larger bounded test	At least 80% first-opportunity and 70% repeat-opportunity full completion; the same thresholds hold within ordinary support; both opportunity minimums are met; the declared support budget is met; no unresolved loss or material balance discrepancy
Revise the service	A named failure in access, comprehension or ordinary operations explains a missed gate and has a feasible, costed repair; retest the affected chain
Stop expansion	An unresolved loss or material balance discrepancy occurs, or the service cannot meet the defined promise and support budget after the scoped repair; resolve affected customer outcomes before further use
Insufficient evidence	Opportunity minimums are missed, need or outcome status is unresolved enough to change a decision, or subsidy and rescue prevent a conclusion about the planned offer

Passing these gates supports the next bounded investment for this route and offer. It does not establish national demand, financial viability at scale, a welfare effect or comparative validation of BMF.

Diagnose a weak result before changing the goal

The following examples are hypothetical diagnoses. They are not M-PESA pilot results.

Observation	Competing explanation to check	Next evidence or action
Senders attempt; recipients cannot obtain cash	Delivery capacity or location fails	Inspect the endpoint and replenishment records before changing motivation messages
People complete only when staff take over	The action requires unsupported skill, or training is poorly designed	Observe which step fails; test a simpler action or costed assistance
First completion is strong; later use is low	Later need is absent, fees change the choice, or the first experience disappoints	Separate missing opportunity from refusal; inspect price and earlier outcome
Borrowers report paying; loan records remain wrong	The institutional handoff fails	Match the episode to the account record; repair reconciliation before more borrower prompts
The service works but costs more to support than it earns	The offer or operating model is not sustainable	Cost the full route; reconsider scope or pricing before expansion

The [measurement lab](#) [12] provides fictional observations for calculation practice. The [test method](#) [21] explains how to adapt the design to another setting.

What happened afterward

Hughes and Lonie report that the launch proposition became **Send Money Home**, with cash conversion, person-to-person transfers and airtime purchasing. Their account also describes subsequent changes to agent controls and balance management. (evidence BS-0079)

In a tenth-anniversary company retrospective, former Safaricom CEO Michael Joseph describes shifting the loan prototype toward transfers, investing in agent training and launching in March 2007. This is later participant testimony, not evidence available to the imagined 2006 meeting. [Read Joseph's account](#) [22].

Later household research addresses a different question. Suri and Jack's 2016 study estimated that increased mobile-money access lifted about 194,000 Kenyan households out of extreme poverty. That is a study estimate from later household evidence, not a count of pilot successes or a result of the proposed test above. [MIT's account of the study](#) [23]. (evidence BS-0007)

The history illustrates how a product direction and its supporting operation can develop together. It cannot establish that this choice was inevitable, that another bounded learning decision was irrational, or that a present-day framework caused the outcome.

Compare your answer

Review your workbook without giving yourself credit for predicting the eventual launch. Did you define a complete action, keep the contrary evidence, identify each actor, specify normal support, preserve missing data and make a decision that can change?

Use the [facilitator rubric](#) [1] to assess those features. Then write one change to your original decision and name the evidence or reasoning that caused it. An unchanged answer is acceptable if it now states its assumptions and reversal condition more clearly.

Instagram A Worked Decision

Worked recommendation: Test photo sharing for a defined group before committing to a new social network. Treat the missing evidence as work to do.

Complete the [student exercise](#) [10] before reading further. The responses below are original instructor analyses. They are not recommendations recovered from Instagram's team, and choosing the historical winner is not the assessment criterion.

The historical outcome

The team focused on photography. Its 2011 account identifies three product problems: image appearance, upload delay and distribution to other services. A [September 2010 preview](#) [24] confirms the prelaunch photo focus. (evidence BS-0080)

Krieger's [2015 retrospective](#) [25] also describes substantial work to support the service after launch. The history combines behavior choice, product design and execution. It supplies no controlled estimate of the contribution of any one choice.

Response A commit to a narrow prototype

Decision record. Prioritize making a chosen phone photo available to an intended audience. Initially serve people who already take photos they sometimes want to share. Build only enough to observe the whole action in a normal sharing situation. This recommendation authorizes a focused learning step; a wider launch still requires a separate decision.

