Cognition 245 (2024) 105717

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Changing your mind about the data: Updating sampling assumptions in inductive inference

Brett K. Hayes [a,*], Joshua Pham [a], Jaimie Lee [a], Andrew Perfors [b], Keith Ransom [b], Saoirse Connor Desai [a,c]

[a] School of Psychology, University of New South Wales, Sydney, Australia
[b] School of Psychological Sciences, University of Melbourne, Australia
[c] School of Psychology, University of Sydney, Australia

ARTICLE INFO

Keywords:
Inductive reasoning
Sampling assumptions
Belief revision
Bayesian models

ABSTRACT

When people use samples of evidence to make inferences, they consider both the sample contents and how the sample was generated ("sampling assumptions"). The current studies examined whether people can update their sampling assumptions – whether they can revise a belief about sample generation that is discovered to be incorrect, and reinterpret old data in light of the new belief. We used a property induction task where learners saw a sample of instances that shared a novel property and then inferred whether it generalized to other items. Assumptions about how the sample was selected were manipulated between conditions: in the property sampling frame condition, items were selected because they shared a property, while in the category sampling frame condition, items were selected because they belonged to a particular category. Experiment 1 found that these frames affected patterns of property generalization regardless of whether they were presented before or after the sample data was observed: in both cases, generalization was narrower under a property than a category frame. In Experiments 2 and 3, an initial category or property frame was presented before the sample, and was later retracted and replaced with the complementary frame. Learners were able to update their beliefs about sample generation, basing their property generalization on the more recent correct frame. These results show that learners can revise incorrect beliefs about data selection and adjust their inductive inferences accordingly.

1. Introduction

People frequently rely on samples of existing evidence to make inferences, predictions and decisions. For example, a person deciding where to stay on holiday may sample online recommendations of "hotels with more than 3-stars". However, our beliefs about such samples can change over time as new information comes to light. Our holiday planner may have initially believed, for example, that the star ratings represented a consensus measure of hotel quality as reported by previous guests. Later she learns that these ratings are often generated by the hotels themselves, as a form of self-promotion (Pitrelli, 2022). Under these circumstances, one might expect the holiday planner to revise their hotel choices even if they receive no new hotel recommendations.

Research on inductive inference has revealed much about how we make inferences from samples of evidence (see Feeney, 2018 for a review). However, relatively little attention has been given to how (or whether) inferences change in light of new information about how the samples were generated. The answer to this question bears on issues like what information is encoded during the inference process and whether people can revise their inferences when they discover that their previous beliefs about the sampling process were incorrect. Such issues are important in information-rich environments where people often receive conflicting reports about how a sample of evidence was generated (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). The focus of this paper is on exploring this question.

1.1. Samples and sampling assumptions in induction

In studies of property induction, learners see a sample of instances that share some novel property (e.g., animals that have a particular type of blood), and use this to infer how far the property generalizes (e.g., whether other animals that haven't been observed also have that type of blood). A long tradition of induction research, dating back to the seminal works of Rips (1975) and Osherson, Smith, Wilkie, Lopez, and Shafir

* Corresponding author at: School of Psychology, University of New South Wales, NSW 2052, Australia.
E-mail addresses: B.hayes@unsw.edu.au (B.K. Hayes), saoirse.connordesai@sydney.edu.au (S.C. Desai).

https://doi.org/10.1016/j.cognition.2024.105717
Received 5 July 2023; Received in revised form 4 January 2024; Accepted 5 January 2024
Available online 18 January 2024

(1990), emphasizes the role of the *sample contents* in the inference process. Such work has revealed important principles that guide how we generalize from samples of evidence to a broader population. These include the similarity of the sample to the target of the inference; learners are more likely to generalize to targets that are similar to the instances observed in the evidence sample (Hayes & Thompson, 2007; Osherson et al., 1990). Another important principle is sample diversity; learners are more likely to infer that a property generalizes to other category members, after observing that a diverse sample of instances (e.g., horses, mice and dolphins) have the property as compared to a less diverse sample (e.g., horses, zebras, donkeys) (Hadjichristidis, Geipel, & Gopalakrishna Pillai, 2022; Heit & Hahn, 2001; Liew, Grisham, & Hayes, 2018).

These principles have been shown to be robust across a range of learning contexts and stimuli (see Hayes & Heit, 2018, for a review). They have also been shown to "scale up" to how people use samples of evidence to draw inferences about consequential issues such as future global warming (Kary, Newell, & Hayes, 2018) or medical diagnosis (Kim & Keil, 2003).

Recent work, however, has highlighted another important component of the inferential process – people's *sampling assumptions* – their *beliefs about how the sample was generated* (Hayes, Navarro, Stephens, Ransom, & Dilevski, 2019; Tenenbaum & Griffiths, 2001). One line of research has shown that people's inferences about a given sample depend on their beliefs about the *intentions* behind sample selection (see Hayes, Liew, Connor Desai, Navarro, & Wen, 2023 for a review). When adults believe that sample contents were selected by someone with helpful intentions and a knowledge of the relevant domain, they are likely to take factors such as the size and diversity of a sample into account when judging how far a property generalizes. However, the impact of these factors diminishes if they believe the sample was generated randomly (Hayes, Navarro, et al., 2019; Navarro, Dry, & Lee, 2012; Ransom, Hendrickson, Perfors, & Navarro, 2018; Ransom, Perfors, Hayes, & Connor Desai, 2022; Ransom, Perfors, & Navarro, 2016). Young children are also more likely to factor in the sample composition into their inferences when samples are selected by a knowledgeable teacher (Rhodes, Gelman, & Brickman, 2010).

A related issue and the focus of the current work, concerns the impact of *sampling frames* – causal constraints on the sampling process that mean that only certain types of instances can be observed. Such constraints often arise because our ability to sample relevant evidence is limited by available time or resources. Hence some type of selective sampling strategy is necessary. For example, when searching for holiday accommodation, one could make the search more tractable by only examining reviews of those hotels managed by a popular hotel chain. In this case, the sample is subject to a *category frame* – inclusion of instances in the sample is dependent on them belonging to a particular category. An alternative *property frame* strategy would be to search according to some relevant property or criterion (e.g., only search for hotels with 5-star reviews). As detailed below, different inferences can arise from the same sample of evidence depending on whether the sample was believed to be subject to a category frame or a property frame. We see this as an important issue to study because many, if not most, samples of evidence that people encounter outside the laboratory are subject to some sort of selection constraint or bias (Hogarth, Lejarraga, & Soyer, 2015).

To better understand the implications of different sampling frames, let us turn to an example that is closer to the scenarios used in the current studies. Imagine that you were tasked with discovering which animals on a previously unexplored island had a particular enzyme in their blood. Resource constraints limit the extent of exploration and so a sampling frame needs to be applied. Category sampling would involve sampling from one animal category and seeing which category members have the enzyme. Property sampling would involve taking an initial sample of animals known to have the enzyme and examining what types of animals were included in this sample. Crucially, the different frames license very different kinds of inferences (see Fig. 1 for an illustration).

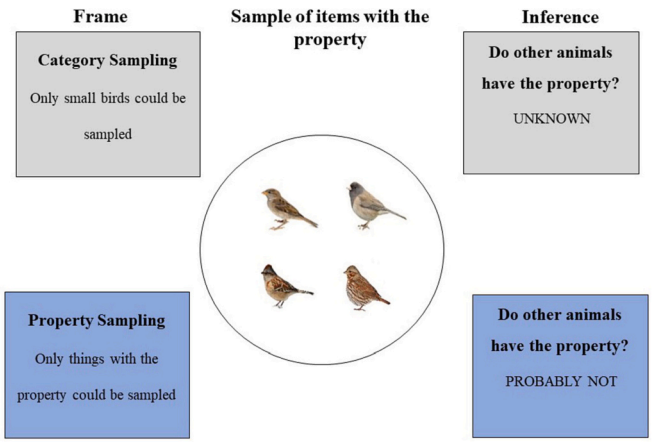

Fig. 1. Example of different sampling frames applied to the same sample of evidence.
Note: The Figure illustrates how category and property frames lead to different inferences when applied to the same sample of items that share a property.

Under category sampling, discovering that most instances in a sample of birds have the enzyme is informative about how the property is distributed *within* that category, but tells us little about whether the property *generalizes* to other animals. The absence of instances from other categories in the sample is attributable to the frame. Under property sampling, however, the same sample data is highly informative for property generalization. In this case, there are no obvious restrictions on the categories that *could* have been included in the sample. The fact that only members of a single category appear in the sample strongly implies that the property does not generalize outside that category. In other words, under category sampling, the absence of evidence in the sample is not informative. Under property sampling, absence of evidence *is* evidence of absence.

Previous research shows that many adult learners are sensitive to the implications of different sampling frames. When presented with a sample containing members of a single category that share some property, those given property frame instructions are less likely to generalize the property to other instances than those given category frame instructions (Hayes et al., 2023; Hayes, Banner, Forrester, & Navarro, 2019; Lawson & Kalish, 2009; Ransom et al., 2022). Observing larger samples of evidence leads to further "tightening" of generalization under property sampling but has less of an effect on property sampling (Hayes et al., 2023; Hayes, Banner, et al., 2019).

1.2. Revising assumptions about the sampling process

One limitation of previous research on the effects of sampling assumptions is that, in most cases, such assumptions are manipulated before the sample is observed; there has been no attempt to examine situations where previous beliefs about how a sample has been generated need to be revised or updated. Situations that call for revision of initial sampling assumptions are, however, common outside the laboratory. For example, an investor may see advertisements from an investment company featuring high-performing mutual funds, assuming that these were a representative sample of company products. They may later discover that these funds were carefully selected for the purpose of advertising and are atypical of the company's portfolio (Koehler & Mercer, 2009). Likewise, in our opening example, our holiday planner would have to reassess their hotel choices when they learned about how star ratings were actually generated. In these cases, people observe a sample of information that seems to fit one set of assumptions about how the sample data were generated, but later learn that a different sampling process was operating.

A crucial question, therefore, is whether people revise their sampling assumptions and adjust their inferences when they receive new information about the sampling process. There is some evidence which suggests that people may struggle with this. Ransom et al. (2022) examined the impact of different sampling assumptions (e.g., intentional vs. random; category frame vs. property frame) on property induction when information about the sampling process was presented before or after observation of sample data. When the sampling process was explained before the data, contrasting assumptions affected subsequent inferences. For example, people presented with a property sampling frame before observing the data sample subsequently showed narrower property generalization than those presented with a category frame, replicating the main finding from previous studies of sampling frames (Hayes, Banner, et al., 2019; Lawson & Kalish, 2009). In contrast, frames presented after the data had less impact – there was little difference in patterns of generalization between category and property frames conditions.

