Do Individuals Treat Their Posterior Beliefs as Sufficient Statistics? An Experimental Investigation

Zhenlin Kang[*] Marina Agranov[†] Kirby Nielsen[‡] Kim Sarnoff[§]

May 21, 2026

Abstract

We experimentally investigate whether individuals view their posterior beliefs as sufficient statistics. After facing a series of updating tasks, we ask individuals whether they would prefer to rely on their posterior belief or would prefer to revisit past signals when receiving new information. A substantial majority (68%) prefer to revisit past signals. Secondary evidence suggests that this perceived posterior insufficiency results from lack of confidence in the posterior belief: Individuals have a stronger preference to revisit signals when they were less confident in their posterior, and a majority of individuals would rely on the Bayesian posterior instead of revisiting signals.

[*] zkang@caltech.edu; Division of the Humanities and Social Sciences, California Institute of Technology

[†] marina.agranov@gmail.com; Division of the Humanities and Social Sciences, California Institute of Technology and NBER

[‡] kirby@caltech.edu; Division of the Humanities and Social Sciences, California Institute of Technology

[§] ksarnoff@princeton.edu; Department of Economics, Princeton University

For helpful comments and suggestions, we thank Aislinn Bohren, Colin Camerer, Luciano Pomatto, and Charlie Sprenger. This project was reviewed and granted an exemption by the Institutional Review Board at the California Institute of Technology under protocol number IR25-1572.

1 Introduction

A fundamental prediction of Bayesian—and many non-Bayesian—models of reasoning is that posterior beliefs are "sufficient statistics" of all prior information: Once a decision-maker (DM) has received some information and updated their belief, they do not need to revisit that information in the future since all information content is contained in the updated posterior. We call this phenomenon "posterior sufficiency". Recent work shows that posterior sufficiency often fails in practice: Individuals reach quite different final beliefs when sufficiency is enforced compared to when it is not (Sarnoff, 2025). However, existing work has neglected the question of whether decision-makers *perceive* their posterior beliefs as sufficient for future updating. In other words, when receiving new information, do individuals prefer to incorporate this new signal into their previously-formed posterior belief, or do they prefer to revisit all past information alongside the new signal? Perceptions of sufficiency need not coincide with actual sufficiency, and it is the perception of sufficiency that may drive relevant behavior outside of the lab. For example, if individuals don't think of their posteriors as sufficient, then they may spend time and effort revisiting past information, relying on memory of previous signals, etc.

We design an experiment to test perceptions of posterior sufficiency. We find a high degree of perceived posterior insufficiency: A substantial majority (68%) of DMs prefer to re-examine the previous signals rather than relying on their posterior belief. However, this does not appear to reflect a rejection of summarizing information in general: Most of these DMs (76%) would prefer to update from the *Bayesian* posterior rather than revisit the previous signals. Thus, it appears that individuals do not view their posterior beliefs as sufficient—not because they are reluctant to rely on summary statistics, but because they do not trust their own posteriors. Consistent with this, we find that the main predictor of choosing to revisit past information rather than relying on one's own posterior is low confidence in the posterior belief. The fact that individuals distrust their own posteriors but are happy to rely on expert summaries presents potentially important avenues for information architecture and welfare-enhancing information summary aids.

Section 2 outlines our experimental design. We recruited 860 subjects through Prolific for an online experiment. In Part 1, subjects complete standard updating tasks and report their associated confidence with each guess. Part 2 gives subjects experience with updating from their posterior belief and updating from previous signals. After this, in Part 3, we tell subjects that they will complete more updating tasks in the future, and we ask them whether they would prefer to update from their posterior belief or revisit the previous signals in these future tasks. This elicitation provides our main measure of whether DMs view their posterior belief as a sufficient statistic. Following this, we elicit the information preference again, but now for one of three randomly selected conditional events: small vs. large signal sample size, questions of low vs. high confidence, or when facing few vs. many future decisions. Finally, we tell subjects that there is a "statistical guess" (i.e., the Bayesian posterior) that the computer could provide them, and we ask whether they would prefer to update from their posterior, update while revisiting the previous signals, or update from the statistical guess.

We present our main results in Section 3. We find that 68% of subjects prefer to revisit past information rather than relying on their posterior belief. This large degree of "perceived insufficiency" of posterior beliefs appears to result from DMs' distrust in their posteriors, which we show in three ways. First, we run an additional condition that is analogous to our main experiment but where future "updating" tasks do not provide any additional signals. In other words, we ask individuals whether they would prefer to revisit the original signals or would rather rely on their previously-formed posterior when the task is *exactly the same* as it had been originally. Here still, 60% of subjects prefer to revisit the signals. This indicates that many DMs doubt their posterior beliefs even when they don't need to incorporate new information, and incorporating new information exacerbates this distrust. Second, in our main treatment, 76% of subjects who prefer to revisit previous signals chose to rely on the Bayesian posterior when given the option. This indicates that a substantial majority of these individuals are not averse to relying on summarized information per se, but they are averse to relying on their own summary. Finally, when we elicit individuals' information preference separately for situations in which they had reported high confidence and low confidence, we find that the preference to revisit past signals is much stronger for situations of low confidence. In contrast, we find that information preference does not depend on the number of signals or on the number of updating tasks, suggesting that "complexity" along these dimensions does not push toward sufficiency. Thus, taken together, our results indicate that individuals do not view their posterior beliefs as sufficient statistics because, and when, they are not confident in the judgments they formed.