Reasoning. The reported interest makes this a reasonable candidate to investigate. It does not show how many people have the need, how often it arises, or whether the proposed route improves their existing practice. The largest avoidable commitment would be building a network before learning whether people want another audience there.

Compare the alternatives

This is an instructor comparison. It assigns no historical scores.

Option	Contribution to the proposed objective	Reason to retain or defer it
Share a place	Gives an audience location-based information.	Retain if the intended group's real need concerns places; a general photo objective does not establish that.
Share a photo	Gives an audience visual information selected by the sender.	Investigate first, while testing the demand outside the original network.
Browse and respond	Helps an audience find value in contributions.	Necessary to some product models, but it cannot supply the contributions by itself.
Share through an existing destination	May reduce preparation and transfer work without requiring another network.	Keep as a serious alternative and a comparison for the prototype.

Define the action and its dependencies

Proposed target: when an intended participant independently wants to share a phone photo, they make a selected image available to their intended audience and can confirm that it arrived. The episode ends with that confirmation or with abandonment. This is an instructional definition, not an Instagram metric.

Part of the chain	Responsible party	What to observe
A relevant sharing occasion arises	Sender	A real intention to share; a research instruction alone does not establish demand.
Select or capture the image	Sender	Whether the image and required skills are available.
Decide what and whom to share with	Sender	Audience choice, hesitation and deliberate refusal.
Prepare and transfer the image	Sender and product	Steps, errors, waiting and any supplied help.
Make the result accessible	Product and destination	Delivery or access confirmation, rather than a button press alone.
Encounter another relevant occasion	Sender and their environment	Another opportunity, a different route chosen, or an unknown opportunity state.

This chain keeps the distinction between [capability and context](#) [20] visible. A person can know how to post and still lack a suitable audience or a usable connection. Some people may appropriately choose not to share.

A concrete next test

The following is a **proposed teaching example**. Its participant count and duration are design choices for a small observation exercise, not historical numbers or universal validation thresholds.

Protocol field	Proposed plan
Purpose	Find where intended users choose, complete or abandon this route during normal sharing occasions.
Recruitment	Invite 12 people who already sometimes share phone photos. Seek people beyond the team's friends. Record invitations, declines and eligibility; do not silently replace people who fail.
Exposure and context	Offer a working prototype for seven days. Let participants keep their usual tools. Observe naturally occurring use at home and away from home where those contexts are relevant.
Completion	Confirm that a participant-approved image is available to the chosen audience. Record attempted shares that stop earlier.
Support	Record every reminder, demonstration and troubleshooting step. Separate normal product support from extra research help.

Protocol field	Proposed plan
First-use record	Keep invited, eligible, exposed, attempted and completed counts separate, with the same stated window.
Repeat record	For conditional repeat completion, count first completers who complete again over first completers with a documented later sharing opportunity. Report unknown opportunities and unknown completions separately, with missing-outcome bounds. If also counting later-opportunity completion across all participants, label that as a separate measure, including earlier failures.
Alternative	Record occasions on which the participant chooses their usual route or decides not to share.

Use participant logs to identify possible sharing occasions, then distinguish directly observed events from later reports. A reported opportunity does not become an observed event merely because it appears in a log. Keep that distinction in the repeat-use record.

Rules written before observation: repair a recurring technical obstacle before interpreting abandonment as lack of demand. Reconsider the audience or route if people consistently choose a functioning alternative that better serves the same occasion. Gather more evidence when exposure, opportunity or completion cannot be determined. A promising small observation exercise can justify a larger test; it cannot establish a population-wide completion rate.

The decision after this exercise is whether to invest in a more rigorous field test and what it should compare. Use the [BMF protocol](#) [21] to specify the operational requirements, uncertainty and pass conditions for that larger commitment. No observations or results are invented here.

Response B investigate the existing route first

An equally defensible response can defer the prototype. It would observe the same intended audience using the tools they already have, map the recurring difficulties, and test whether a new service is needed at all. Its strongest reason is the absence of representative usage evidence. Its cost is delaying direct experience with a possible improvement.

This response becomes weak if “more research” has no endpoint. A credible record names the observation that would justify building: for example, a repeated, consequential step that the team can remove while preserving the person's intended audience. It also names a stopping result, such as finding no meaningful improvement the proposed service can offer that population.

