FROM THE RESEARCH BENCH

Algorithm trust changes with the task

Algorithm trust is task-dependent: review objectivity, observed error, advice source, expertise, and reliance before calling a system trusted.

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher

PUBLISHED
1 September 2026

READING TIME
9 min

WORDS
1,888

https://isoglu.com/insights/journal/algorithm-trust-changes-with-the-task/

CONTENTS

MANAGEMENT SUMMARY

People can appreciate algorithmic advice in one setting, resist it after seeing a visible error, and rely on it differently when a task feels objective or subjective. Castelo, Bos, and Lehmann show that perceived task objectivity changes algorithm reliance across laboratory and field studies. Dietvorst, Simmons, and Massey show that observing an algorithm err can reduce confidence and future reliance even when the algorithm has outperformed a person. Logg, Minson, and Moore document algorithm appreciation, while identifying limits when people compare an algorithm with their own estimate or have forecasting expertise. This article reconciles aversion and appreciation as boundary conditions and builds a trust-by-task review card. It does not evaluate a current vendor, model, employee, or deployment.

Keywords: Algorithm trust · Algorithm aversion · Algorithm appreciation · Decision support

Algorithm trust is not one stable attitude. It changes with the task, the observed error, and the person who is making the decision.

The short answer is that a system should not be described as trusted or distrusted without naming the decision boundary. People may use algorithmic advice more for an objective task, withdraw from an algorithm after seeing it make a mistake, and still prefer algorithmic advice to human advice in another setting. Those findings are compatible once task objectivity, experience, comparison point, and expertise are treated as conditions rather than averaged into a single trust score.

Castelo, Bos, and Lehmann show that perceived task objectivity changes reliance on algorithms. Dietvorst, Simmons, and Massey show a specific aversion response after people observe an algorithm err. Logg, Minson, and Moore show algorithm appreciation in several estimation and forecasting settings, with weaker appreciation when people compare the algorithm with their own estimate or bring forecasting expertise. Together, the sources support a task-and-experience map, not a universal rule for human reliance.

Why must algorithmic evaluation begin with task structure rather than system capabilities?

Imagine the same decision-support interface being used for two questions. The first asks for a numerical forecast with a defined outcome and a historical benchmark. The second asks for a judgment about a person's taste, fit, or emotional response. The interface has not changed. The decision context has.

Before asking whether the user trusts the algorithm, name the fields below:

Table 1 Why must algorithmic evaluation begin with task structure rather than system capabilities?

Field	Question	Why it matters
Task character	Is the task perceived as objective, subjective, or mixed?	Perceived objectivity changes willingness to use algorithmic advice.
Advice source	Is the advice labelled as algorithmic, human, or unknown?	The source label can change adherence independently of the numerical advice.
Observed performance	Has the decision-maker seen the system make an error?	A visible mistake can change reliance, even when comparative performance remains stronger.
Comparison point	Is the advice compared with a person's estimate or with no alternative?	Appreciation can weaken when the user's own judgment is the direct competitor.
Expertise	Does the decision-maker have forecasting or domain expertise?	Expertise can change how strongly a person follows algorithmic advice.
Consequence	What happens if the advice is followed or ignored?	Reliance is a decision, not merely an attitude response.

Table from this essay. Sources and interpretation are given in the article.

The fields prevent a common category error. A user can trust an algorithm for a numeric forecast and still reject it for a subjective recommendation. A user can appreciate a labelled algorithm in a controlled experiment and re-

duce reliance after a salient error in practice. Neither observation proves that the user's general attitude has changed across all tasks.

Figure 1 The algorithm trust-by-task map

Review field	Example boundary	Permitted interpretation	Stop signal
Task character	Objective, subjective, or mixed	"Reliance is being reviewed for this task."	The article generalizes from one task to all tasks.
Observed performance	No observed error, or a visible error	"Experience with this performance may affect reliance."	One error is treated as proof of overall inferiority.
Advice source	Algorithmic, human, or unknown label	"The source label is part of the decision context."	Source labels are omitted while attitude is compared.
Expertise and comparison	Own estimate, human benchmark, or no direct rival	"The reliance response is bounded by the comparison and expertise."	A novice and an expert are treated as the same decision-maker.
Release question	What reliance decision is being made?	"Use, override, review, or test the advice at this boundary."	"Users trust the algorithm" is the final finding.

Author's decision framework grounded in Castelo, Bos and Lehmann (2019), Dietvorst, Simmons and Massey (2015), and Logg, Minson and Moore (2019). The prompts are synthetic and do not score a current system.

How does perceived task objectivity alter human baseline trust in algorithms?

Castelo, Bos, and Lehmann report that algorithm reliance varied with perceived task objectivity across four online laboratory studies with more than 1,400 participants and two online field studies with more than 56,000 participants. Consumers trusted and used algorithms less for tasks perceived as subjective than for tasks perceived as objective. The result is a boundary condition on reliance, not a claim that all subjective tasks reject algorithms or all objective tasks accept them.

The authors also report that increasing perceived task objectivity or increasing the algorithm's affective human-likeness increased trust and use for otherwise subjective tasks. This matters because teams often describe a task as objective based on the data format while users experience it as interpretive. A rating, ranking, or forecast may look numeric and still involve assumptions about taste, fit, fairness, or identity. The decision interface does not remove that perception.

The practical implication is to measure or discuss perceived task character before interpreting a reliance rate. If a team changes the explanation, interface, or social framing of a system, it may also change how the task is perceived. That does not mean the algorithm has become more accurate. It means the conditions of use have changed.

How do observable algorithmic errors trigger disproportionate algorithm aversion?