The final portion of our analysis addresses the secondary question of whether individuals are well-calibrated in their information choices. We find evidence of naivety: The preference for revisiting previous signals is unrelated to whether the subject was closer to the Bayesian benchmark when updating using the previous signals compared to when updating from their posterior. Instead, confidence is the main predictor of information choices: Individuals who are more confident in updating when they see the previous signals compared to when they see their posterior belief are more likely to prefer revisiting previous signals, even though this does not necessarily improve the accuracy of their beliefs. We discuss how this relates to recent work on metacognition (e.g., Enke and Graeber, 2023) below.

We see a few important implications of our results. First, a benefit of our experiment is the ability to present the choice of revisiting signals versus relying on one's own posterior on equal footing, and to present individuals with their chosen information in a costless and unbiased manner. However, outside the lab, revising previous signals might be more costly than relying on one's posterior since individuals might have to spend time and effort relocating old information, retrieving it from memory, etc. If individuals still prefer to revisit information when it is costly to do so, then it's possible that decision fatigue, satisficing (i.e., revisiting a few signals rather than the full set of prior information), memory biases (i.e., asymmetric recall as in Zimmermann, 2020), or salience of information might introduce bias and errors into updating from previous signals. [1]

[1]For instance, Bohren et al. (2024) shows that imperfect recall and attention to more salient payoff can distort mental representation of asset outcomes in different ways.

Second, our results suggest that individuals are willing to trust experts' summaries even when they don't trust their own summaries, and are indeed better at forming new beliefs when relying on the experts' summaries. This suggests scope for information design and other decision aids that can help individuals incorporate information in a rational manner.

Related Literature. Posterior sufficiency is a property that a general updating rule may or may not satisfy. For the Bayesian updating rule, posterior sufficiency is automatically satisfied: The posterior is a one-to-one function of the prior given the signal, so the posterior alone contains all the information needed to continue updating. For a general non-Bayesian updating rule, this need not hold. For instance, an updating rule in which the decision-maker's posterior is a strictly monotone transformation of the Bayesian posterior would satisfy posterior sufficiency,[2] while one in which the posterior is the weighted average of the prior and the Bayesian posterior (as in Epstein et al., 2010) would violate it. For a theoretical exploration of posterior sufficiency, see Cripps (2018) who links it to the divisibility property of updating rules; divisibility guarantees that processing information in pieces is equivalent to processing all the signals at once.[3] Our work naturally relates to explorations of posterior (in)sufficiency in updating, but our primary question is whether individuals *perceive* their posterior to be a sufficient statistic, not whether it actually is or is not.

In terms of empirical work, our paper relates most directly to work on belief updating and Bayesian reasoning. A large literature documents systematic deviations from the Bayesian benchmark in standard updating environments, including base-rate neglect, conservatism, and order dependence; see Benjamin (2019) for a recent survey. Surprisingly, this literature has paid little attention to posterior sufficiency, with Möbius et al. (2022) a notable exception. Sarnoff (2025) addresses this gap and tests posterior sufficiency directly, finding that a majority of subjects violate the principle, which we also observe in our data. Kieren et al. (2024) experimentally identify attention and memory as a candidate explanation for the gap between posterior beliefs in sequential vs. simultaneous updating problems (i.e., as-if posterior insufficiency). Relatedly, Raymond and Wittrock (2024) analyze, both theoretically and experimentally, when it is optimal for a memory-constrained agent to form a posterior belief versus remember the previous signals as a function of uncertainty about payoff-relevant states and the number of signals. Finally, Esponda and Xu (2026) study the slightly different principle of forecast sufficiency: that decisions should not depend on whether a probability is objective or is subjectively forecast, but should depend only on the value of this probability. Experimentally, they find that investment decisions under risk vs. ambiguity do not satisfy this property. Our work complements

[2]See Section 1.2 in Cripps (2018) for more examples of divisible non-Bayesian updating rules.

[3]In particular, Cripps (2018) shows that any divisible updating rule must have the structure in which posterior μ_1 can be represented as $\mu_1 = F^{-1}$ [Bayesian update of $F(\mu)$] where F is a bijection and μ is the prior. In words, you start with prior μ, you apply F to get a shadow prior $F(\mu)$, then you update the shadow prior using standard Bayes' rule to get a shadow posterior, and, finally, you apply F^{-1} to the shadow posterior to get back your actual updated belief μ_1. Obviously, when F is the identify function, $F(\mu) = \mu$, you recover standard Bayesian updating. When F is something else you get a non-Bayesian but divisible updating rule. The key point for posterior sufficiency is that because F is a bijection, μ_1 is a one-to-one function of μ given the signal, so the posterior μ_1 uniquely identifies the prior μ, and you can continue updating from μ_1 alone.

this recent interest in "as if" posterior insufficiency by showing that individuals do not *want* to rely on posterior beliefs as summary statistics and by documenting the mechanism behind this preference.