Response A accepts more prototype cost to obtain behavioral evidence. Response B accepts more delay to understand existing practice. Compare those commitments and their reversal conditions rather than treating agreement with history as proof of reasoning quality.

Check your response

A strong response does this	Repair this error
Defines an observable action for a specific person and occasion.	“Use Instagram” leaves the behavior and success point unclear.
Keeps posting, audience activity and product features distinct.	“Build filters” is a solution choice without a complete target action.
Names missing evidence and a finding that changes the decision.	Later growth cannot supply an earlier comparison that was never observed.
Includes the existing route and necessary support.	A test that removes competing tools or supplies unrecorded help can mislead.
Makes a limited next commitment.	A favorable reaction to a prototype does not settle sustained use or business viability.

Transfer question, invented scenario: the next product serves employees who may share work images only with a specified internal audience. Which parts of your proposed chain remain useful, and which parts of a public-network strategy would you reconsider? Explain the change using the new condition, not a claim about Instagram's users.

Source chronology and limits

Record	Publication date	How it is used
Product preview [24]	20 September 2010	Contemporaneous observation before public release.
Founders' Stanford talk [26] and transcript [27]	11 May 2011	First-person recollection of the earlier decision.
Krieger's retrospective [25]	7 October 2015	Later account of operating the product.

Sources checked **5 September 2026**. The packet's objectives, candidate comparison, protocol and assessment are instructor work. No recovered source shows the use of DRIVE, BMF or BFA, and the public history does not validate the integrated method.

Original teaching material by Jason Hreha, version 1.0, licensed CC BY 4.0. External sources retain their own terms. See [teaching notes and reuse guidance](#) [1].

Verify A Worked Decision

Worked recommendation: Observe the complete journey in one service before expanding recruitment. Keep a usable route available while identifying where applicants fail to reach the task's end.

Complete the [student exercise](#) [11] first. This is an instructor's proposed response to evidence available **after 5 March 2019**. It does not report a government trial or use later events to prove the recommendation correct.

Repair the briefing

Supported replacement: “The reported 48% concerns Verify sign-up in a single attempt under the source's counting rules. We need separate evidence of completion across the intended service journey.” The NAO's report supplies the limits behind that correction. (evidence BS-0081)

The original briefing makes an additional unsupported move: it assumes improving an intermediate event will improve the final task. Even if both measures later rose, a before-and-after pattern would still need a suitable comparison before attributing the change to a particular intervention.

Historical outcome boundary: this case begins after an audit of disappointing program performance. The audit is the starting evidence. The investigation and routes below are teaching proposals, so there is no historical result to reveal for them.

A completed decision record

Field	Instructor response
Useful outcome	An eligible applicant submits the selected online application and receives confirmation of acceptance for processing, with required identity assurance maintained.
Target action	During an application episode, the applicant supplies the required information, completes the permitted identity route and submits the request.
Population	People beginning their first application to the selected service; record whether they already hold a usable identity account.
Immediate commitment	Investigate the full journey and record the support it requires before increasing recruitment into that route.
Strongest alternative	Start a bounded assisted-route pilot alongside observation if current access problems justify earlier support investment.
Key uncertainty	Which stage prevents useful completion, for which applicants, under which conditions?
Limit on this decision	No claim that a route has achieved BMF, that extra support will work, or that the program has one dominant failure cause.

The selected behavior contains several necessary actions. The purpose of the chain is to expose them, not to assign all responsibility to the applicant.

Define the chain and allocate the work

The following is an original model for the exercise's application service. It is not a reproduction of a government workflow diagram.

