

After the Assistant Gets It Wrong: Earning Permission Again

An executive guide to recovery, renewed authorization and continued reliance in consumer AI

Consumer AI Adoption Gap · Research brief 08
October 5, 2026

© 2026 Eldris Inc. All rights reserved.

Executive brief

An AI assistant orders the wrong item. The charge is within the spending limit, but the size is incorrect. The product apologizes, initiates a return and offers to try again. Has it repaired the transaction? Has it established that the next purchase should be delegated? Those are separate questions.

Repair restores a situation. Renewed permission authorizes a future action. A successful refund does not establish that the assistant can now select the right item. A reassuring explanation does not establish that the refund has completed. A user may accept the remedy while choosing a narrower role for the assistant.

This report is for consumer AI product leaders, commerce platforms and research teams designing what happens after a failed delegated action. Its bounded decision concerns a routine purchase: how should an assistant recover from buying the wrong product variant, and what permission should it request for the next purchase?

The central recommendation is to evaluate recovery on two tracks: **what was repaired and what the user subsequently authorized**. Record the remaining effort, cost and uncertainty as well as the permission granted, narrowed or withdrawn. Retention and satisfaction alone cannot establish an effective recovery design.

Three principles follow:

- Diagnose the failure. An incorrect recommendation, an unauthorized action, a failed execution and a poor remedy require different responses.
- Show verified state. Distinguish a requested cancellation from a canceled order, and a promised refund from a received refund.
- Make the next permission a real choice. Allow a narrower scope or independent completion without treating either as a failure of the user.

The objective is justified continued reliance. It may mean renewed delegation, approval before purchase or recommendation only. Sometimes the appropriate outcome is to stop using the assistant for the task.

The failure changes the user's decision

Before a mistake, a consumer evaluates the expected convenience and consequences of delegation. After a mistake, that person also has evidence about how the system performs, how the provider responds and how much work a failure creates.

The wrong-variant example is constructed, not an observed company case. Assume the user specified a product and size, authorized a purchase within a defined budget and received a different size. That incident concerns execution against the authorized request. If the assistant selected an unapproved substitute, authority may also be implicated. If the return fails, recovery becomes a further issue.

A single general trust question compresses those judgments. The user could believe that the assistant finds useful products while refusing to let it purchase them. That distinction is commercially useful: the product may retain a valuable advisory role even when transaction authority is withdrawn.

The relevant next question is: **“What would you let this assistant do now, given what happened and what has changed?”** The answer needs a task, an action and a meaningful alternative.

Decision implication: Measure the permission boundary before and after the incident. Avoid interpreting a return to the application as restored purchasing authority.

Control is a research lead, not a universal remedy

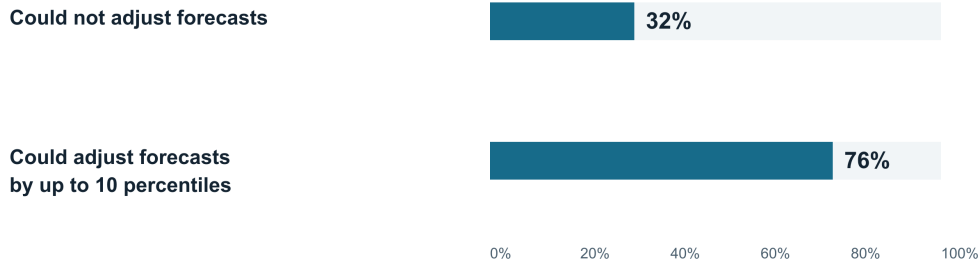
In Study 1 of Dietvorst, Simmons and Massey's forecasting research, 32% of participants in a condition without adjustment chose the model; 76% did so when they could adjust its forecasts by up to 10 percentiles.

Participants knew the model was imperfect. This was a forecasting experiment, not an agentic purchase or recovery study. [Published study](#).

PUBLISHED FORECASTING EXPERIMENT / SELECTED CONDITIONS

Limited control changed model choice

Share choosing the model · Participants knew it was imperfect



A research lead for control design, not a prediction about AI shopping.

Study 1: 288 completers across four conditions. Selected rounded point estimates.