Dietvorst, Simmons, and Massey report algorithm aversion across five studies after participants observed algorithmic forecasters perform. Participants lost confidence in an algorithm more quickly than in a human after the two made the same mistake, even when the algorithm had outperformed the human. Seeing the algorithm perform also reduced the likelihood that participants would tie future incentives to it rather than to an inferior human forecaster.

This finding identifies a particular experience: the decision-maker sees the system make an error. It does not show that users should never rely on algorithms, and it does not show that every error has the same effect. Error visibility, error magnitude, task stakes, and the available comparison all belong in the review record.

The distinction matters for operational design. A team may want a system to be evaluated against a human benchmark, monitored for calibration, and reviewed after errors. Those controls are not evidence that the system is trusted. They are ways to make reliance and correction observable. A release note should say whether the system is being used, reviewed, overridden, or tested after an error.

The strongest unsupported sentence is “users abandon algorithms when they err.” The permitted sentence is narrower: in the reported studies, observing algorithmic errors reduced confidence and reliance more sharply than observing comparable human errors, even when the algorithm had performed better. That sentence preserves the source boundary and leaves room for the task and comparison conditions.

Under what specific operational conditions do users exhibit algorithm appreciation?

Logg, Minson, and Moore report algorithm appreciation across six experiments. Lay participants adhered more to advice when they believed it came from an algorithm than when they believed it came from a person. The reported settings include numeric estimates and forecasts, including judgments about song popularity and romantic attraction.

The result can sound like the opposite of algorithm aversion until the comparison is made explicit. In the appreciation studies, the question is often whether people follow advice from a labelled source. In the aversion studies, the question is how people react after observing an algorithm make a mistake. A source label, an observed performance record, and a task framing are different treatments.

Logg and colleagues also report limits. Algorithm appreciation weakened when people chose between the algorithm’s estimate and their own estimate, or when participants had forecasting expertise. The direct competitor may be the user’s own judgment rather than a generic human forecaster. That changes the decision. A person can prefer an algorithm to an unnamed human average and still override it when the algorithm conflicts with a well-supported expert estimate.

The review should therefore record the comparison point. “Algorithmic advice was preferred” is incomplete without saying preferred to what, under which task, with which expertise, and after which experience.

Why are algorithm aversion and appreciation context-dependent behaviors?

The three sources can be assembled into a simple sequence:

1. The task is framed as objective or subjective.
2. The advice source is labelled or left ambiguous.
3. The decision-maker observes performance, including any visible error.
4. The advice is compared with a person, the user's own estimate, or no direct alternative.
5. Expertise and stakes shape whether the person follows, checks, or overrides the advice.

This sequence explains why a single trust survey is a weak release criterion. It compresses task, history, comparison, and expertise into one number. A high score can coexist with low reliance in a high-stakes task. A low score can coexist with useful algorithmic input when the user is required to review and correct it.

The map also provides a better test for a new decision-support process. Ask whether the intended user should follow the advice automatically, review it before acting, use it as one input among several, or treat it as a benchmark. Those are different reliance policies. They should not be hidden behind the word trust.

What phased governance sequence should govern automated decision system rollout?

Use this sequence when deciding whether evidence supports a claim about reliance:

1. Define the task and record whether users experience it as objective, subjective, or mixed.
2. State what the algorithm is being compared with: a human forecast, a user's own estimate, or no direct rival.
3. Record whether users have observed an error and how the error was presented.
4. Separate advice adherence from accuracy, calibration, and outcome quality.
5. Describe relevant user expertise rather than assuming that all users respond alike.
6. Decide whether the policy is automatic use, human review, override, or controlled testing.
7. Write the strongest supported sentence and the stronger universal sentence that remains outside the evidence.

This is not a recipe for manufacturing trust. It is a way to keep the decision object visible. A team can choose a review-first policy even when algorithm appreciation is high. It can also test a useful algorithm in a subjective setting without claiming that the task has become objective.

Where are the empirical boundaries of human-algorithm collaboration research?

The three sources do not evaluate a current vendor, model, employee, or organizational deployment. They do not establish one universal algorithm-trust trait, a guaranteed response to every error, or a general rule that algorithms should replace human judgment. They show that perceived task objectivity, observed errors, source labels, comparison points, and expertise can condition reliance in the reported settings.

The stopping rule is concrete. Do not release “users trust the algorithm” until the task, observed performance, comparison point, expertise, and reliance policy are named. If one is missing, the system may still be accurate or useful. The trust claim is not yet bounded enough to publish.

The task boundary connects to hybrid intelligence and its capability boundary and the customer model that cannot predict every outcome.

Castelo, N., M. W. Bos, and D. R. Lehmann. (2019). Task-Dependent Algorithm Aversion. Journal of Marketing Research, 56(5), 809-825. DOI

Dietvorst, B. J., J. P. Simmons, and C. Massey. (2015). Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. Journal of Experimental Psychology: General, 144(1), 114-126. DOI

Logg, J. M., J. A. Minson, and D. A. Moore. (2019). Algorithm Appreciation: People Prefer Algorithmic to Human Judgment. Organizational Behavior and Human Decision Processes, 151, 90-103. DOI

END OF ESSAY

HOW TO CITE THIS ESSAY

Isoglu, S. (2026, September 1). Algorithm trust changes with the task. Sinan Isoglu. https://isoglu.com/insights/journal/algorithm-trust-changes-with-the-task/

KEEP READING

From the research bench

What is data lineage? The transformations behind a result

https://isoglu.com/insights/journal/what-is-data-lineage/

From the research bench

What is interference? When one unit changes another unit's outcome

https://isoglu.com/insights/journal/what-is-interference/

From the research bench

What is measurement error? The gap between a construct and its indicator

https://isoglu.com/insights/journal/what-is-measurement-error/

Sinan Isoglu, MBA (Quantic)
Commercial growth leader, lecturer and doctoral researcher