Stage	Work required	Observation and possible failure
Enter the service	Applicant identifies the relevant service and starts the application journey.	Record entry before identity-provider selection. Wrong-service visits and unresolved eligibility need explicit handling.
Select a permitted identity route	Applicant and service establish whether a usable account or another route is available.	Record the route offered, chosen or declined. Avoid assuming every entrant needs a new account.
Supply necessary evidence	Applicant gathers and provides the required information; support may assist with permitted steps.	Distinguish missing information, lack of access, inability to complete the step and deliberate refusal.
Complete identity checking	Provider performs the required checks.	Record successful and unsuccessful attempts with reasons where known.
Reach the application	Provider and service pass the result; the service matches it to the relevant records.	Observe a successful handoff rather than assuming that verification guarantees access.
Submit and confirm	Applicant finishes the application; the service confirms acceptance for processing.	Count this as the exercise's completed task. A later service decision is a separate outcome.
Use support or another route	Applicant and staff resolve a problem or continue through an available alternative.	Record completion, time and staff work across routes; do not classify every switch as failure.

This model tests [the whole behavior chain](#) [28]. It also makes an option requiring less applicant work visible: a permitted existing step may be reused, or work may move to the provider. Whether that option is technically and operationally available must be established.

Specify the observations before collecting them

The following plan is **an instructional example**, not a study conducted by Verify. Its timing is a design assumption to adjust to the real service.

Start with a process study. During a two-week recruitment period, invite consenting eligible entrants across the service's normal arrival periods. Include relevant differences in device access, available documentation, prior accounts and need for assistance. Follow each included application for seven days after entry. Record who was invited, who declined and why inclusion was limited where that information is available. This selected observation group cannot establish a population-wide rate.

Link the episode. Use a study reference to connect the same application across its stages and repeat attempts. Do not collect copies of identity documents merely to count progress. Keep application episodes distinct from people and individual attempts.

Record the operating conditions. Keep the service's required identity checks and existing assistance available. Record extra research help separately. Staff effort is part of evaluating an assisted service; supported completion does not establish unaided completion.

Measure	Definition for this proposed study
Confirmed task completion	Included application episodes with an acceptance confirmation within seven days, divided by all included eligible application episodes.
Unresolved outcome share	Included episodes whose final state cannot be established, divided by the same denominator. Report alongside completion.
Stage progress	Episodes reaching each named checkpoint, using both the full cohort and the previous checkpoint as explicitly different denominators.
Repeated attempts	Additional attempts within an existing application episode; report separately from new people or applications.
Supported completion	Confirmed completions with recorded assistance, with the type and amount of assistance visible.
Alternate-route completion	Included episodes completed through another permitted route within the window.
Coverage	Counts invited and included, the contexts represented, and known differences between participants and the intended service population.

No result values are supplied. An unknown final state is not a confirmed failure. The observed completion fraction is incomplete when outcomes remain unresolved; present that limitation alongside the number. Use the [measurement lab](#) [12] to practice the counting rules with fictional records.

Decide what the observations would permit

These rules guide the next investigation. They do not make a small observation study a causal comparison.

Possible finding	Defensible next decision
Events cannot be connected across the journey.	Repair observation first. Do not select a winning route from incomparable denominators.
A recurring obstacle appears before identity checking.	Test a targeted preparation or support change. Treat the proposed cause as a hypothesis and observe the later task as well.
Identity checking succeeds but applicants cannot reach the application.	Prioritize investigation of the service handoff and record matching before increasing recruitment.
An existing route completes the task with acceptable assurance and effort.	Retain it as a serious alternative. More new registrations have no independent value under this objective.
An assisted route appears usable.	Test whether its staffing, availability and cost can be sustained; do not infer unsupported self-service performance.

Possible finding	Defensible next decision
Too few relevant episodes are observed, or outcomes remain uncertain.	Record an inconclusive result and the specific evidence still needed.

Before expanding a proposed change, the service owner must set an acceptable completion requirement, observation window, assurance condition and support-cost limit. A comparative trial would also need an appropriate allocation method, adequate sample and uncertainty analysis. The audit does not supply those local decision requirements. A completed recommendation can therefore authorize the next investigation while explicitly withholding a broader rollout decision.

A defensible alternative

Another response could prioritize an assisted-route pilot immediately, while collecting the same end-to-end observations. This may be reasonable if the service owner gives greater weight to providing current applicants with a workable route than to delaying support costs. That priority is an explicit decision assumption; the packet supplies no local cost-benefit calculation.

The alternative must keep the required assurance, record staff work and preserve observation of the final task. It cannot claim that additional help is automatically effective. A comparison with current practice would still be necessary to estimate the change in completion attributable to the pilot.