We also relate to recent work on metacognition and confidence. The perceived insufficiency of posterior beliefs that we document appears to stem from significant underconfidence in one's reasoning ability. This provides potentially promising opportunities for information design and decision tools to help individuals incorporate information confidently. Furthermore, Enke and Graeber (2023) show that people with higher confidence in Bayesian updating tasks are closer to Bayesian benchmarks. Our data shows no systematic relationship between performance and perceived posterior sufficiency, but the strong correlation between confidence and perceived sufficiency suggests underlying metacognitive motivations.

Finally, our work contributes to the literature on the preferences and demand for instrumental and non-instrumental information. Studies on instrumental information have focused on whether the features of information structures and the misuse of information are related to the demand for information (Ambuehl and Li, 2018; Charness et al., 2021; Llorente-Saguer et al., 2024; Guan, 2024; Guan et al., 2025), while work on non-instrumental information focuses on the motivations behind information acquisition and avoidance in situations where the information is not decision relevant (Ahlbrecht and Weber, 1996; Eliaz and Schotter, 2010; Ely et al., 2015; Ganguly and Tasoff, 2017; Nielsen, 2020; Falk and Zimmermann, 2023; Masatlioglu et al., 2023). We find that confidence in using information is the key driver of information demand underlying perceived posterior insufficiency, similar to Eliaz and Schotter (2010)'s finding that individuals demand non-instrumental information in situations where this information can make them feel more confident about their decisions.

2 Experimental Design

Our experiment consists of four main parts and a short survey. The first two parts present individuals with standard belief updating problems and allow us to measure deviations from Bayesian updating across various formats. Part 3 elicits individuals' information preferences. Part 4 presents new updating problems; we use this only to incentivize the information choices made in Part 3. All of our main decisions are incentivized; we discuss details below. Furthermore, we require subjects to correctly pass comprehension questions before each of the four main parts. We include a full set of screenshots in the online appendix.

As we explain below, we have two treatments: the *NewSignal* treatment and the *NoNewSignal* treatment. We describe the design for the *NewSignal* treatment first as this is our main focus. Afterwards, we contrast the *NoNewSignal* treatment.

2.1 Part 1: Updating Tasks

We frame all updating tasks similar to the standard "book bag and poker chips" design. There are two bags, Bag A and Bag B. Bag A always contains 6 red balls and 4 blue balls, while Bag B always contains 4 red balls and 6 blue balls. We tell subjects that the computer will first select either Bag A or Bag B with given chance (the prior), and then will draw one or more balls with replacement from the secretly-selected bag (the signals). After observing the draws, we ask subjects to report their best guess of the chance that the secretly-selected bag is Bag A and to report their confidence in this guess.[4] Subjects face 20 such updating tasks in Part 1.

For the purposes of our main research questions, Part 1 updating tasks only serve to elicit subjects' posterior beliefs that they might rely on in the future. As a result, we tell subjects that their guesses will be relevant for future tasks, so they should report them carefully.[5]

Parameters The specific parameterizations of these tasks allow us to test for as-if posterior sufficiency in belief updating. However, since this is not the main focus of our paper, we refer the interested reader to Online Appendix B for design details and C.2 for findings. What is relevant is that induced prior beliefs span at least the range of 10% to 90% chance that Bag A is the secretly-selected bag, and the sample size ranges from small (1 selected ball) to large (15 selected balls).

2.2 Part 2: Framed Updating Tasks

Part 2 consists of ten more updating tasks constructed in the following way. For five of the updating tasks from Part 1, we construct two new types of questions each: the "update-from-posterior" question and the "update-from-previous-signals" question. In the update-from-posterior questions, we tell subjects their guess from the question in Part 1 (without showing the original prior and signals), present one new signal, and elicit their updated belief. In the update-from-previous-signals questions, we show subjects the original prior and signals from the same question in Part 1 (without showing their posterior), present the same new signal, and elicit their updated belief. Posterior sufficiency implies that the reported posterior should be the same in the update-from-posterior question and in the update-from-previous-signals question.

2.3 Part 3: Information Choice

Having experienced both "posterior sufficient" and "posterior insufficient" forms of updating in Part 2 (i.e., the updating-from-posterior and update-from-previous-signals tasks, respectively), Part 3 presents subjects with our main elicitation. We explain to subjects that they will face some more updating questions exactly like those in Part 2, and we ask them whether they would rather rely on their Part 1 posterior or whether they would prefer to revisit the previous signals. In other words, they are exactly choosing whether they want future updating problems to be of the update-from-posterior or update-from-previous-signals format from Part 2. We use the term "perceived posterior insufficiency" to refer to the choice of revisiting the previous signals rather than relying on the Part 1 posterior.