Source: Dietvorst, Simmons and Massey, Management Science, 2018, Figure 2.

© Eldris Inc

Figure 1. Selected conditions from Study 1; 288 participants completed forecasting across four conditions. Published rounded percentages; uncertainty bars are not reproduced. The result motivates testing control, not forecasting commerce behavior.

The distinction matters. Adjusting a forecast affects an answer before use. Canceling a purchase depends on whether an external action can still be stopped. Approving a checkout can prevent some mistakes, but only if the relevant details are visible and the person can assess them. A control that feels reassuring may fail to address the actual cause.

For the wrong-size purchase, a spending cap would limit the charge but would not establish correct variant selection. Confirmation showing the exact item and size could provide an opportunity to catch a mismatch. Whether people notice it, and whether the additional effort is worthwhile, must be tested.

Decision implication: Match the control to the failure mechanism. Use public research to select plausible designs, then measure their effects in the actual delegated task.

Classify the failure before choosing the remedy

The framework below is original Eldris analysis. Its categories can overlap within an incident. They separate the decisions a product team must make rather than assigning a universal severity score.

ELDRIS FAILURE AND RECOVERY DESIGN MAP

Different failures require different evidence

Original analytical guidance · Categories may overlap within one incident

FAILURE QUESTION	IMMEDIATE RESPONSE	UNRESOLVED EVIDENCE
Wrong recommendation?	Correct the information; identify consequences.	Can the assistant apply the relevant preferences?
Exceeded permission?	Stop or narrow the affected action; verify boundaries.	Does the authorization limit now work?
Failed execution?	Verify external state before retrying or repairing.	Can the authorized action be completed reliably?
Failed recovery?	Track pending work; provide accountable support.	Has the remedy actually completed?
A remedy can resolve one part of an incident while leaving another open.		
Not a universal severity score or a validated predictor of renewed permission.		

© Eldris Inc

Figure 2. Four failure questions and the evidence a remedy should address. A repair can resolve one part of an incident while leaving another unresolved.

Was the recommendation wrong?

The assistant proposed an unsuitable item or gave incorrect information, but the user retained control of the purchase. Correct the recommendation and show the relevant basis for the correction. Investigate whether the issue concerns missing preferences, incorrect product data or poor interpretation.

The next permission might remain advisory. There is no completed transaction to undo, but the user may already have spent effort checking or acted on the information elsewhere. Identify the consequence rather than assuming it was harmless.

Did the assistant exceed permission?

The assistant bought an unapproved substitute, exceeded a limit or acted after permission was withdrawn. A transaction remedy does not resolve the permission problem. Determine what authority was granted, what the system enforced and why the action occurred.

A stronger apology cannot substitute for evidence that the relevant boundary now works. Stop or narrow the affected action while that question remains unresolved. A product should not request the same broad delegation as though the incident concerned only an unfortunate outcome.

Did execution fail within the task?

The assistant attempted the authorized purchase but submitted the wrong variant, created a duplicate or reported completion when an order had not completed. Verify the actual external state before retrying. An uncertain checkout outcome can make a second attempt create a second order.

The remedy depends on what exists: an order to cancel, an item to return or an incomplete action to finish. Product-side activity logs should be reconciled with merchant or payment confirmations where relevant. An assistant's own statement of success is insufficient verification.

Did recovery create a second failure?

The refund remains pending, the consumer must repeat information or the assistant cannot reach the merchant. Track these as recovery outcomes rather than hiding them inside the original incident.

Give the user a clear route to a responsible support function when the assistant cannot resolve the case. Record outstanding work and ownership. A case is not complete merely because the conversation has ended.

NIST's Generative AI Profile recommends documented incident handling, tracking errors and assessing whether recovery procedures were followed and effective. This provides an operational foundation; it is not evidence about which consumer remedy restores permission. [NIST guidance](#).

Decision implication: Identify both the initial failure and any recovery failure. Specify which mechanism has been addressed and which remains uncertain.

Build recovery around observable state

The recovery experience should make five facts understandable: what the assistant did, how it differed from the request, what has been confirmed, what remains pending and what the user needs to do.