Assess the reasoning

A sound response distinguishes the applicant's goal from the registration metric, names all necessary actors, retains unknowns and defines the next decision. It can favor investigation or a bounded pilot if it explains the different commitments.

Revise a response that treats the projection as a control group, labels every departure a motivation problem, excludes early failures to improve a reported percentage, or counts an account as a completed public service. These errors concern what the evidence permits, independent of which route you prefer.

Sources and chronology

The [NAO investigation](#) [29] and its [findings release](#) [30] were published on **5 March 2019**. The detailed measurement boundary is in the [full report, printed page 16, paragraphs 1.16 and 1.17](#) [31]. They describe the same audit, not independent replications. Sources checked **5 September 2026**.

The action chain, study plan, decision rules and alternative response above are instructor analyses. The audit does not show government use of DRIVE, BMF or BFA, and this reconstruction does not validate those methods.

Original teaching material by Jason Hreha, version 1.0, licensed CC BY 4.0. External sources retain their own terms. See [teaching notes and reuse guidance](#) [1].

Measurement Lab Answers

Purpose. Check the calculations and compare the reasoning after completing the lab. The arithmetic has specified answers; several next decisions can be defensible if their assumptions and evidence limits are explicit.

Return to the [student exercise](#) [12] if you have not recorded a response. The public-source sections are instructor analysis. All handover numbers below are invented teaching data, not research findings.

A Wikipedia separate first quality and return measures

The activation analysis counts a first edit to an Article or Article Talk page within 24 hours of registration; constructive activation adds no reversion within 48 hours. Retention involves another edit on a different day in the following two weeks after activation. (evidence BS-0082)

A return rate among activated accounts asks a different question from retained editors per original account. More people can reach the first stage without the conditional return rate improving. Name the population and conditioning rule beside either rate.

The report found improved activation but did not directly detect a retention change. It inferred an increase in retained users through activation. The randomized contrast covers the feature bundle, and exclusion of people who toggled features limits a pure intention-to-treat interpretation. It cannot isolate suggested tasks or establish long-term retention. See the [original report](#) [32].

Defensible next test. Compare a specified task or return-support change while holding the other features constant. Define the intended newcomers, available editing occasions, repeat window, and quality measure before assignment. Retain the original assigned groups for the primary comparison and record actual feature use separately. Longer follow-up and adequate precision would be needed for a longer-term claim. This is a proposed design, not Wikimedia's historical protocol.

B California EITC a treatment result is not a full diagnosis

The research reports no detected increase in filing or claiming from the tested outreach, despite engagement with linked resources. That narrows the claim to these interventions and settings; it does not establish that no workable route to claiming exists. (evidence BS-0083) See the [journal record](#) [33] and [identified working paper](#) [34].

Proposed chain: message delivery, attention, eligibility check, access to information and documents, help if needed, preparation, submission, and recorded claim. The chain is an instructor reconstruction. A website visit measures one intermediate step. An outcome comparison needs the assigned populations and administrative results; restricting it to visitors would select a subgroup that chose to visit.

Possible remaining explanations include difficulty preparing the return, inability to reach usable help, and incomplete delivery to some recipients. A null overall result does not identify which explanation dominates. An offer of assistance should not be coded as received assistance.

Defensible decision. Stop extending the tested outreach on an unsupported expectation of increased claims. If further investment is justified, investigate completion barriers and specify a feasible support change before testing it. "Change support" can describe that next service decision; it is not a claim that more

intensive help has already been shown effective by this study. A reasoned stop on further investment is also defensible under a stated resource limit.

C Fictional data expected counts

The [CSV](#) [35] contains 24 invented team-occasion records from 12 teams. Each team appears twice. Download the [machine-readable answer values](#) [36] for checking calculations.

Count	Expected value
Eligible roster records	24
Unique fictional teams	12
Known applicable opportunities	19
Known occasions without an opportunity	3
Unknown opportunities	2
Attempts on known opportunities	16
Known completions on known opportunities	11
Known noncompletions on known opportunities	7
Unknown completion on a known opportunity	1

The two unknown opportunities belong to T12. Their attempt and completion states are also unknown. Keep them visible outside the opportunity-specific calculations. T05's second record has a known opportunity and attempt but an unknown completion.