[4] To be more specific, we ask "*How sure are you in your guess?*" Recent work cautions the use of confidence judgments for identifying "mistakes" and improvable choices (Bernheim et al., 2026). We are not using confidence for this purpose, and we do not take a strong stand on how subjects interpret our confidence elicitation.

[5] As described in the following subsections, subjects might revisit their own guess from Part 1 plus a new signal and update their belief again. This means that their guesses from Part 1 will affect their performance and their incentives in the follow-up tasks.

We present this to subjects as a binary choice and they indicate their preferred option. After this, we elicit a measure of indifference: We ask subjects whether they would be willing to switch their choice if offered a small $0.05 bonus to do so.

Conditional Elicitations. After this, we elicit subjects' *conditional* information preference (and subsequent indifference measure) in one of three scenarios: small vs. large sample, low vs. high confidence, and few vs. many future decisions. Specifically, taking small vs. large sample as an example, we ask subjects "Suppose that in Part 4 you will revisit 2 rounds from Part 1, where you saw relatively few balls drawn (2-5 balls). Which information would you like to carry over from Part 1 for those questions?" Separately, we would ask them the same question, but for "relatively many balls drawn (12-15 balls)."

Expanding the Choice Menu. Finally, after the unconditional and conditional elicitations, we tell subjects that there exists a *Statistical Guess* to the tasks completed in Part 1 and 2 (i.e., the Bayesian posterior) that can be obtained by computer simulation. Specifically, we tell subjects that the computer will simulate the data generating process in a given task a billion times, keep the cases where the exact sequence of balls is observed, and calculate the fraction of times that Bag A was the secretly-selected bag. We ask subjects to choose between their own posterior, the previous signals, and the statistical guess; this choice would be implemented on new updating tasks in Part 4 if this response were randomly selected to be implemented. We did not break indifference in this expanded choice menu.

Incentives. We incentivize all of our elicitations in standard ways. First, we tell subjects that one of the guesses from Part 1, 2, or 4 will be selected for payment. The payment is determined by the Becker–Degroot–Marschak (BDM) mechanism (Becker et al., 1964), which gives them the best chance of receiving payment if they truthfully report their guesses.

Furthermore, we implement one of the information preference elicitations from Part 3. We randomly select either the unconditional, conditional, or expanded-menu question and subjects see the information that they selected in this question for their future guesses in Part 4.[6]

2.4 Treatments

The procedures above describe our main treatment that we refer to as the *NewSignal* treatment; "new signal" refers to the fact that subjects see an additional signal in Part 2, and, when they are reporting their information preferences, they know that they will also see a new signal in Part 4. We also run the *noNewSignal* treatment. This treatment is exactly the same as the *NewSignal* treatment, but subjects do not see a new signal in Part 2 or Part 4, and they know they will not see a new signal in Part 4 when they are reporting their information preferences.[7]

[6] If we randomly select to implement the question in which we break indifference, then we randomly select whether to offer the $0.05 bonus or not, and implement the information accordingly.

[7] To ensure subjects understand whether or not a new ball will arrive, we ask them to complete a comprehension question regarding whether or not they will expect a new ball in Part 4. Additionally, we ask them to predict (referred to as "prediction" question hereafter) whether they will make a different guess in Part 4 from Part 1

The *noNewSignal* treatment allows us to assess the extent to which any perceived posterior insufficiency is inherent to future updating. In other words, if individuals do not want to rely on their posterior as a sufficient statistic when they need to incorporate new information, are they even willing to rely on their posterior when they do *not* need to incorporate new information?

2.5 Implementation Details

We recruit 860 subjects through Prolific; 427 subjects participate in the *NewSignal* treatment and 433 participate in the *noNewSignal* treatment. Subjects receive a base payment of \$7. We also randomly select one comprehension question and reward \$0.05 for accuracy.[8] The median completion time was around 32 min.

3 Results

Our main question concerns the Part 3 information choices: Do subjects prefer to update from their posterior belief, or do they prefer to revisit the original signals?[9] Since we elicited this as a weak preference before breaking indifference, we consider subjects with weak and strict preference separately.[10]

Figure 1 presents our main result: We show the percentage of subjects, in both the *NewSignal* and *noNewSignal* treatments, who choose to update from their posterior versus choose to update from all of the previous signals. The light shaded portion of the bar indicates weak preference, while the darker portion of the bar indicates strict preference. When we include both weak and strict preferences, we find that 68% of subjects in the *NewSignal* treatment would rather rely on the previous signals than their posterior (67.92% vs. 50%, two-sided proportion test, $p < 0.001$). A similar pattern holds among subjects with strict preference: 30% strictly prefer to rely on the previous signals while only 13% strictly prefer to rely on posterior.[11]

This lack of perceived posterior sufficiency is present in the *noNewSignal* treatment as well, but the direction is weaker. Here, we see about 60% of subjects prefer to rely on the previous signals while 40% prefer to rely on their posterior (59.82% vs 50%, two-sided proportion test, $p < 0.001$). Among those with strict preference, the proportions are not significantly different

and why in a multiple-choice format. Subjects can select from choices, including "No; they should be the same because I will not see any new balls" (only 5% of subjects in the *NewSignal* treatment chose this option) and "Yes; they might be different because I will see a new ball" (only 13% of subjects in the *noNewSignal* treatment chose this option).