For the wrong-size purchase, the product might confirm the submitted variant and the merchant's cancellation response. If cancellation is unavailable, it should identify the return process and remaining steps. Avoid presenting an initiated remedy as a completed remedy.

Separate immediate containment from restoration. Pausing future purchases can contain exposure. It does not recover a charge or eliminate the work of returning an item. Measure the user's actual burden: time, repeated contacts, out-of-pocket cost and unresolved uncertainty.

Explain what is known about the cause and what has changed. Where the cause is still under investigation, say so. A plausible narrative generated by the assistant is not a verified diagnosis. A correction should be supported by relevant evidence, such as a fixed variant mapping or a tested authorization check.

Clear language and an apology may be appropriate, but neither establishes technical repair. Human-AI interaction research recommends supporting correction and making system behavior and limitations understandable. Those principles inform recovery design without proving that a particular message restores trust. [Interaction guidelines](#).

Decision implication: Define completion using the relevant external outcome. Measure the burden left with the consumer, not just the speed of the first response.

Repair and permission need separate records

A consumer can accept a refund and refuse another automated purchase. A consumer can also approve another purchase before a refund completes. Those combinations show why the two outcomes should remain visible.

ELDRIS RECOVERY AND AUTHORIZATION FRAMEWORK

Repair and permission need separate records

Original conceptual guidance · The same incident can produce different outcomes

REPAIR RECORD	NEXT PERMISSION RECORD
What happened?	Recommendations only
What external state is confirmed?	Cart prepared for approval
What has been repaired?	Purchase within defined limits
What remains pending?	Independent completion
What burden remains with the user?	Permission withdrawn

A completed refund does not automatically restore purchasing authority.

Record actual state and the chosen scope. Revocation may not reverse completed actions.

Permission alternatives are illustrative and must reflect genuine product capabilities.

© Eldris Inc

Figure 3. Two parallel records for the same incident. The choices are illustrative; their availability must reflect the product's genuine capabilities. Repair completion does not automatically restore prior authority.

For the next purchase, offer a small set of legible alternatives where the product supports them: recommendations only, preparation of a cart for approval, or purchase within specified limits. Include a way to complete the task independently. Record the scope chosen, not only whether the person pressed a continue button.

The request should describe the actual action. A generic invitation to “try again” could refer to searching, preparing a cart or completing checkout. That ambiguity makes the permission difficult to interpret and the measurement less useful.

Revoking future permission also needs a defined effect. Show which scheduled or in-progress actions are stopped, which have already completed and which require separate cancellation. Revocation is an authority change; it cannot guarantee reversal of every external consequence.

Decision implication: Keep transaction status and authorization status separate. Do not silently restore a revoked permission because the remedy succeeded or the user returned.

Do not optimize reassurance at the expense of judgment

The strongest-looking recovery screen may increase willingness to continue without improving reliability. That is an important product risk and a measurement problem.

Microsoft's research on appropriate reliance recommends helping people understand capabilities, identify mistakes and verify outputs. It also reports that apparent verification aids can backfire in some circumstances. Its findings concern generative AI experiences, including retrieval-based applications; they do not establish a recovery effect in commerce. [Appropriate reliance research](#).

For the purchase workflow, evaluate what users understood: the item bought, the remedy's status and the permission they are granting next. Test whether they can identify a mismatch in a later purchase. Renewed permission with mistaken understanding is different from informed continuation.

The opposite error is to treat every withdrawal as product failure. Declining an action that remains unreliable can be an appropriate response. Users may also retain the assistant for lower-consequence tasks. Recovery should preserve that option when feasible.

Repeated failures require their own evaluation. Evidence from one recoverable wrong-variant purchase cannot establish tolerance of repeated mistakes, privacy failures or actions outside permission. The consequence, recurrence and repair burden all change the decision.

Decision implication: Evaluate informed authorization and successful subsequent outcomes together. A higher willingness score is insufficient if the design also increases inappropriate reliance.

Test the next permission decision

A useful study begins with one task, one defined failure and a small number of recovery alternatives. Decide whether the study is testing disclosure, repair service or renewed-permission controls. Changing all three together can compare complete packages, but it cannot isolate which component caused a difference.