More broadly, the finding that initial sample frames have more of an impact on people's inferences than frames presented after a sample is consistent with the findings of "primacy effects" in impression formation (e.g., Anderson, 1965, 1973; van Overwalle & Labiouse, 2004) and the perseverance of initial beliefs in self- and social-perceptions (e.g., Peake & Cervone, 1989; Ross, Lepper, & Hubbard, 1975). Ross et al. (1975), for example, found that the effects of initial positive or negative feedback about performance on a novel task that participants completed themselves or saw others complete, persevered despite subsequent debriefings that discredited the feedback. Primacy effects have also been reported in contingency learning, where initial evidence has a greater influence on causal judgments than does later evidence (Dennis & Ahn, 2001).

In a seminal review, Hogarth and Einhorn (1992) found that the effects of the order of evidence presentation on belief revision were complex, and depended on a range of factors, such as the amount and complexity of the evidence and whether judgments are made on a trial-by-trial basis as new evidence is encountered or after all the evidence has been observed. Primacy effects were most often observed when judgments were only required after data presentation was complete – which has been the procedure used in previous sampling frames studies.

An important implication of this work is that beliefs about the sampling process may change the way that incoming sample data is encoded, but have little effect on the subsequent retrieval of sample data. Ransom et al. (2022) outlined several possible variants of this *encoding-only* hypothesis. For example, it may be that early in the sampling process, people generate competing hypotheses about how far a property generalizes (e.g., "only members of the category observed in the sample have the property", "members of categories similar to the sample have the property", "all categories in this domain have the property") and update their beliefs about the likelihood of each hypothesis as each new sample instance is encountered. According to this view, the sampling assumptions that are in place before the sample is observed change the way the sample data are encoded in memory. Hence, initial sampling assumptions will dominate subsequent inferences from the sample – even when the original sampling assumption is retracted (i.e., acknowledging that the original assumption was false or incorrect).

Findings from work on belief updating in a range of contexts including narrative comprehension (e.g., Kendeou, Butterfuss, Kim, & Van Boekel, 2019; Kendeou, Smith, & O'Brien, 2013; Rapp & Kendeou, 2007) and juror decision-making (e.g., Harris & Hahn, 2009; Lagnado & Harvey, 2008; Shengelia & Lagnado, 2021), however, suggest that this view may be overly pessimistic. For example, Rapp and Kendeou (2007) asked participants to read stories and rate their trait impressions of characters at various points in the narrative. Participants' first impressions (e.g., that Travis was "clumsy" because he fell over repeatedly at a dance club) were revised when a subsequent refutation of the impression was encountered, especially when this refutation contained an alternative causal explanation (e.g., the dance floor had just been waxed). Similarly, Mickelberg, Walker, Ecker, Howe, and Perfors (2023), found that person impressions based on behavioral descriptions were revised when these were found to be incorrect, and that this was true regardless of whether the misinformation was positive or negative.

In related work, Lagnado and Harvey (2008) presented mock jurors with sequentially presented evidence in fictitious criminal cases. In some conditions, new evidence discredited evidence presented earlier in the sequence (e.g., after hearing that a neighbor had seen the suspect near the crime scene, it was learned that the neighbor had a long-standing grudge against the suspect). The discrediting evidence reduced jurors' estimates of the suspect's guilt. Moreover, this effect was stronger when the discrediting information came at the end of the sequence rather than when it appeared in the middle of the sequence.

Other relevant findings come from studies of the continued influence effect (CIE; Connor Desai & Reimers, 2019, 2023; Connor Desai, Pilditch, & Madsen, 2020; Ecker, Lewandowsky, Swire, & Chang, 2011; Johnson & Seifert, 1994). In this paradigm, an (incorrect) explanation for an event is given that is consistent with the observed evidence (e.g., a plane crash was caused by a terrorist attack), but is later retracted. Memory for the details of the event and its explanation are then probed. A common finding is that the original beliefs about the cause of the event persist in memory when the retraction involves a denial of the original explanation (see Ecker et al., 2022; Walter & Tukachinsky, 2020, for reviews). This continued influence can extend to inferences about the event that were implied but not were not included in the original description (Ecker et al., 2011). Crucially, however, the effects of the original explanation are greatly reduced when the retraction contains a plausible causal alternative (e.g., in the plane crash scenario, it is subsequently found that the plane had a faulty fuel tank) (Ecker et al., 2011; Ecker, Lewandowsky, Cheung, & Maybery, 2015; Johnson & Seifert, 1994; Kan, Pizzonia, Drummey, & Mikkelsen, 2021; Rich & Zaragoza, 2020).

1.3. The current studies

This library of previous work suggests that people may be able to re-evaluate previously observed evidence in the light of new *causal* explanations of the data generation process. That is, they may be able to modify their property inferences to take account of information about the sampling process that is encountered after the sample data have been observed. The current studies aimed to test this hypothesis using the sampling frames paradigm.

The first step involved a re-examination of whether people can apply sampling frames retrospectively (i.e., after the sample data has been observed) when the situation incorporates real-world mechanistic or category knowledge. As previously noted, Ransom et al. (2022) found that sampling frames that were presented after the sample data had a much smaller effect than those presented before. However, they used relatively abstract stimulus materials whose category structure was unfamiliar to participants before the experiments: in one study, for example, the sample was made up of "small rocks discovered on the planet Sodor", and participants were subsequently asked to make inferences about rocks of other sizes from that planet. Hence, it is possible that participants' attention was focused primarily on learning about the sample rather than reasoning about existing knowledge based on the sampling frame.

Experiment 1 in the current series, therefore, re-examined the effects of presenting frames before and after the sample data using causal explanations of sample selection applied to more familiar (animal) categories. We found that people used sampling frames to guide property inferences even when they were presented *after* the sample of evidence was observed.

Experiments 2 and 3 addressed the more novel question of whether people could not only revise their inferences in light of additional information about the sampling frame, but go further and *retract a frame*

that had been presented previously and replace it with an alternative "correct" frame. The property inferences in these "frames-switch" conditions were compared to "no switch" conditions where the sampling frame remained consistent throughout learning. We predicted that if people were able to successfully shift between frames, then the patterns of generalization in the switch conditions should be based on the more recent *correct* frame rather than the original frame.

1.4. Why it is important to understand frame shifting

Determining whether people successfully shift between sampling frames is important for a number of reasons. Theoretically, this offers a further test of the encoding-only hypothesis of Ransom et al. (2022). Finding a successful retraction of an early frame – and subsequent shift towards inferences based on a later frame – would challenge this hypothesis, pointing to a role of sampling assumptions in both the encoding *and* retrieval of sample data.

This work will also inform further development of a formal model of inference. Previous findings concerning the effects of sampling frames on inductive inference are well-captured by a Bayesian computational model with biased samples (Hayes, Banner, et al., 2019; Ransom et al., 2022). Extending ideas originally proposed by Tenenbaum and Griffiths (2001), the model assumes that a learner's task is to decide among a set of hypotheses about the true extension of a property p (e.g., "only birds have p", "birds and mammals have p", "all animals have p"), based on a sample of exemplars from category x that have the property. As shown in Eq. 1, the posterior probability for each hypothesis h under a sampling frame s is a joint function of the base-rate probability of each hypothesis $P(h)$ and the likelihood of the data being observed given the hypothesis and the sampling frame. An important feature of this model is that the likelihood function varies depending on one's sampling assumptions. Different likelihood functions are applied for samples subject to a category frame or a property frame (see Hayes, Banner, et al., 2019; Ransom et al., 2022 for details). In other words, the way that a learner evaluates the probability of rival hypotheses about property generalization as they observe sample members, will change depending on the type of frame being applied.

$$P(h|\ \mathrm{x},\ s) \propto P(\mathrm{x}\ |h,\ s)P(h) \quad (1).$$

The model predicts the differences in property generalization observed under different sampling frames. It also generates several novel predictions (e.g., when instances outside the sampled category are rare, there will be less tightening of generalization under a property frame). These have been confirmed in empirical work (Hayes et al., 2023; Hayes, Banner, et al., 2019; Ransom et al., 2022).

This Bayesian model can be characterized as a computational or ideal-observer model of how generalization from samples of evidence should proceed (Marr, 1982). This model is agnostic about when and how model components, like calculating likelihoods based on relevant sampling assumptions, occur during the course of the generalization process. Clarifying the impact of sampling frames during data encoding and retrieval will contribute towards the development of a more detailed *process* model of how learners approach the task of inductive generalization (Tauber, Navarro, Perfors, & Steyvers, 2017).

More generally, our tests of switching between sampling frames have important implications for our ability to retrospectively revise our beliefs about sample data and its implications. If we find that participants are successful in switching frames, then this suggests that we can reduce the negative effects of exposure to biased samples by providing alternative explanations of the evidence-generation process after the evidence is observed ("debunking", see Ecker et al., 2022 for a review).

2. Experiment 1

This study re-examined the effects on inductive inference of sampling frames presented before or after a sample of evidence is observed. We used familiar animal categories in the observed sample and as inductive test items. Compared to the stimuli used by Ransom et al. (2022), these categories should be more familiar to participants as well as more coherent, in that participants will have more prior knowledge of the similarity of instances within a category and differences between categories (cf. Malt, 1995; Murphy & Medin, 1985).

All participants viewed the same sample of evidence (small birds with the novel property of "plaxium blood" – see Fig. 2A) and were asked to judge whether the property generalized to a range of other types of birds and animals. Sampling frames, in the form of causal explanations of the sample selection process, were presented before or after observation of the sample. In the category sampling conditions, instances were selected on the basis of category membership (i.e., only small birds could have been observed). In the property sampling conditions, instances were selected because they possessed the novel property (i.e., only things with plaxium could have been observed).

Previous empirical work (e.g., Lawson & Kalish, 2009), and the Bayesian model of inference (Hayes, Banner, et al., 2019; Ransom et al., 2022), predict less property generalization to items outside the sampled category under a property frame as compared with a category frame. The key question was whether this frames effect would be found in both the frames-before and frames-after conditions.