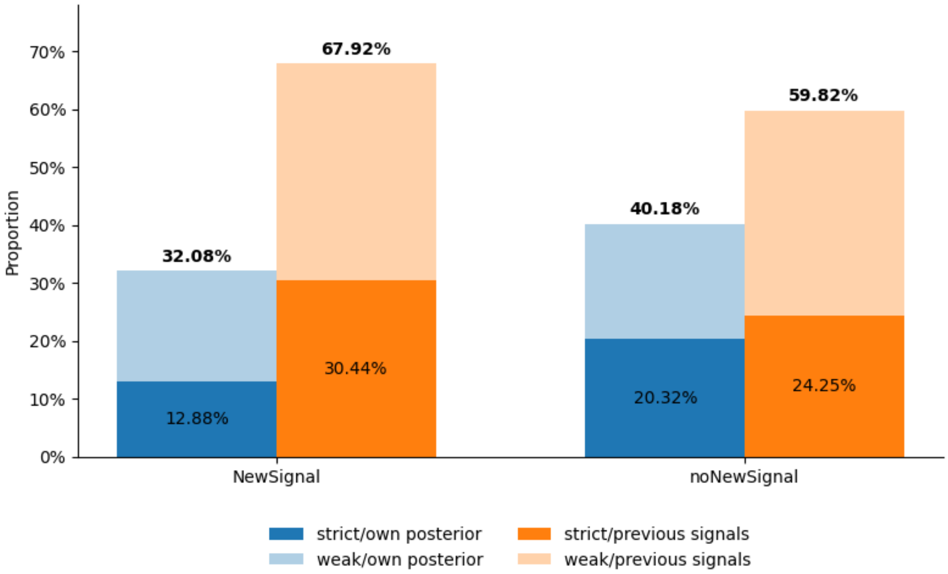
[8] We present results for the full sample. The findings are robust when restricting the sample to subjects who passed all comprehension questions in two attempts and did not choose any incorrect option in the "prediction" question (See online Appendix A and E.1)

[9] Throughout the results, we "revisit the original signals" and other similar phrases as shorthand, but note that this always includes revisiting the original prior along with the previous signals.

[10] While one could consider combining those with weak preference for relying on their posterior and those with weak preference for revisiting past signals into a single indifference class, we find that the data patterns look very similar for those with weak and strict preference—and therefore look very different across the two weak preference groups—which justifies analyzing them separately.

[11] Conditional on having a strict preference, those who prefer to revisit previous signals constitute the majority (70.27% vs. 50%, two-sided proportion test, $p < 0.001$).

Figure 1: Distribution of Information Preferences in the Unconditional Binary-Choice Menu



Notes: Light blue and orange portions represent the proportion of subjects who weakly prefer their own posterior and previous signals respectively, indicated by the numbers above the bars. The dark blue and orange portions represent the subjects who strictly prefer own guess and previous signals, respectively, out of all the subjects within the treatment.

from equal split: 20% strictly prefer to rely on the previous signals compared to 24% who strictly prefer to rely on posterior.

Putting this together, our results present strong evidence that individuals do not want to rely on their posterior beliefs. The higher rate of perceived posterior insufficiency in the *NewSignal* treatment relative to the *noNewSignal* treatment suggests that the expectation of incorporating future information increases subjects' unwillingness to rely on their posterior beliefs (two-sided proportion tests, $p = 0.013$ for weak preference and $p = 0.041$ for strict preference). However, even when they will not incorporate any new information, a majority of subjects still choose not to rely on their posterior beliefs, pointing to a more fundamental distrust of one's own beliefs as a summary statistic of past information.

Result 1 *Perceived posterior insufficiency is widespread and is significantly more common in the NewSignal treatment (about 70%) than in the noNewSignal treatment (about 60%).*

3.1 Do Individuals Reject Summary Statistics In General?

Having established our main result, we turn to secondary results to understand the determinants of this preference. One possibility is that individuals reject summary statistics wholesale and would always prefer to see the full set of information—either because they do not believe updating from a summary statistic is ever the correct procedure, or they prefer to update from signals nevertheless. The alternative is that individuals would accept a summary statistic if they were certain that it accurately incorporated all past information. Secondary features of our experimental design allow us to test between these two hypotheses. In particular, recall that, after we elicited the information choice (and conditional information preferences, discussed below), we expanded the choice set to include a "statistical guess," i.e., the Bayesian posterior. If those who initially selected to update from past signals switch to updating from the Bayesian posterior in this expanded menu, then it would indicate that they do not simply reject summary statistics in general.