Completion calculations

Percentages are rounded to two decimal places. All figures in this section come from the invented file.

Question	Calculation	Result
Attempts per known opportunity	16 / 19	84.21%
Completion among known outcomes on known opportunities	11 / 18	61.11%
Completion on known opportunities, unresolved completion fails	11 / 19	57.89%
Completion on known opportunities, unresolved completion succeeds	12 / 19	63.16%
Completion among attempts with known outcomes	11 / 15	73.33%
Conditional completion, unresolved attempt fails	11 / 16	68.75%
Conditional completion, unresolved attempt succeeds	12 / 16	75.00%

Question	Calculation	Result
Observed completions per scheduled roster record	11 / 24	45.83%

The 73.33% figure omits opportunities with no attempt and the unresolved attempted outcome. It answers a conditional question. It is not the proportion of opportunities successfully completed.

The 45.83% figure uses a legitimate but different denominator: every scheduled roster record. It includes known occasions without a task to do and unknown opportunities. Do not label it completion per applicable opportunity.

The bounds account only for the one unresolved completion among known opportunities. They do not resolve T12's unknown opportunities, sampling uncertainty, repeated-team dependence, or selection. No inferential confidence interval or causal estimate is supplied by this teaching file.

Repeat completion

First completers are T01 through T06. Five have a known second opportunity; T06 has no unfinished work on its second occasion and is excluded from this particular denominator. Of those five, T01, T02, and T03 complete again; T04 does not; T05 is unresolved.

Repeat measure	Calculation	Result
Repeat among known second outcomes	3 / 4	75.00%
Lower bound among known second opportunities	3 / 5	60.00%
Upper bound among known second opportunities	4 / 5	80.00%

Counting T06 as a failed repeat would confuse absence of a task with failure to do it. Counting every team as a first completer would answer a different question. These are two occasions, so no conclusion about long-term repetition follows.

Support and recorded effort

Support condition	Known opportunities	Completions	Unresolved completion	Recorded staff minutes
Normal	14	7	0	25
Exceptional	5	4	1	63

Normal-support completion is 7/14, or **50.00%**. Exceptional-support completion is bounded by 4/5 and 5/5, or **80.00% to 100.00%**. Excluding the unresolved exceptional record produces 4/4, which must be identified as a known-outcome subset.

Support was not randomly assigned. The table cannot show that coaching caused the difference. It also cannot show that exceptional help can be supplied at normal operating capacity.

Recorded staff effort totals **88 minutes**, with two unknown staff-time records. Include effort on failures and on T05's unresolved attempt. The recorded total already exceeds the illustrative 60-minute budget by **28 minutes**. Unknown minutes can increase that total. The ratio of recorded staff effort to known completions is 88/11, or **8 minutes per known completion**. It is not the average coaching duration of successful teams or the service's full cost.

Worked next decision

Change support, with a limited diagnostic step before another trial. The normal-support rate is below the invented 70% operating requirement, and recorded help already exceeds the budget. The file does not support expansion, a claim of BMF, or a causal conclusion about coaching.

Inspect the handoff at acknowledgement and the unresolved task-owner step. Compare the practical requirements of a scheduled handover overlap, a clearer ownership rule, and the current written route. Repair the missing records where possible. Estimate normal staffing and queueing for the proposed change, then agree a new test and decision rule before seeing its results.

This recommendation is instructor analysis. A response that stops the route because no affordable repair fits the stated constraints can also be defensible. Reselection needs evidence favoring an alternative action. "Inconclusive" is appropriate for the precise repeat rate and coaching effect; those uncertainties do not erase the observed normal-support and recorded-budget failures.

Original answer guide and fictional calculations: Jason Hreha, version 1, 2026-09-05, CC BY 4.0. Linked sources retain their own rights. No real participant study is reported.

References and online companions

Numbered links identify sources or companion pages used in the text. External material retains its own rights. These links do not imply endorsement. The pack contains original teaching synthesis, not copies of third-party casebooks or proprietary exhibits.

M PESA source use

The original teaching material follows the [teaching-pack reuse terms](#) [1]. The [student source chronology](#) [37] identifies the pilot sources. Later company and study sources above are outcome evidence. External publications retain their own rights. Sources checked 5 September 2026.