2.1. Method

2.1.1. Participants

Two hundred and eighty-one participants were recruited using the online research platform, Prolific (130 females, 147 males, 4 other; $M_{AGE} = 33.89$ years, age range: 18–75). Approximately equal numbers were randomly allocated to one of four groups: category-frames before, property frames-before, category-frames after, property frames-after. Participants were paid £1.00 on completion. No participants were excluded from this or any subsequent experiment.[1]

2.1.2. Procedure

The experiment was run online using jsPsych (de Leeuw, 2015) and conducted through the Prolific platform. Pictures of animals used during sampling and test phases were originally sourced from Google Images and modified using Adobe Photoshop to remove background features. All participants were told that they were acting as researchers exploring the wildlife of a little-known island. They were told that there were many types of animals on the island and shown pictorial examples. Their task was to use a sample of animals to infer which animals on the island had 'plaxium blood' and which did not.

The experiment commenced with a warm-up phase intended to familiarize participants with the task goals and general procedure. Participants were informed about a preliminary expedition which had collected two animals that had the novel property of plaxium blood. No information was provided about how the sample was collected in this phase. A comprehension test consisting of two multiple choice questions followed, with any incorrect responses returning participants to the initial instructions. Two sample instances (pictures of small birds) were then presented, one at a time. On each sampling trial, the participant clicked a dialogue box to view the sample instance with the text

[1] Sample sizes in the current experiments were based on several considerations. We followed the general design strategy for studies employing Bayesian data analysis suggested by Schönbrodt and Wagenmakers (2018). This involved collecting participant data until a desired level of conclusive evidence was found for the main effect of sampling frames (i.e., the difference between generalization ratings for category and property sampling). As suggested by Schönbrodt and Wagenmakers (2018), a threshold of a Bayes Factor of 6 or more for/against this effect was used. This general strategy, however, was moderated by the resources for payment of participants that were available at the time each study was run. Fewer funds were available at the time that Experiment 2 was run, resulting in smaller cell sizes.

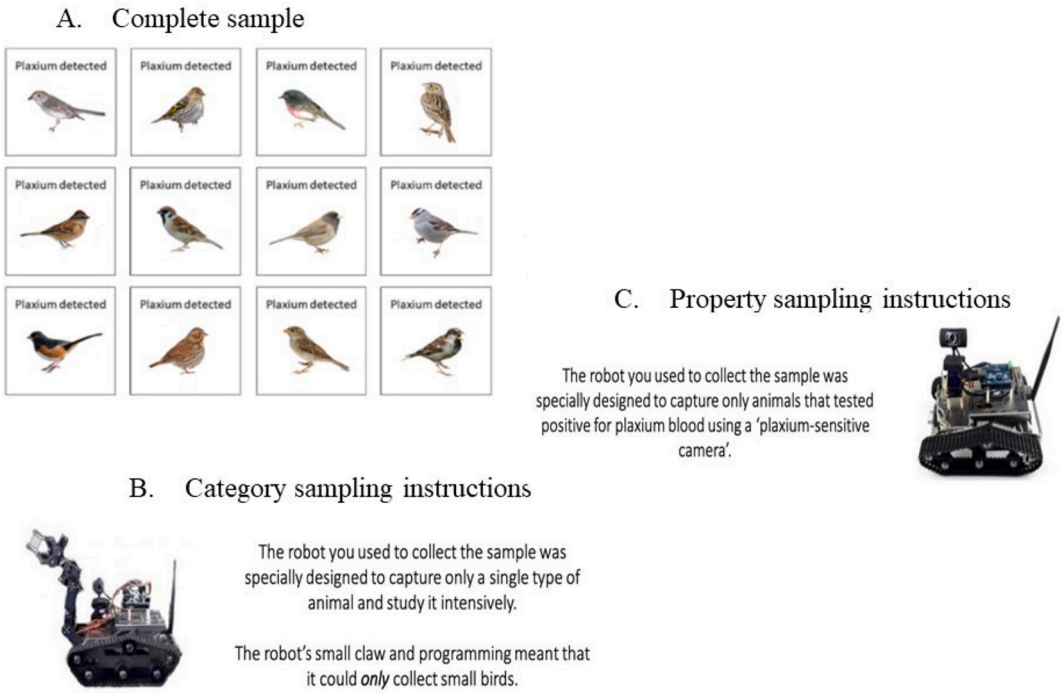

Fig. 2. The sample used in all conditions (A), category sampling instructions (B) and property sampling (C) instructions.
Note: Frame instructions were presented either between the warm-up and main sampling phase (frame before) or after the complete sample had been observed (frame after).

'plaxium detected' displayed above it. Participants could view each sample instance for as long as they wanted, with a 2 s delay between each instance. After viewing the sample, a warm-up generalization test was presented where participants were asked to rate the likelihood of each of six test animals having plaxium blood. The test items varied in similarity to the sample category of small birds (in decreasing order, *sparrow, pigeon, owl, emu, mouse, lizard*).[2] Generalization ratings were made using an on-screen 10-point slider scale (1 = "very unlikely" and 10 = "very likely"). Test items appeared one at a time in a random order. Based on previous work (e.g., Hayes, Banner, et al., 2019), no differences in generalization patterns between the experimental groups were expected after observing the small warm-up sample.

After the warm-up, participants in the frames-before condition were given sampling frames instructions before proceeding to the main sampling phase. In the category-frame condition, it was explained that a subsequent sample would be collected by a robot with a small claw which meant they could only collect small birds. They were also told about an alternative robot with a plaxium-sensitive camera, but that this robot *would not* be used to collect the sample. In the property-frame condition, these instructions were reversed. They were told that the sample would be collected by a robot equipped with a 'plaxium-sensitive camera', which meant that it would only collect things with plaxium (see Fig. 2 for examples of frames instructions). Those in the frames-after condition were presented with the same sampling frames instructions, but these were presented *after* the main sampling phase. The wording of the instructions was the same for the frames-before and frames-after conditions, except for the use of different tenses (future for frames before; past for frames after). We included the description of both robots in the sampling frames instructions in the interests of consistency with instructions used in later retraction studies. Participant understanding of the frames instructions was checked using a multiple-choice quiz administered immediately after these instructions. If any question in the quiz was answered incorrectly the correct instructions were repeated.

In the main sampling phase, participants were presented with another ten sample instances of small birds, in addition to the two small birds from the warm-up phase. These sampling trials followed the same procedure as the warm-up. Pictures remained on screen as subsequent sample instances were presented so that, by the end of this phase, twelve "plaxium positive" small birds were present (see Fig. 2A).

The final property generalization test phase was presented either immediately after the completion of the main sample phase (frames before) or after completion of this phase and presentation of the sampling frames instructions (frames after). Each of six generalization test items appeared in random order and participants rated the likelihood of plaxium blood being found in each. The generalization test items were the same as those used in the warm-up phase.

2.2. Results and discussion

Bayesian approaches were used to analyze the data, carried out with the JASP v0.17 package (JASP Team, 2023). Such analyses have the advantage of quantifying the strength of evidence both in favor of differences between conditions, and of null effects. The Bayes Factor (*BF*) comparing two hypotheses is a ratio that expresses the relative probability of observing the data under one hypothesis compared to another hypothesis. We use the notation BF_{10} to refer to Bayes Factors that support the alternative hypothesis and BF_{01} to refer to Bayes Factors that support the null hypothesis. For example, $BF_{10} = 10$ indicates the data is 10 times more likely to have come from the alternative hypothesis than the null hypothesis. Conventionally, a *BF* between 3 and 20 represents positive evidence for the alternative (or null) hypotheses respectively, a *BF* between 21 and 150 represents strong evidence, and a *BF* above 150 represents very strong evidence (Kass & Raftery, 1995).

[2] This rank ordering of the similarity of the test items to the sample category was established in previous work where participants gave pairwise similarity ratings for these stimuli (Hayes, Banner, et al., 2019).

Property generalization ratings were analyzed with Bayesian analyses of variance, using Cauchy default priors ($r = 0.707$; Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017).[3] Recall that no sampling frames information was presented in the warm-up phase, and participants were exposed to minimal sampling data (two small birds). Our analysis revealed that, in this phase, people made property inferences based on the similarity of test items to the sample, with more generalization to items with high similarity (see https://osf.io/qjgxt/ for data), $BF_{10} > 150$. However, as expected for this phase, there was positive evidence against an impact of sampling frames on these inferences, $BF_{01} = 16.129$.

The more notable results concern property generalization ratings in the final test phase, which are shown in Fig. 3. These data were analyzed using a 2 (sampling frames condition) x 2 (frame timing) x 6 (test items) Bayesian mixed-model analysis of variance with repeated measures on the last factor. We computed Bayes Factors comparing the likelihood of models that included each of the main effects of the frames condition, frame timing and test items, and their two-way and three-way interactions, to equivalent models that omitted one of these effects. The analysis found very strong evidence of a main effect of the similarity of the test item to the observed sample, with more similar items receiving higher generalization ratings, all $BF_{10} > 150$). The analysis also found very strong evidence of an overall effect of sampling frames, with narrower generalization to test items following property sampling as compared to category sampling, $BF_{10} = 3610.03$ (category frame: $M = 5.149$; posterior 95% credible intervals 4.968–5.532; property frame: $M = 4.381$, posterior 95% credible intervals 4.202–4.556). There was also positive evidence that this effect interacted with test items, $BF_{10} = 16.537$. Fig. 3 shows that the effect of sampling frames on property inferences was most pronounced for the most novel test items (i.e., those with lower similarity to the observed sample).

As suggested by the Figure, however, frame timing had little impact on these effects. There was weak evidence against an interaction between the frame condition and frame timing, $BF_{01} = 2.817$, and positive evidence against a three-way interaction of frame condition, timing and test item, $BF_{01} = 29.412$.

This experiment re-examined the effects of the timing of the presentation of sampling frames on the inferences that people draw from a sample of category members that share a novel property. Like many previous studies (Hayes et al., 2023; Hayes, Banner, et al., 2019; Lawson & Kalish, 2009), we found that participants factored sampling frames into their inferences when the frames were presented before the sample was observed. Generalization of a property to novel items was narrower under property than category sampling. Contrary to Ransom et al. (2022), however, we found a similar impact of sampling frames on inference when the frames were presented *after* the sample. This indicates that people can apply frames retrospectively to samples of evidence that they observe and grasp their implications for property inference.

Our more positive findings regarding the retrospective application of frames are most likely due to our use of more familiar categories during the sampling and generalization test phases. This may have made it easier for participants to identify the animal category to which the sample members belonged. In turn, this may have facilitated understanding of the inferential implications of the subsequent frame (e.g., in property sampling, that the absence of members of other categories was highly informative about how far the property generalizes).