For the wrong-variant purchase, an initial comparison could hold the incident, repair outcome and demonstrated purchasing capability constant while varying the next authorization design:

- Restore the same purchasing scope only after a clear renewed authorization.
- Offer a cart requiring confirmation before checkout.
- Let the user select a narrower scope from a short set of alternatives.

These are proposed research conditions, not tested recommendations. Every condition should clearly communicate the failure and verified remedy; the comparison should not rely on concealing material information. Follow the permission choice with a consequential task decision where appropriate, rather than stopping at a stated intention.

Measure at least four outcomes: understanding of the incident and remedy, authority granted or withdrawn, completion of the next task and the burden placed on the user. Record correction and verification behavior where it is observable. Select the primary decision measure in advance and size the study around a meaningful difference.

Analyze all participants assigned to each recovery design. Examining only those who return would exclude people most affected by the failure. Follow up on later relevant occasions, with missing responses and departure reported. A controlled task can reveal immediate choices; continuing reliance requires evidence beyond the first session.

A research panel can provide evidence about consumer responses when recruitment and tasks match the intended population. It cannot verify a merchant's operational reliability or establish field retention on its own. Pair behavioral research with evidence that the actual repair, permission and revocation mechanisms work.

Decision implication: Choose the study around the design decision the team can change. Keep behavioral response and operational reliability as complementary forms of evidence.

What to require before asking the user to delegate again

Prepare a short decision record for each material failure category:

1. **Incident:** the authorized action, actual outcome and verified external state.
2. **Remedy:** completed repair, pending steps, costs and burden remaining with the user.
3. **Cause and change:** established findings, unresolved questions and evidence for the proposed correction.
4. **Next permission:** the action requested, limits, alternatives and effect of revocation.
5. **Consumer evidence:** understanding, permission choices, subsequent outcomes and observed recurrence.
6. **Release decision:** what can resume, what remains restricted and what will trigger reevaluation.

Renewed delegation is supportable when the operating evidence and consumer evidence justify the defined scope. Narrower permission may preserve useful assistance while limiting unresolved exposure. Where neither repair nor reliability is established, postpone the request for renewed delegation and resolve the outstanding condition.

For product leaders, the goal is not to make the failure disappear from the conversation. It is to demonstrate what has been repaired, what can now be relied on and what the person can choose. **The next authorization is a new decision, not an automatic return to the previous state.**

Method and limitations

This brief selectively reviews primary sources checked on October 5, 2026. It is not an original consumer study, systematic review or estimate of recovery performance. The forecasting experiment concerns a statistical model and a laboratory task; its result cannot be transferred directly to purchase delegation after an incident.

NIST supplies operational guidance. Interaction and appropriate-reliance research inform questions to examine; they do not identify the best consumer recovery package. The failure map, purchase example, parallel records and study proposal are original Eldris analysis. No actual purchase, refund, company incident or new participant result is claimed.

The report supports product planning and evaluation. Determining which recovery design changes informed permission requires task-specific research. Establishing sustained reliance and reliable remedies additionally requires follow-up and operational evidence. Recovery may appropriately end with reduced or withdrawn authority.

Sources

1. **Dietvorst, B. J., Simmons, J. P., and Massey, C. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them.** *Management Science*, 64(3), 1155-1170, March 2018; first published online November 4, 2016. Study 1 and Figure 2 supply the selected model-choice percentages. [Published paper](#).
2. **NIST. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1.** July 2024. MANAGE 4.3 and related actions address incident communication, recovery assessment and tracking errors. [Profile](#).
3. **Amershi and colleagues. Guidelines for Human-AI Interaction.** CHI 2019. Guidelines address correction, capability communication and user control. [Publication and paper](#).
4. **Vorvoreanu, M., Passi, S., Dhanorkar, S., Heger, A., and Walker, K. Fostering Appropriate Reliance on GenAI: Lessons Learned from Early Research.** Microsoft Technical Report MSR-TR-2025-4, 2025. Research synthesis and guidance on user understanding, verification and overreliance. [Technical report](#).