1. teaching notes

<https://behavioralstrategy.com/education/teaching-notes/>

2. complete pack

<https://behavioralstrategy.com/downloads/teaching-pack/v1/teaching-pack.zip>

3. individual files

<https://behavioralstrategy.com/downloads/#public-teaching-pack>

4. six-part rubric

<https://behavioralstrategy.com/education/teaching-notes/#discussion-rubric>

5. host observation record

<https://behavioralstrategy.com/education/teaching-notes/#host-observation-record>

6. CC BY 4.0

<https://behavioralstrategy.com/education/teaching-notes/#reuse-and-adaptation>

7. workshop

<https://behavioralstrategy.com/education/workshop/>

8. solo route

<https://behavioralstrategy.com/education/solo-practice/>

9. M-PESA packet

<https://behavioralstrategy.com/education/mpesa-decision/>

10. Instagram packet

<https://behavioralstrategy.com/education/instagram-decision/>

11. Verify packet

<https://behavioralstrategy.com/education/verify-decision/>

12. Measurement lab

<https://behavioralstrategy.com/education/measurement-lab/>

13. evidence ledger

<https://behavioralstrategy.com/evidence/ledger/>

14. editable text record

<https://behavioralstrategy.com/downloads/teaching-pack/v1/host-observation.txt>

15. corrections page

<https://behavioralstrategy.com/corrections/>

16. CC BY 4.0

<https://creativecommons.org/licenses/by/4.0/>

17. license map

<https://behavioralstrategy.com/licensing/>

18. manifest

<https://behavioralstrategy.com/downloads/teaching-pack/v1/manifest.json>

19. Behavior Market Fit

<https://behavioralstrategy.com/glossary/behavior-market-fit/>

20. Behavior Fit Assessment

<https://behavioralstrategy.com/glossary/behavior-fit-assessment/>

21. test method

<https://behavioralstrategy.com/methodology/validate-behavior-market-fit/>

22. Read Joseph's account

<https://www.vodafone.com/news/newsroom/digital-society/m-pesa-created>

23. MIT's account of the study

<https://news.mit.edu/2016/mobile-money-kenyans-out-poverty-1208>

24. September 2010 preview

<https://techcrunch.com/2010/09/20/instagram/>

25. 2015 retrospective

<https://www.wired.com/2015/10/five-years-of-building-instagram/>

26. Founders' Stanford talk

<https://stvp.stanford.edu/av/stanford-startup-entire-talk>

27. transcript

<https://stvp.stanford.edu/node/12536/printable/print>

28. the whole behavior chain

<https://behavioralstrategy.com/education/decision-workbook/#step-3>

29. NAO investigation

<https://www.nao.org.uk/reports/investigation-into-verify/>

30. findings release

<https://www.nao.org.uk/press-releases/investigation-into-verify/>

31. full report, printed page 16, paragraphs 1.16 and 1.17

<https://www.nao.org.uk/wp-content/uploads/2019/03/Investigation-into-verify.pdf#page=18>

32. original report

https://www.mediawiki.org/wiki/Growth/Personalized_first_day/Newcomer_tasks/Experiment_analysis_November_2020

33. journal record

<https://www.aeaweb.org/articles?id=10.1257/pol.20200603>

34. identified working paper

<https://www.capolicylab.org/wp-content/uploads/2020/11/Increasing-Take-Up-of-the-Earned-Income-Tax-Credit.pdf>

35. CSV

<https://behavioralstrategy.com/downloads/teaching-pack/v1/measurement-records.csv>

36. machine-readable answer values

<https://behavioralstrategy.com/downloads/teaching-pack/v1/measurement-answers.json>

37. student source chronology

<https://behavioralstrategy.com/education/mpesa-decision/#source-chronology-and-reading-boundary>

Attribution and reuse

Original teaching content, exercises, and invented data: Jason Hreha, Behavioral Strategy, version 1, September 5 2026. Licensed CC BY 4.0 for the original material. Linked source text, marks, and exhibits retain their own rights. Preserve attribution, identify changes, and do not imply endorsement. No general site license or third-party rights are changed by this pack.