3. Experiment 2

Having established that it is possible for people to retrospectively apply sampling assumptions when making property inferences, we now turn to the question of whether people can *update* these assumptions – replacing beliefs about how a sample was generated with a new set of beliefs. We examined whether people are able to make property inferences based on a new sampling frame that is presented after the sample data and contradicts an earlier frame. Learners were tasked with inferring the extension of a novel property ("plaxium blood") after observing a sample of items (small birds) with that property. In no-switch conditions, a category or property sampling frame was presented before the sample data and was never retracted (similar to the frames-before conditions in Experiment 1). In the switch conditions (category-to-property and property-to-category groups), an initial sampling frame was presented, but this frame was retracted and replaced with a new frame after the sample data. The crucial question was whether property generalization in the switch groups was based on the initial frame or the more recent, retracted frame. If participants can successfully retract and replace the old with the new frame, then property generalization in the category-to-property group should resemble that in the property-only group, and generalization in the property-to-category group should resemble that in the category-only group.

3.1. Method

3.1.1. Participants

Two hundred participants (98 females, 100 males, 2 other, $M_{AGE} = 35.72$ years, age range: 18–74) were recruited using the online research platform, Prolific. Participants were paid £1.00 on completion. Equal numbers were randomly allocated to four groups: category-only (no switch), property-only (no switch), category-to-property, property-to-category.

3.1.2. Procedure

The same stimulus materials were used for sampling phases and generalization tests as in Experiment 1. The experiment proceeded through five phases: 1) warm-up, 2) initial sampling frames instructions, 3) main sampling, 4) second frames instructions, 5) property generalization test. Fig. 4 summarizes the experimental protocol.

The warm-up phase was identical to that in Experiment 1. After the warm-up, participants were provided with initial frames instructions that were similar to those used in the frames-before condition of the previous experiment. All groups were introduced to two robots, each with a function consistent with either category sampling or property sampling. All participants were told about each type of robot, but were informed that subsequent samples would be collected by only one type of robot (see Appendix for verbatim instructions). Those in the category-only and category-to-property conditions were initially told that their sample would be collected by the robot that could only sample small birds. Those in the property-only and property-to-category conditions were told that their sample would be collected by the robot that could only sample things with plaxium. A multiple-choice comprehension test checked participants' understanding of the task, with incorrect responses again returning participants to the relevant instructions.[4]

The main sampling phase followed in which participants were presented with another ten sample instances of small birds. After the sample had been observed a second sampling frames instructions phase was presented. Those in the no switch conditions were reminded about the

[3] Robustness tests confirmed that the main results reported for this and subsequent experiments were maintained when wider Cauchy priors were used (see https://osf.io/qjgxt/).

[4] An exploratory "no frames" baseline condition was also included in this experiment ($n = 50$). This group was given no specific information about the way that sample instances were collected. The aim was to examine whether peoples' default beliefs about the sampling process were closer to category or property sampling. The results, however were inconclusive with weak evidence for the null hypothesis when generalization ratings in this condition were compared to the category-only condition, $BF_{01} = 2.155$, and the property only condition, $BF_{01} = 5.195$ (see https://osf.io/qjgxt/ for full data and analyses).

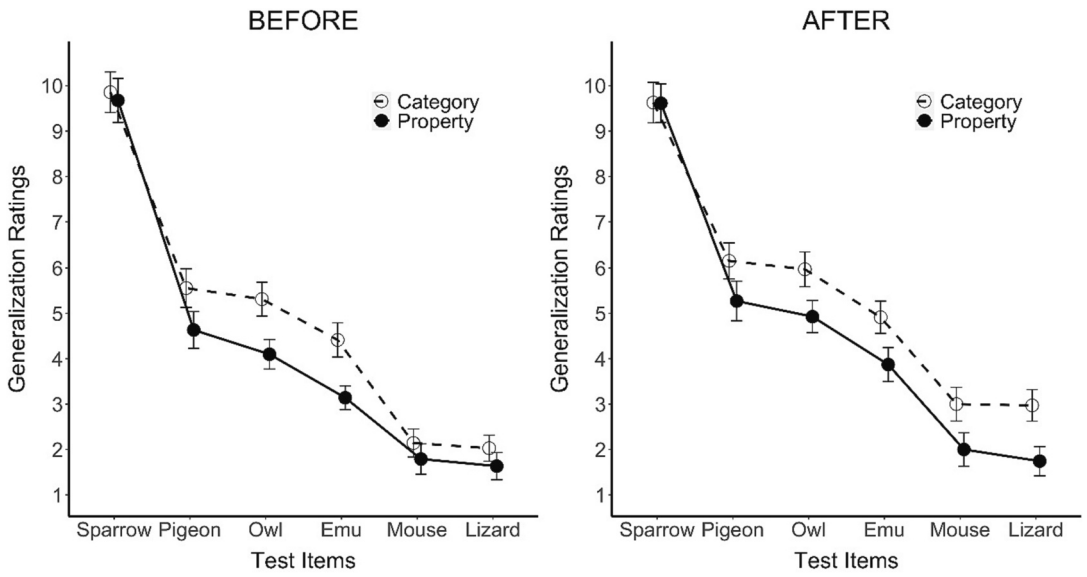

Fig. 3. Experiment 1. Mean test phase property generalization ratings in the frames before and after conditions.
Note: In this experiment, category or property sampling frames were presented before or after a sample of small birds with plaxium. The Figure shows subsequent property generalization ratings to items from the same category (sparrow) and novel categories varying in similarity to the sample. There was less generalization to novel categories under property sampling than under category sampling in both timing conditions. 95% confidence intervals are shown for all means.

relevant robot used to collect the sample. Those in the switch conditions were notified of a 'mix-up' in the information they received about the robot responsible for the sample data (see Appendix for details). They were told that the alternative robot was the actual source of the observed sample instances. They were told to disregard their previous conclusions and use this new sampling frame to inform their inductive inferences (see Fig. 5 for screenshots of the retraction instructions).

To help them grasp the implications of the new frame, participants were presented with a brief 'reminder' of the sample showing the images of the twelve sample instances with the correct robot beside them for 10 s. Those in the no switch conditions did not receive additional frame instructions, but were provided with the reminder of the sample (without a robot picture). A further multiple-choice comprehension test was given to all participants to confirm their understanding of the designated sampling frame. If any response was incorrect, the correct instructions were repeated. In the final property generalization test, each of the six test items appeared in random order and participants rated the likelihood of plaxium blood being found in each.

After the conclusion of the experiment, as a manipulation check, those in the switch conditions were asked to rate the plausibility of the frames retraction on a 3-point scale (1 = "not at all plausible/believable", 2 = "moderately plausible/believable", 3 = "very plausible/believable").

3.2. Results and discussion

Property generalization ratings were again analyzed with Bayesian analysis of variance using Cauchy default priors. A preliminary analysis confirmed that, in the warm-up phase, property generalization ratings were influenced by the similarity of the test item to the sample, $BF_{10} > 150$, but there was positive evidence against an effect of sampling frames, $BF_{01} = 3.61$.

Property generalization ratings in the final test are shown in Fig. 6. These data were analyzed using a 4 (sampling frames condition) x 6 (test items) Bayesian analysis of variance with repeated measures on the second factor. We computed Bayes Factors comparing the likelihood of models that included each of the main effects of frames condition and test item, and their two-way interactions to equivalent models that omitted one of these effects. The analysis found very strong evidence of a main effect of the similarity of the test item to the observed sample, with more similar test items receiving higher generalization ratings, all $BF_{10} > 150$.

The more interesting questions concerned differences in ratings between sampling frames groups, which were examined through a series of planned comparisons, as well as possible interactions between these group effects with the test item factor. A comparison of ratings in the category-only and property-only frame conditions found positive evidence for a main effect of category frame, such that generalization ratings were generally lower under property framing (property-only: $M = 4.655$, posterior 95% credible intervals = 4.404–4.897) than category framing (category-only: $M = 5.317$, posterior 95% credible intervals = 5.064–5.549), $BF_{10} = 6.658$. This again replicates the frames effect found in the frames-before condition of Experiment 1 and previous studies (Hayes, Navarro, et al., 2019; Lawson & Kalish, 2009; Ransom et al., 2022). There was evidence against an interaction between this frames effect and the test item factor, $BF_{01} = 7.464$.

We next compared ratings in the category-to-property condition with those in the category-only and property-only conditions. There was positive evidence of a main effect such that generalization ratings in the category-to-property condition ($M = 4.587$, posterior 95% credible intervals = 4.309–4.842) were lower than those in the category only condition, $BF_{10} = 6.233$. As suggested by Fig. 6, there was also positive evidence that this effect interacted with test item, $BF_{10} = 5.168$. In contrast, there was positive evidence of no difference between generalization ratings in the category-to-property and property-only conditions, $BF_{01} = 6.901$. Hence, inductive generalization in the category-to-property condition was based on the new property frame rather than the initial category frame.

A complementary pattern was found in comparisons between ratings in the property-to-category condition and those in category-only and property-only groups. As shown in Fig. 6, there was positive evidence that generalization ratings in the property-to-category condition ($M = 5.427$, posterior 95% credible intervals = 5.159–5.672) were higher than those in the property-only condition, $BF_{10} = 16.219$. This effect interacted with test item, such that group differences in ratings were larger for test items with the lowest similarity to the sample, $BF_{10} = 42.587$. In contrast, generalization ratings in the property-to-category condition did not differ from those in the category-only condition,