Figure 2 overlays the responses from this expanded choice set onto the main results from Figure 1, focusing on the *NewSignal* treatment for ease of exposition. Among subjects who initially chose to rely on the previous signals, about 76% would prefer to rely on the Bayesian posterior when given the option. Furthermore, among subjects who initially chose to rely on their own guess, two-thirds or more also would prefer the Bayesian posterior to their own guess.[12] Thus, we find that a vast majority of subjects would prefer to rely on a simple summary statistic—the Bayesian posterior—if given the option (two-sided proportion test, $p < 0.001$). Importantly, willingness to rely on the Bayesian posterior is optimal: We designed Part 1 questions to compare updating from the Bayesian posterior, updating from one's own posterior, and updating from the previous signals, and we find that individuals are closer to Bayesian benchmarks when updating from the Bayesian posterior (see online Appendix C.2 for details).

There are two natural reasons why individuals might be willing to rely on the Bayesian posterior even if they were unwilling to rely on their own posterior. The first is that, by introducing the Bayesian posterior into the menu, we necessarily have to explain it, and this could potentially teach subjects the concept of a sufficient statistic or otherwise change their preferences for using one.[13] The second is that individuals distrust their own posterior, but do not necessarily distrust summary statistics in general, so they are happy to rely on the Bayesian posterior when given the option. Two pieces of evidence are consistent with this latter mechanism.

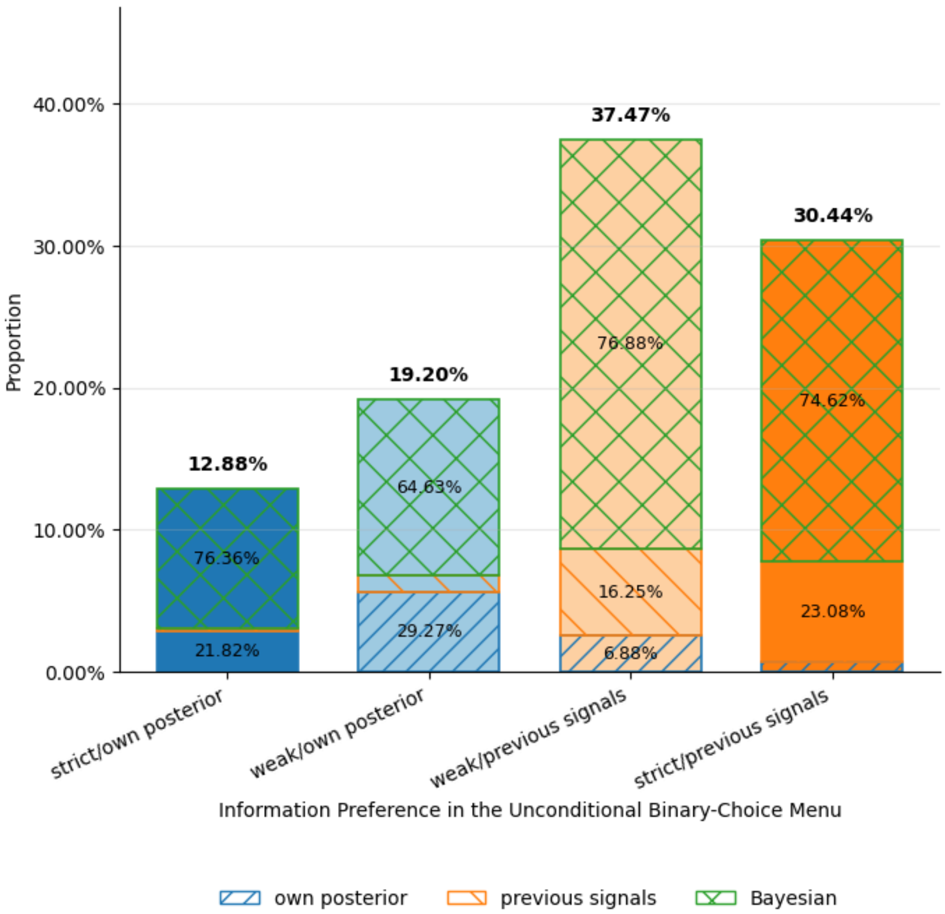
First, we asked subjects whether they would prefer to rely on their own posterior or revisit the previous signals separately for situations of low versus high confidence. We find that 72% of subjects prefer to revisit previous signals in situations of low confidence, but this preference weakens substantially when confidence is high; here, only 56% prefer to revisit previous signals (McNemar test, $p = 0.002$).[14] Thus, this conditional elicitation reiterates that low confidence in one's posterior is a significant driver of the preference to revisit previous signals. Second, we find further support that subjects did not reject the concept of summary statistics in general from the free-response questions in our post-experiment survey. As an example, one subject who chose to revisit previous signals originally yet switched to the Bayesian posterior when this

[12] Notably, 15–20% of participants who weakly prefer to revisit the previous signals in the binary choice still choose this even when the statistical guess is available. Among those who weakly prefer their own posterior in the binary choice, an even larger share (close to 30%) still prefer to rely on it instead of the statistical guess.

[13] Note, however, that this might also have taught them that their own posterior was sufficient as well, so it's not clear how this would change the preference for relying on one's own posterior.

[14] Among individuals with strict preferences, 28% prefer to revisit signals and 19% prefer their own posterior in situations of low confidence; 34% prefer to revisit signals and 15% prefer their own posterior in situations of high confidence.