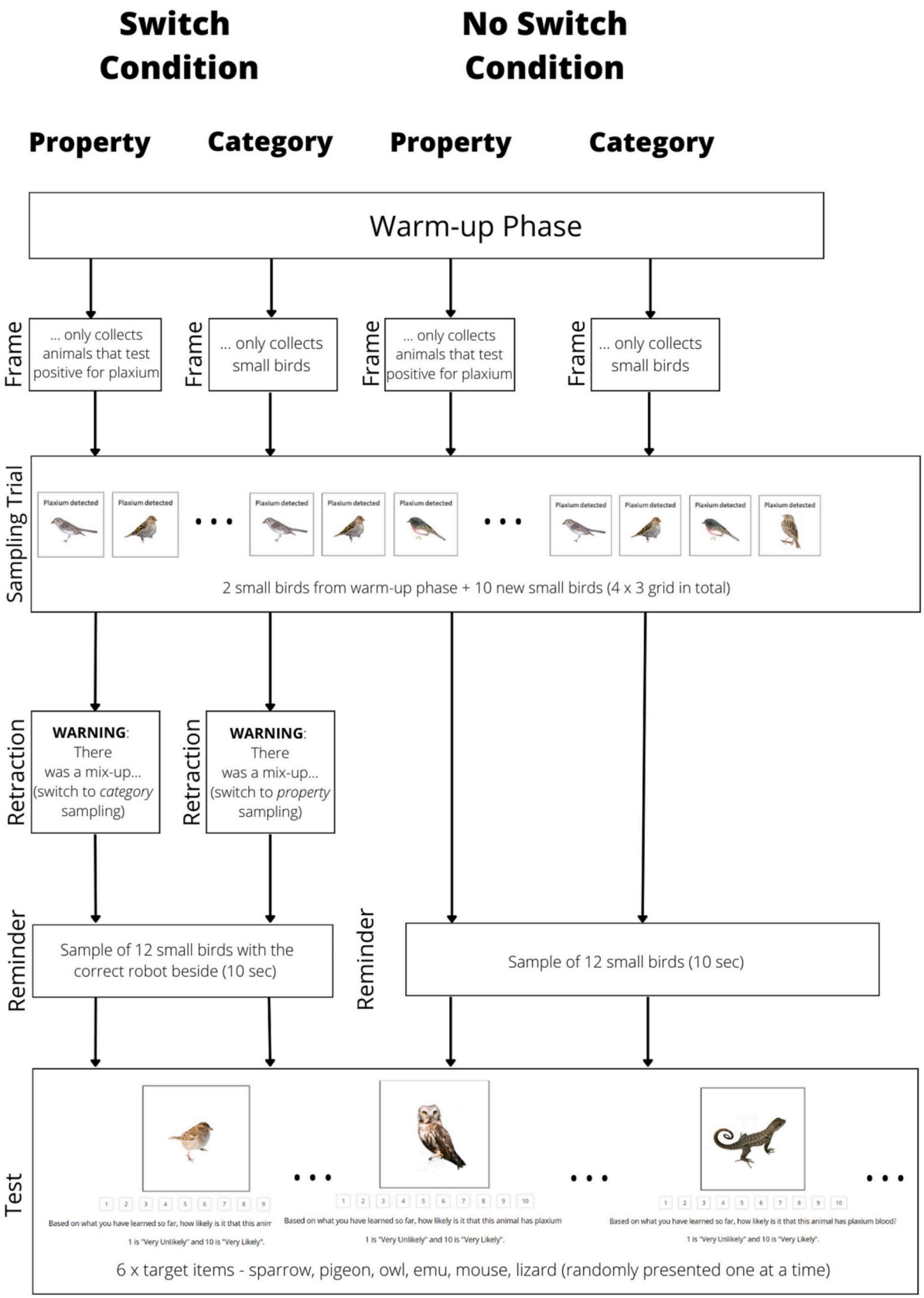

Fig. 4. Timeline summary of Experiment 2.

Fig. 5. Experiment 2. Frames retraction instructions for the switch conditions.

BF_{01} = 6.675. Generalization in the property-to-category condition appeared to be based on the new category frame rather than the initial property frame.

3.2.1. Manipulation check

The manipulation-check on the plausibility of the retracted frames explanations found that most participants in both the category-to-property condition (M = 2.2, SEM = 0.094) and the property-to-category condition (M = 2.1, SEM = 0.096), found the retractions to be at least moderately plausible, and that there was positive evidence of no difference in ratings between these groups, BF_{01} = 4.744.

Experiment 2 examined whether people presented with a sampling frame before they observed a sample of evidence could shift to a new frame when the previous frame was said to be incorrect and replaced with an alternative. We found positive evidence of such switching: those presented with a property frame, which was subsequently retracted and replaced by a category frame, showed patterns of generalization that were more in line with Bayesian predictions about category sampling. Likewise, those presented with an initial category frame successfully shifted to a property frame.

These data may seem surprising in the light of previous work that has highlighted the persistent effects of prior beliefs on belief updating in the light of new evidence (e.g., Hogarth & Einhorn, 1992). They also contrast with much of the work demonstrating the continued influence of initial beliefs on memory about the causes of events (e.g., Ecker et al., 2022). It is important to note, however, that in the current studies we replaced an initial *causal* explanation of the sampling process with an alternative causal explanation. Work on the continued influence effect has shown that such causal reinterpretations of witnessed evidence are most likely to attenuate the effect (e.g., Ecker et al., 2011; Ecker et al., 2015; Rich & Zaragoza, 2020). We also note that the current work goes beyond previous demonstrations of successful retraction of causal information in the continued influence effect. In the CIE, new causal explanations are used to reinterpret old evidence. In the current work, the new frame presented after the sample led to a change in the way property inferences were made about members of categories that were not part of the sample of evidence.

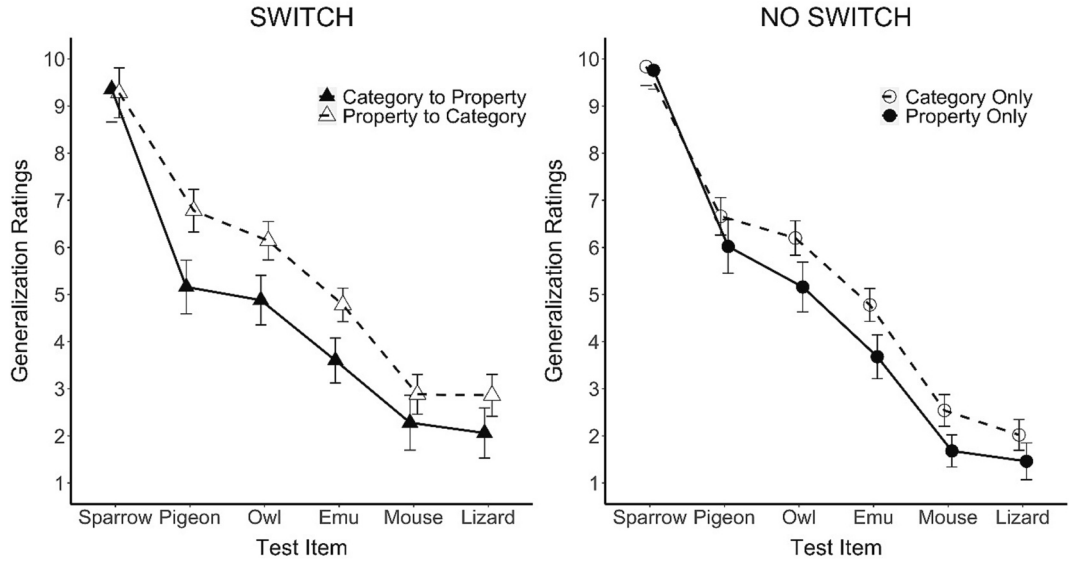

Fig. 6. Experiment 2. Mean test phase property generalization ratings in the switch and no switch conditions.
Note: In the switch conditions, participants were presented with either a category or property frame before observing a sample of small birds with plaxium, but this was retracted and replaced with the alternative frame after the sample. In the no switch condition, there was no retraction of the original frame. Generalization to novel items in the switch conditions reflected updating of sampling assumptions in line with the more recent, alternative frame, i.e., generalization in the category to property condition was similar to that in the property only condition; generalization in the property to category condition was similar to that in the category only condition. 95% confidence intervals are shown for all means.

One aspect of our belief revision procedure that deserves consideration is the use of an explicit instruction in the switch conditions to "disregard your previous conclusions and use … new information". We acknowledge that such explicit instructions to replace old with new frames are likely to be rare in everyday learning situations. In our earlier investor and holiday planning examples, a new frame is unlikely to be accompanied by an explicit denial of the old frame. We note, however, that evidence from studies of the continued influence effect shows that mere negation or denial of an initial explanation of an event is unlikely to eliminate the effect of that explanation on subsequent memory (e.g., Ecker et al., 2011; Johnson & Seifert, 1994). Hence, the presence of our "disregard previous conclusions" instructions is unlikely to be the key factor that led to the successful frame switching observed in this study.

The results from this experiment suggest that people show considerable flexibility in the way they update and apply their beliefs about sample selection when asked to make sample-based inferences. It should be acknowledged, however, that there are other aspects of the Experiment 2 procedure that may have facilitated this process. In particular, all groups, including the crucial switch conditions, were provided with a brief reminder of the sample data when given the frame. This reminder could have provided the learner with an opportunity to re-encode the sample data in the light of the new frame that had just been presented. In other words, we cannot be certain that those in the switch conditions were applying the new frame *retrospectively* to previously encoded sample data.

An additional factor that may have facilitated the shift between frames was the presentation of *both* sampling frames before the sample was observed. The way that each of robots selected samples – based on category membership or having the property of interest, was explained to all participants in the initial frame stage. They were then told which of the robots selected the sample that they were going to observe. Knowing about the sampling alternatives from the outset may have made it easier for those in the switch conditions to subsequently adopt the alternative frame.

Previous work on "knowledge restructuring" in category learning (e.g., Kalish, Lewandowsky, & Davies, 2005), for example, has shown that in order for people to switch from a simple category rule that leads to modest categorization accuracy to an optimal but more complex rule, they need advance knowledge of that rule. Other work on belief revision in scientific reasoning also suggests that having prior knowledge of alternative theories makes it easier to re-interpret new evidence that is ambiguous or is inconsistent with a given theory (Chinn & Brewer, 1998; Ganea, Larsen, & Venkadasalam, 2021).

4. Experiment 3

Experiment 3 re-examined whether people can retract and replace a previously presented sampling frame with a new frame presented after the sample data were observed. In this case, we removed two of the procedural features that may have facilitated switching between frames in Experiment 2. In this study, no reminder about the sample composition was provided when the new frame was introduced. Moreover, half the participants were only told about a single frame before the sample data was presented. These participants were unaware of the alternative method of sampling before the data was observed. Those in the switch conditions were only informed about the alternative frame after seeing the data.

This study was originally conceived of and run as two separate experiments, that successively removed features from the Experiment 2 design (i.e., the first study removed the reminder about the observed sample, the second removed this feature as well as the initial explanation of the alternative frame). However, the designs were so similar we decided to combine the data from these studies. Whether an early explanation of the alternative sampling frame was presented was included as a factor in the data analyses. Combining the data from these studies increased the likelihood that our Bayesian analyses would yield *conclusive* evidence for or against the effects of interest (Schönbrodt & Wagenmakers, 2018).

4.1. Method

4.1.1. Participants

A total of five hundred and sixty participants (240 females, 301 males, 19 non-binary/other, M_{AGE} = 37.93 years, age range: 18–76 years) were recruited using Prolific ($n = 280$ in each of two experimental runs). Approximately equal numbers were allocated to the four experimental conditions. There were no exclusions.

4.1.2. Procedure

The procedure was very similar to Experiment 2, except that we removed the brief reminder of the sample that was presented before the final generalization test. For half the participants, we also modified the instructions for all frames conditions so that, at the initial frame stage, only a single type of frame mechanism (category or property) was described. In the switch conditions for these participants, the alternative sampling frame was only introduced at the retraction stage.

4.2. Results and discussion

A preliminary analysis again confirmed that there were no differences between sampling frames conditions in property generalization ratings during the warm-up phase, BF_{01} = 13.157. Property generalization ratings in the final test phase are shown in Fig. 7. These data were again analyzed using a Bayesian analysis of variance with the same design as in Experiment 2, except that we added a factor that reflected whether or not initial instructions about the alternative frame were included.

We again found very strong evidence of an effect of test item similarity to the sample on property generalization, $BF_{10} > 150$. Comparing generalization ratings in the property-only and category-only conditions, there was very strong evidence of a frames effect, BF_{10} = 6613.502, with narrower generalization under a property frame (M = 4.649; 95% posterior credible interval: 4.464–4.813) than under a category frame (M = 5.393; posterior 95% credible interval: 5.212–5.564). There was also positive evidence that this effect interacted with test item similarity, so that effect of the sampling frames was larger for inferences about the more novel test items (i.e., those that were less similar to the training sample), BF_{10} = 19.98.

Fig. 7 shows that generalization ratings in the category-to-property condition were lower than those in the category-only condition, (M = 4.699; 95% posterior credible interval: 4.513–4. 863). There was very strong evidence for this difference, BF_{10} = 165.157. There was inconclusive evidence of an interaction between this retraction effect and test item similarity, BF_{10} = 0.543. Generalization ratings in the category-to-property condition did not differ from those in the property-only condition, BF_{01} = 22.393. This indicates a shift towards generalization based on the final rather than the initial frame.

Evidence of a shift from the initial to the final frame was also found in comparisons of the property-to-category condition with the category-only and property-only groups. Generalization ratings in the property-to-category condition (M = 5.525; 95% posterior credible interval: 5.341–5.69) were higher than those in the property-only condition, with very strong evidence for this difference $BF_{10} > 150$. This effect interacted with test item, such that group differences were only found for novel test items, not for the small bird item, $BF_{10} > 150$. There was a trend for generalization ratings in the property-to-category condition to be higher than in the category-only condition, but the evidence for this difference was weak, BF_{10} = 2.933.

There was no evidence that any of the effects reported above were affected by whether an explanation of the alternative sampling frame was included prior to the sampling phase. The strength of the evidence for a null effect of this factor as a main effect or an interaction, ranged from positive, BF_{01} = 4.049 to very strong, $BF_{01} > 150$.

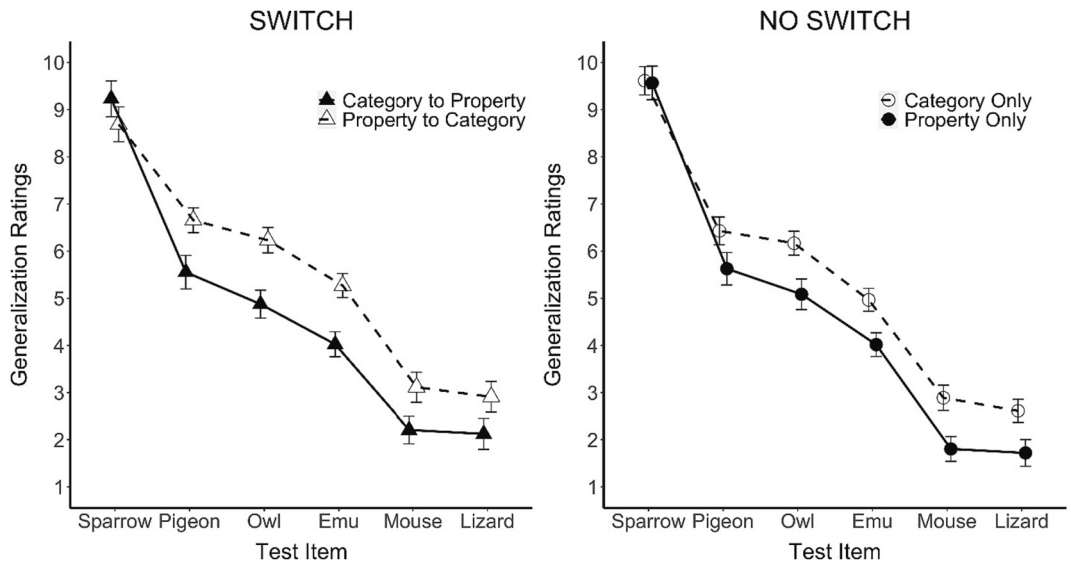

Fig. 7. Experiment 3. Mean test phase property generalization ratings in the switch and no switch conditions.
Note: As in the previous study, generalization to novel items in the switch conditions reflected updating of sampling assumptions in line with the more recent, alternative frame, i.e., generalization in the category to property condition was similar to that in the property only condition; generalization in the property to category condition was similar to that in the category only condition. 95% confidence intervals are shown for all means.

In sum, we again found clear evidence of successful retraction and replacement of an earlier frame with a later frame. This effect persisted when participants were not reminded about the sample contents before making generalization judgments, and when they were unaware of the possibility of an alternative frame prior to sampling.

5. General discussion

A considerable body of work supports the view that inductive reasoning involves consideration of both the contents of an evidence sample *and* assumptions about how that sample was generated (Hayes, Banner, et al., 2019; Hayes, Navarro, et al., 2019; Ransom et al., 2016; Ransom et al., 2018; Ransom et al., 2022). The current studies reaffirmed this general finding – showing that people were sensitive to constraints on the sample selection process when making sample-based inductive inferences. When sample members were selected because they had the property of interest (property sampling) and were found to belong to a single category, there was little property generalization to other categories. In contrast, when sampling was restricted to members from a single category (category sampling), people were more likely to infer that the property of interest may be present in other categories.

Crucially, the current work also examined the novel question of whether people could *revise* their sampling assumptions when making property inferences – shifting from the sampling frame that was in place before the sample data was observed to a new frame introduced after sampling. Experiment 1 established that people were capable of applying sampling frames to their inferences regardless of whether these were presented before or after sampling. Experiments 2 and 3 found strong evidence that people can revise their beliefs about the sampling process when initial beliefs were retracted. In both of these studies, people were successful in switching from an initial property frame to a category frame – shifting from narrow generalization of a novel property to broader generalization. Likewise, people also successfully switched from a category to a property frame – shifting from broad to narrow generalization.

These studies reinforce the crucial role of people's beliefs about the sampling process in how they draw inferences from a sample of evidence. A key novel contribution, however, is highlighting how flexibly these beliefs can be applied. People were sensitive to retractions of old frames which were found to be incorrect, replacing these with new frames presented after the sample was observed. These findings represent a significant advance on previous work, where information about the sampling process has typically only been presented before the sample remains unchanged over the course of learning.

Contrary to the findings of Ransom et al. (2022), the current work suggests that sampling assumptions can affect the *retrieval of previously observed sample data,* as well as the way that it is encoded. In Experiment 1, frames that were only learned about after the sample data had a similar impact on property inferences to those observed before the data. This suggests that, in the frames-after case, the sample data were stored in "raw" form (i.e., without a specific frame being applied) and then retrieved from memory when the frame was presented and final generalization judgments were required. A similar process may have operated in Experiments 2 and 3, with the sample data stored separately from the initial sampling frame. The frame was then only applied to the sample for final generalization judgments. In the switch conditions, this was the more recently presented frame. An alternative possibility is that, in the switch conditions, the initial frame may have affected the way the sample was encoded and represented, but this representation was revised in the light of the new frame. Differentiating between these alternative accounts of how sample data are represented and combined with sampling frames information is an important goal for future work.

Our findings concerning the impact of causal explanations of the sampling process on property generalization are in line with previous accounts that have highlighted the important role of causal knowledge in inductive reasoning (e.g., Hayes & Thompson, 2007; Medin, Coley, Storms, & Hayes, 2003; Rehder, 2017). What is novel about the current work is that it shows that people can switch between different causal explanations of sample generation, with appropriate shifts in inferences based on the sample.

More broadly, our findings are consistent with previous demonstrations of updating of beliefs in domains such as narrative comprehension (e.g., Kendeou et al., 2019) and forensic decision-making (e.g., Lagnado & Harvey, 2008). In line with this work, we have shown that people can re-interpret sample evidence based on information received after the sample is observed. We believe that a key element in such updating is the provision of a new explanation of data generation that is just as, or more coherent, than a previous explanation (cf. Thagard, 2008). In this regard, our findings of a successful switch between sampling frames could be seen as analogous to conceptual change from an initial explanatory

theory of the observed data to a new theory (e.g., Carey, 2009; Chinn & Brewer, 1998; Vosniadou, 2013).

5.1. Implications for formal models of inductive inference

The Bayesian model developed by Hayes, Banner, et al. (2019), explains property inference by calculating the conditional probabilities of different hypotheses about how far a property generalizes, given the sample data. Property and category sampling frames are implemented via different likelihood functions. The model makes no commitment to when these likelihood functions are evaluated. In that respect, the model is compatible with both the Ransom et al. (2022) results and the current findings.

If the ultimate goal, however, is to develop a process model of inference, then we need to incorporate assumptions about when and how sampling frames affect the way that sample data is represented and used. One possible direction is to formalize models of the relationships between sampling frames and sample data in terms of causal Bayes nets. This approach can capture and model inferences from samples that are subject to different types of constraints on sample selection, including sampling frames (Bareinboim, Tian, & Pearl, 2014; Kemp, Navarro, & Hayes, 2023). It may be possible to add assumptions to these Bayes nets about the temporal relations between the frames and the sample (e.g., whether the frame constraints are known about before or after the sample data is presented), and then test the predictions of the resulting Bayes net models against our inference data.