Figure 2: Distribution of Preferences in the Expanded Menu Conditional on the Preferences in the Unconditional Binary-Choice Menu



Notes: The numbers above the bars indicate the proportion of subjects with the corresponding preference in the unconditional binary choice menu. The numbers within the top stack bar represent the proportion of subjects who prefer the *Statistical Guess* conditional on their preferences in the unconditional binary choice menu.

option was provided said "I felt statistical guess would be a good aggregate of data available for me to make a correct future choice;" another said: "It is way better for an algorithm or computer to guess what the chances were than me. At least with that anchoring, I felt like I could get closer than with my own guesses." See Appendix D for a full classification of free response text and correlation with subjects' choices.

Taken together, these results present a clear mechanism behind posterior insufficiency: Individuals prefer to revisit past information when they are not confident in the judgments they formed in light of this information.

Result 2 *A majority of subjects who do not perceive their posterior as sufficient would rely on the Bayesian posterior. Furthermore, perceived posterior insufficiency is more prominent when subjects are less confident in their posterior beliefs.*

3.2 Are Individuals Less Likely to Revisit Past Information When It Is Relatively More Costly To Do So?

The high degree of perceived posterior insufficiency we observe in our experiment might be counterintuitive from a "complexity" perspective: It might seem relatively more difficult to revisit many signals rather than relying on a simpler summary statistic. We included our two other conditional information preference elicitations (many vs. few signals and many vs. few future questions) to explore this potential sensitivity.

We asked individuals whether they would rather rely on their posterior or on the previous signals separately for questions with many signals and questions with few signals. We find no significant difference in information preference between the two (see online Appendix C.1 for details). Second, we asked individuals whether they would rather rely on their posterior or on the previous signals separately for answering many future questions versus answering few future questions. We again find no significant difference between the two (see online Appendix C.1 for details). Thus, while it might be the case that revisiting signals is more difficult when there are many signals or when there are many questions, individuals' uncertainty in their posteriors appears strong enough to dominate the complexity of revisiting the information in our data.[15]

Result 3 *Individuals are equally likely to prefer revisiting previous signals when there are few vs. many signals, and are equally likely to prefer revisiting previous signals when there are few vs. many future questions.*

3.3 Are Individuals Sophisticated in Their Information Preferences?

Given that it appears subjects doubt their own posterior beliefs and that this drives posterior insufficiency, a natural question is whether this reluctance to rely on one's own posterior is well-

[15]We find further support for this finding from the free-response questions in our end-of-experiment survey: While about 10% of those who chose to revisit their own posterior stated that they did so to avoid re-processing the previous evidence, 65% of subjects chose to revisit previous signals stated that they did so because they felt uncertain about their own posterior or felt more comfortable with the original chances and signals. See online Appendix D for details.

founded. We assess sophistication under the following assumption: A sophisticated subject should select the information that maximizes their likelihood of earning the bonus in Part 4. In other words, subjects whose guesses were closer to the Bayesian benchmark in the update-from-previous-signals questions compared to the update-from-posterior questions in Part 2 should be those subjects who are less willing to rely on their posterior beliefs for future guesses.[16]

Recall that we construct the "update-from-posterior" and "update-from-previous-signals" questions by presenting the information from the same Part 1 questions in two different ways. In both cases, the additional signal is identical and therefore the implied Bayesian posterior is the same, so we can compare the performance between the two question types. We construct a relative performance measure by first calculating the deviations from the Bayesian benchmark in the update-from-posterior question and in the paired update-from-previous-signals question, taking the difference between these, and then averaging across all five pairs for each subject. We construct a relative confidence measure (i.e., whether a subject was relatively more confident in the update-from-posterior questions or in the paired update-from-previous-signals questions) in a similar way. We then regress an indicator for weakly preferring previous signals on these relative performance and relative confidence measures. A positive relative performance means that the subject was relatively better at updating when relying on their posterior, and we would expect a negative correlation with the dependent variable if the subject were sophisticated.[17]

Table 1: Marginal Effects on Information Preferences

			$\mathbb{P}$(Chose Previous Signals)	
Variable	Baseline	Δ	Marginal Effect	p-value
Part 2				
Relative performance	0.000	0.100	-0.015 [-0.062, 0.032]	0.522
Relative confidence	0.143	0.100	0.009 [0.001, 0.016]	0.020

Notes: Control variables include an indicator for Bachelor's degree and above, an indicator for STEM major, the interaction term between the two, and a categorical variable for the reported level of Bayesian knowledge. "Baseline" column contains the values at which we evaluate the marginal effects: we set relative performance at 0 (i.e., subjects equally good or bad at both question types), and relative confidence at its median. We set the control variables at the modal values. "Δ" is the marginal change in the covariate. 95% confidence intervals, computed using the delta method, are reported in square brackets below the marginal effects. Since logit function is monotonically increasing, we test the significance of each coefficient rather than the marginal effect using Wald test and report the resulting p-values in this table.