Although Bayesian approaches have most often been used to model the impact of sampling assumptions on property inferences, it is worth considering whether other models of inference could be extended to capture these effects. One well-known account, the Similarity-Coverage (Sim-Cov) model (Osherson et al., 1990), assumes that property inference is driven by two complementary components; i) the similarity between sample instances and targets, and, ii) coverage or the similarity between the sample instances and members of a higher-level category that includes both the sample and targets. The relative weight given to each component is determined by a free parameter α. The narrower generalization associated with a property frame compared to a category frame could be captured by increasing the value of this parameter, giving more weight to sample-target similarity. Shifts between frames presented before and after the sample data could likewise be captured by a change in a relative weighting of similarity and coverage. Note, however, that Sim-Cov model does not offer any principled explanation for *why* such shifts should take place.

Another issue that remains open to further investigation is how people represent the knowledge that they acquire from observing samples subject to different types of frames. The Bayesian model of biased sampling (Hayes, Banner, et al., 2019), makes no explicit representational assumptions. It is possible that some people may represent this knowledge as a summary rule. For example, under property sampling, after observing a sample like those used in the current experiments one might infer that "only small birds have this property". Notably, however, our results show that, if people do form summary rules, they are not static – they are readily revised and replaced with new rules in line with changes in beliefs about sample generation.

If people do form summary rules after observing the data, then it might be argued (as did a reviewer) that there is an alternative "anchor and adjust" account of our results. The suggestion is that learners' inferences are anchored to a static rule that is adjusted incrementally as new sample data are encountered (cf. Hogarth & Einhorn, 1992). However, it is hard to see how an anchor and adjust approach alone, with no mental model of sampling, could explain how and why people who have initially anchored on a given rule (e.g., "only birds have this property" following property sampling) then shift to a different rule ("it's possible that both birds and non-birds may have the property") when an alternative category frame is presented. This shift, observed in Experiments 2 and 3, implies that people encode both a) how the sample was generated, and b) the implications of this generative process for the observed sample, and also use both sources of information to evaluate rival hypotheses about far a property generalizes. Without all of these components, they would not *change* their generalization in the way we observe.

Moreover, the anchor and adjust account appears inconsistent with a body of previous research which shows that whether or not people adjust their inferences as they observe more data depends on their sampling assumptions (Hayes et al., 2023; Hayes, Banner, et al., 2019; Ransom et al., 2022). For example, under property sampling, people are more likely to believe that the property does not generalize beyond the observed category as more category members that have the property are observed (i.e., as the sample size increases). In contrast, increasing the size of the observed sample has little impact on the inferences made under category sampling.

5.2. Adjusting for biased sampling

The current work highlights how flexible people can be when using information about sample generation processes to make inferences from sample data. These results challenge a strong version of the metacognitive myopia argument (Fiedler, Prager, & McCaughey, 2023), which holds that people generally neglect sample generation when making judgments and decisions. That said, it is clear that there are many situations in which people do fail to factor biased sample selection into their inferences (e.g., Feiler, Tong, & Larrick, 2013; Fiedler, Brinkmann, Betsch, & Wild, 2000). In explaining this discrepancy, we should be mindful of the conditions under which our participants exhibited sensitivity to sampling frames. In each study, all participants saw the same sample data, and were given detailed causal information about different sample generation processes. Our results suggest that people are able to draw sensible inferences from causal explanations of sample selection applied to familiar categories and are capable of updating their inferences when new causal explanations replace old ones. Notably though, observation of the sample was passive – participants had no control over sample selection. In contrast, when participants have control over the sampling process, and use a sampling strategy that leads to a biased or unrepresentative sample (e.g., Fiedler et al., 2000; Le Mens & Denrell, 2011), they struggle to correct for these biases. In other words, learners appear to understand the implications of sampling biases that have been applied to an existing data set. But when the biases in the sample arise from their own sampling decisions, learners find it harder to adjust their inferences.

Another possible boundary condition on sensitivity to selection biases relates to sample composition. In the current studies, the sample was always made up of discrete members of a familiar category (i.e., each was a unique picture from the category of "small birds" that was easy to distinguish from other pictures). Previous work suggests that learners find it easier to draw inferences from samples composed of discrete instances (Xie, Navarro, & Hayes, 2021). In contrast, they struggle with drawing appropriate inferences from samples that contain repeated presentations of the same instance – often overweighting the evidential value of these repetitions (cf. Connor Desai, Xie, & Hayes, 2022; Unkelbach, 2007; Yousif, Aboody, & Keil, 2019).

5.3. The continued influence effect and retraction of misinformation

We have noted parallels between the influence of initial and revised sampling assumptions and studies of the continued influence of misinformation on event memory (e.g., Ecker et al., 2011; Johnson & Seifert, 1994). Two different theoretical accounts have been proposed to explain the continued influence effect (Chan, Jones, Hall Jamieson, & Albarracín, 2017; Lewandowsky et al., 2012). In the selective retrieval account, continued influence occurs when correct and incorrect information are simultaneously stored in memory; upon retrieval, misinformation is activated but inadequately suppressed. In the model-updating account,

removing misinformation leaves a gap in people's mental models. People prefer a coherent (incorrect) mental model to an incoherent (correct) mental model, so the misinformation is maintained. A correction may not fill the mental gap left by removing misinformation unless it provides an alternative explanation for the event's outcome (e.g., Ecker et al., 2011). Corrections should provide information that can effectively replace the refuted mental model components without compromising the coherence of the existing elements (Lewandowsky et al., 2012).

The current results are consistent with the model-updating account. We found that the influence of an initial sampling frame on property inferences could be overridden by an alternative causal explanation provided after the data were observed. In other words, people were able to shift between different representations or "mental models" of how the sample data were generated. The current work extends this model updating account beyond the reinterpretation of existing evidence. We have shown that revising one's model of sample selection can change the property inferences that one makes about novel items that were not part of the original sample.

More generally, our results suggest that it may be possible to counter the negative effects of misleading or distorted information via "debunking" or retrospective re-evaluation of that information based on an alternate causal model. Meta-analytic reviews of debunking interventions (Chan et al., 2017; Walter & Tukachinsky, 2020), have shown that such interventions are more likely to succeed when they offer a coherent, alternative explanation of the misinformation.

Reviews of research on the retraction of misinformation also suggest issues of interest for future studies of the updating of sampling assumptions. For example, Walter and Tukachinsky (2020) found that it becomes more difficult to counteract the effects of misinformation with increasing delays between exposure to the misinformation and its correction. In the current work, there were relatively short retention intervals between the presentation of the initial sampling frame, the sample data, and the final sample frame. Future work, therefore, should vary these intervals, and examine whether people are able to successfully update their sampling frames when there is a longer delay between exposures to the sample data and the revised frame.

5.4. Conclusions

When we make inferences based on samples of evidence, we consider not just the contents of the sample, but also how the sample was generated. The current work has shown that people can revise and update their beliefs about the sample generation process – switching from an initial explanation of sampling constraints to a revised explanation received after the sample was observed. Such flexibility has important benefits – in particular, it means that old, incorrect beliefs about the data can be revised, and that people can adjust their inductive inferences accordingly.

CRediT authorship contribution statement

Brett K. Hayes: Conceptualization, Methodology, Supervision, Writing – original draft, Funding acquisition. **Joshua Pham:** Conceptualization, Investigation, Methodology, Writing – review & editing. **Jaimie Lee:** Data curation, Software. **Andrew Perfors:** Conceptualization, Writing – review & editing. **Keith Ransom:** Conceptualization, Writing – review & editing. **Saoirse Connor Desai:** Conceptualization, Methodology, Software, Visualization, Writing – original draft.

Declaration of competing interest

The authors have no known conflicts of interest to disclose.

Data availability

I have shared the data at the Open science Archive https://osf.io/qjgxt/

Acknowledgements

This work was supported by Australian Research Council Discovery Grants DP190101224 and DP220101592 to BKH. We thank Won Jae Lee for assistance in the preparation of the manuscript.

Appendix A

A.1. Experiment 2 Instructions

Frames Condition	Instructions presented after the warm up phase
Category Sampling (including Category only and Category to Property)	In this phase you will see a sample collected by the robot on the left (only collects small birds). You discover that this robot had already been used to bring back a collection of 10 small, sparrow-like birds that it found on the island. Given time constrains, you decide to use this sample in your investigation and test the birds for plaxium blood.
Property Sampling (including Property only and Property to Category)	In this phase you will see a sample collected by the robot on the right (only collects animals that test positive for plaxium). You discover that this robot had already been used to bring back a collection of 10 animals that had tested positive for the presence of plaxium blood. You decide to use this sample in your investigation and examine the types of animals present.
	Instructions presented after the complete sample was observed (switch conditions only)
Category-to-Property	WARNING: There was a mix-up in the information you were given about which robot collected the samples that you have seen. The samples you saw were actually collected by the robot designed to collect only animals that tested positive for plaxium blood using an inbuilt 'plaxium-sensitive camera'. You realise that you must disregard your previous conclusions and use this new information to answer your research question about which animals on the island have plaxium. Here is the sample of animals that the robot collected with the correct robot beside it.
Property-to-Category	WARNING: There was a mix-up in the information you were given about which robot collected the samples that you have seen. The samples you saw were actually collected by the robot designed to collect only samples of small birds. You realise that you must disregard your previous conclusions and use this new information to answer your research question about which animals on the island have plaxium. Here is the sample of animals that the robot collected with the correct robot beside it.

References

Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology, 2*(1), 1–9. https://doi.org/10.1037/h0021966

Anderson, N. H. (1973). Serial position curves in impression formation. *Journal of Experimental Psychology, 97*(1), 8. https://doi.org/10.1037/h0033774

Bareinboim, E., Tian, J., & Pearl, J. (2014). Recovering from selection bias in causal and statistical inference. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence* (pp. 2410–2416).

Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science, 28*(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching, 35*(6), 623–654. https://doi.org/10.1002/(SICI)1098-2736(199808)35:6<623

Connor Desai, S., Pilditch, T. D., & Madsen, J. K. (2020). The rational continued influence of misinformation. *Cognition, 205*, Article 104453. https://doi.org/10.1016/j.cognition.2020.104453

Connor Desai, S., & Reimers, S. (2019). Comparing the use of open and closed questions for web-based measures of the continued-influence effect. *Behavior Research Methods, 51*, 1426–1440. https://doi.org/10.3758/s13428-018-1066-z

Connor Desai, S., & Reimers, S. (2023). Does explaining the origins of misinformation improve the effectiveness of a given correction? *Memory & Cognition, 51*(2), 422–436. https://doi.org/10.3758/s13421-022-01354-7

Connor Desai, S., Xie, B., & Hayes, B. K. (2022). Getting to the source of the illusion of consensus. *Cognition, 223*, Article 105023. https://doi.org/10.1016/j.cognition.2022.105023

Dennis, M. J., & Ahn, W. K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition, 29*(1), 152–164. https://doi.org/10.3758/BF03195749

Ecker, U. K. H., Lewandowsky, S., Cheung, C. S. C., & Maybery, M. T. (2015). He did it! She did it! No, she did not! Multiple causal explanations and the continued influence of misinformation. *Journal of Memory and Language, 85*, 101–115.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review, 18*, 570–578. https://doi.org/10.3758/s13423-011-0065-1

Feeney, A. (2018). Forty years of progress on category-based inductive reasoning. In L. J. Ball, & V. A. Thompson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 167–185). Routledge/Taylor & Francis Group.

Feiler, D. C., Tong, J. D., & Larrick, R. P. (2013). Biased judgment in censored environments. *Management Science, 59*(3), 573–591.

Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General, 129*(3), 399. https://doi.org/10.1037/0096-3445.129.3.399

Fiedler, K., Prager, J., & McCaughey, L. (2023). Metacognitive myopia: A major obstacle on the way to rationality. *Current Directions in Psychological Science, 32*(1), 49–56. https://doi.org/10.1177/096372142211269

Ganea, P. A., Larsen, N. E., & Venkadasalam, V. P. (2021). The role of alternative theories and anomalous evidence in children's scientific belief revision. *Child Development, 92*(3), 1137–1153. https://doi.org/10.1111/cdev.13481

Hadjichristidis, C., Geipel, J., & Gopalakrishna Pillai, K. (2022). Diversity effects in subjective probability judgment. *Thinking & Reasoning, 28*(2), 290–319. https://doi.org/10.1080/13546783.2021.2000494

Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1366–1373. https://doi.org/10.1037/a0016567

Hayes, B. K., Banner, S., Forrester, S., & Navarro, D. J. (2019). Selective sampling and inductive inference: Drawing inferences based on observed and missing evidence. *Cognitive Psychology, 113*, Article 101221. https://doi.org/10.1016/j.cogpsych.2019.05.003

Hayes, B. K., & Heit, E. (2018). Inductive reasoning 2.0. *Wiley interdisciplinary reviews. Cognitive Science, 9*(3), Article e1459. https://doi.org/10.1002/wcs.1459

Hayes, B. K., Liew, S. X., Connor Desai, S., Navarro, D. J., & Wen, Y. (2023). Who is sensitive to selection biases in inductive reasoning? *Journal of Experimental Psychology: Learning, Memory, and Cognition., 49*(2), 284–300. https://doi.org/10.1037/xlm0001171

Hayes, B. K., Navarro, D. J., Stephens, R. G., Ransom, K., & Dilevski, N. (2019). The diversity effect in inductive reasoning depends on sampling assumptions. *Psychonomic Bulletin & Review, 26*, 1043–1050. https://doi.org/10.3758/s13423-018-1562-2

Hayes, B. K., & Thompson, S. (2007). Causal relations and feature similarity in children's inductive reasoning. *Journal of Experimental Psychology: General, 136*, 470–484. https://doi.org/10.1037/0096-3445.136.3.470

Heit, E., & Hahn, U. (2001). Diversity-based reasoning in children. *Cognitive Psychology, 43*(4), 243–273. https://doi.org/10.1006/cogp.2001.0757

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*(1), 1–55. https://doi.org/10.1016/0010-0285(92)90002-J

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science, 24*(5), 379–385. https://doi.org/10.1177/0963721415591878

JASP Team. (2023). *JASP (Version 0.17.1) [Computer Software]*.

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420. https://doi.org/10.1037/0278-7393.20.6.1420

Kalish, M. L., Lewandowsky, S., & Davies, M. (2005). Error-driven knowledge restructuring in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 846. https://doi.org/10.1037/0278-7393.31.5.846

Kan, I. P., Pizzonia, K. L., Drummey, A. B., & Mikkelsen, E. J. (2021). Exploring factors that mitigate the continued influence of misinformation. *Cognitive Research: Principles and Implications, 6*(1), 1–33. https://doi.org/10.1186/s41235-021-00335-9

Kary, A., Newell, B. R., & Hayes, B. K. (2018). What makes for compelling science? Evidential diversity in the evaluation of scientific arguments. *Global Environmental Change, 49*, 186–196. https://doi.org/10.1016/j.gloenvcha.2018.01.004

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kemp, C., Navarro, D. J., & Hayes, B. K. (2023). *Adjustment for selection bias in intuitive reasoning* (Manuscript in preparation).

Kendeou, P., Butterfuss, R., Kim, J., & Van Boekel, M. (2019). Knowledge revision through the lenses of the three-pronged approach. *Memory & Cognition, 47*, 33–46. https://doi.org/10.3758/s13421-018-0848-y

Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 854–865. https://doi.org/10.1037/a0029468

Kim, N. S., & Keil, F. C. (2003). From symptoms to causes: Diversity effects in diagnostic reasoning. *Memory & Cognition, 31*(1), 155–165. https://doi.org/10.3758/BF03196090

Koehler, J. J., & Mercer, M. (2009). Selection neglect in mutual fund advertisements. *Management Science, 55*(7), 1107–1121. https://doi.org/10.1287/mnsc.1090.1013

Lagnado, D. A., & Harvey, N. (2008). The impact of discredited evidence. *Psychonomic Bulletin & Review, 15*(6), 1166–1173. https://doi.org/10.3758/PBR.15.6.1166

Lawson, C. A., & Kalish, C. W. (2009). Sample selection and inductive generalization. *Memory & Cognition, 37*(5), 596–607. https://doi.org/10.3758/MC.37.5.596

Le Mens, G., & Denrell, J. (2011). Rational learning and information sampling: On the "naivety" assumption in sampling explanations of judgments biases. *Psychological Review, 118*(2), 379–392. https://doi.org/10.1037/a0023010

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Liew, J., Grisham, J. R., & Hayes, B. K. (2018). Inductive and deductive reasoning in obsessive-compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 59*, 79–86. https://doi.org/10.1016/j.jbtep.2017.12.001

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology, 29*(2), 85–148. https://doi.org/10.1006/cogp.1995.1013

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT press.

Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. K. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review, 10*(3), 517–532. https://doi.org/10.3758/BF03196515

Mickelberg, A. J., Walker, B., Ecker, U. K. H., Howe, P. D. L., & Perfors, A. (2023). *Does mud always stick? No Evidence for Continued Influence of Valenced Misinformation on Person Impressions* (Manuscript under review).

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*(3), 289. https://doi.org/10.1037/0033-295X.92.3.289

Navarro, D. J., Dry, M. J., & Lee, M. D. (2012). Sampling assumptions in inductive generalization. *Cognitive Science, 36*(2), 187–223. https://doi.org/10.1111/j.1551-6709.2011.01212.x

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*(2), 185. https://doi.org/10.1037/0033-295X.97.2.185

van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review, 8*, 28–61.

Peake, P. K., & Cervone, D. (1989). Sequence anchoring and self-efficacy: Primacy effects in the consideration of possibilities. *Social Cognition, 7*(1), 31–50. https://doi.org/10.1521/soco.1989.7.1.31

Pitrelli, M. (2022, December). *What do hotel 'star' ratings really mean?* CNBC: Here's a breakdown. https://www.cnbc.com/2022/12/11/.

Ransom, K., Hendrickson, A. T., Perfors, A., & Navarro, D. J. (2018). Representational and sampling assumptions drive individual differences in single category generalization. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society* (pp. 931–935). Austin, TX: Cognitive Science Society.

Ransom, K. J., Perfors, A., Hayes, B. K., & Connor Desai, S. (2022). What do our sampling assumptions affect: How we encode data or how we reason from it? *Journal of Experimental Psychology: Learning, Memory, and Cognition.*. https://doi.org/10.1037/xlm0001149

Ransom, K. J., Perfors, A., & Navarro, D. J. (2016). Leaping to conclusions: Why premise relevance affects argument strength. *Cognitive Science, 40*(7), 1775–1796. https://doi.org/10.1111/cogs.12308

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition, 35*(8), 2019–2032. https://doi.org/10.3758/BF03192934

Rehder, B. (2017). Concepts as causal models: Induction. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 377–413). Oxford University Press.

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery. *Developmental Science, 13*(3), 421–429. https://doi.org/10.1111/j.1467-7687.2009.00896.x

Rich, P. R., & Zaragoza, M. S. (2020). Correcting misinformation in news stories: An investigation of correction timing and correction durability. *Journal of Applied Research in Memory and Cognition, 9*(3), 310–322. https://doi.org/10.1016/j.jarmac.2020.04.001

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 665–681. https://doi.org/10.1016/S0022-5371(75)80055-7

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*(5), 880. https://doi.org/10.1037/0022-3514.32.5.880

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E. J. (2017). Bayesian analysis of factorial designs. *Psychological Methods, 22*(2), 304. https://doi.org/10.1037/met0000057

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review, 25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Shengelia, T., & Lagnado, D. (2021). Are jurors intuitive statisticians? Bayesian causal reasoning in legal contexts. *Frontiers in Psychology, 11*, Article 519262. https://doi.org/10.3389/fpsyg.2020.519262

Tauber, S., Navarro, D. J., Perfors, A., & Steyvers, M. (2017). Bayesian models of cognition revisited: Setting optimality aside and letting data drive psychological theory. *Psychological Review, 124*(4), 410. https://doi.org/10.1037/rev0000052

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences, 24*(4), 629–640. https://doi.org/10.1017/S0140525X01000061

Thagard, P. (2008). Explanatory coherence. In J. E. Adler, & L. J. Rips (Eds.), *Reasoning: Studies of human inference and its foundations* (pp. 471–513). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.024.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 219. https://doi.org/10.1037/0278-7393.33.1.219

Vosniadou, S. (2013). Conceptual change in learning and instruction: The framework theory approach. In *International handbook of research on conceptual change* (pp. 11–30). Routledge. https://doi.org/10.1007/s11191-010-9283-6.

Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research, 47*(2), 155–177. https://doi.org/10.1177/0093650219854600

Xie, B., Navarro, D., & Hayes, B. K. (2021). Adding types, but not tokens, affects property induction. *Cognitive Science, 44*, Article e12895. https://doi.org/10.1111/cogs.12895

Yousif, S. R., Aboody, R., & Keil, F. C. (2019). The illusion of consensus: A failure to distinguish between true and false consensus. *Psychological Science, 30*(8), 1195–1204. https://doi.org/10.1177/0956797619856844