Table 1 presents these results. We find that relative performance in Part 2 is not predictive of preferences for previous signals. Thus, it is not the case that the subjects who prefer to revisit

[16] Only 8% of subjects are equally close to the Bayesian benchmark in both types of questions. The rest are roughly equally split between those who are closer to the Bayesian benchmark in the update-from-previous-signals question and those who are closer to the Bayesian benchmark in the update-from-posterior questions. See Online Appendix C.3 for details.

[17] In online Appendix E.2, we show the results using alternative relative performance measures, including control variables from Part 1, or replacing the outcome variable with ordered preference that differentiates weak and strict preferences. Our results are robust across these alternative models.

previous signals are those who performed better in the update-from-previous-signals questions compared to the update-from-posterior questions. Instead, consistent with our analysis above, relative confidence is the key driver: Subjects who prefer to revisit previous signals are those who felt relatively more confident in the update-from-previous-signals questions compared to the update-from-posterior questions, regardless of their actual performance.[18] Taken together, our results indicate that subjects are not well-calibrated in their information choices.

Result 4 *Information choices are not sophisticated: Subjects who are farther from the Bayesian posterior in the update-from-previous-signals questions relative to the update-from-posterior questions are no less likely to prefer revisiting previous signals.*

4 Discussion

We find that a majority of individuals do not view their posterior beliefs as sufficient statistics and instead would rather revisit past signals when incorporating new information. These individuals would rely on the Bayesian posterior, but appear to lack confidence in the sufficiency of their own beliefs. This finding is in contrast to Bayesian and many non-Bayesian models of reasoning and has a few important implications. First, perceived posterior insufficiency might exacerbate other biases such as memory distortions, overreaction to salient signals, or decision fatigue, since relying on past information presents more opportunities for these biases to emerge. Second, willingness to rely on the Bayesian posterior suggests that Bayesian summaries of information and related decision aids could substantially improve future belief updating and individual welfare. Finally, the fact that perceptions of posterior insufficiency result from metacognitive uncertainty about individuals' own reasoning, rather than reflections of the accuracy of updating, suggests room for information design and other channels to increase decision-makers' confidence in their own reasoning abilities.

References

Ahlbrecht, M. and Weber, M. (1996). The resolution of uncertainty: An experimental study. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 593–607.

Ambuehl, S. and Li, S. (2018). Belief updating and the demand for information. *Games and Economic Behavior*, 109:21–39.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:69–186.

[18]This further suggests a mismatch between relative performance and confidence (see online Appendix C.4 for formal analysis).

Bernheim, B. D., Lucia, A., Nielsen, K., and Sprenger, C. D. (2026). When are decisions improvable? an evaluation of diagnostic methods. *NBER Working Paper*.

Bohren, J. A., Hascher, J., Imas, A., Ungeheuer, M., and Weber, M. (2024). A cognitive foundation for perceiving uncertainty. Technical report, National Bureau of Economic Research.

Charness, G., Oprea, R., and Yuksel, S. (2021). How do people choose between biased information sources? evidence from a laboratory experiment. *Journal of the European Economic Association*, 19(3):1656–1691.

Cripps, M. W. (2018). Divisible updating. *Manuscript, UCL*.

Eliaz, K. and Schotter, A. (2010). Paying for confidence: An experimental study of the demand for non-instrumental information. *Games and Economic Behavior*, 70(2):304–324.

Ely, J., Frankel, A., and Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1):215–260.

Enke, B. and Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics*, 138(4):2021–2067.

Epstein, L. G., Noor, J., and Sandroni, A. (2010). Non-bayesian learning. *The B.E. Journal of Theoretical Economics*, 10(1).

Esponda, I. and Xu, L. (2026). Forecast errors and forecast sufficiency under ambiguity.

Falk, A. and Zimmermann, F. (2023). Attention and dread: Experimental evidence on preferences for information. *Management Science*.

Ganguly, A. and Tasoff, J. (2017). Fantasy and dread: The demand for information and the consumption utility of the future. *Management Science*, 63(12):4037–4060.

Guan, M. (2024). Choosing between information bundles.

Guan, M., Oprea, R., and Yuksel, S. (2025). Beyond instrumental value: How complexity shapes information demand. Working paper.

Kieren, P., Mertes, L., and Weber, M. (2024). Information partitioning, learning, and beliefs. *Learning, and Beliefs (March 15, 2024)*.

Llorente-Saguer, A., Oliveros, S., et al. (2024). Beyond value: on the role of symmetryin demand for information. Technical report.

Masatlioglu, Y., Orhun, Y., and Raymond, C. (2023). Intrinsic information preferences and skewness. *American Economic Review*, 113(10):2615–2644.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2022). Managing self-confidence: Theory and experimental evidence. *Management Science*, 68(11):7793–7817.

Nielsen, K. (2020). Preferences for the resolution of uncertainty and the timing of information. *Journal of Economic Theory*, 189:105090.

Raymond, C. and Wittrock, L. (2024). Optimal memory with sequential learning: Signals or posterior beliefs. *Available at SSRN 4765215*.

Sarnoff, K. (2025). The structure of sequential updating. *Working Paper*.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–61.