Behavioural Public Policy (2026), 1–32
doi:10.1017/bpp.2025.10028

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Negative anecdotes reduce policy support: evidence from three experimental studies on communicating policy (in)effectiveness

Amy Rodger[1], Greta Arancia Sanna[2], Vanessa Cheung[2], Nichola Raihani[2,3] and David Lagnado[2]

[1]Usher Institute, University of Edinburgh, UK; [2]Experimental Psychology, University College London, London, UK and [3]School of Psychology, University of Auckland, Auckland, New Zealand
Corresponding author: Amy Rodger; Email: a.rodger@ed.ac.uk

(Received 25 March 2025; revised 17 November 2025; accepted 25 November 2025)

Abstract
Public support is crucial for the success of policy interventions that aim to change behaviour. While communicating evidence of policy effectiveness can increase support, it remains unclear which type of evidence is most effective. Statistical evidence is often seen as objective and persuasive, yet personal anecdotes can strongly influence beliefs. We examined how statistical and anecdotal evidence affect policy perceptions. In three online experiments with representative UK samples ($N = 908$), we showed participants different types of evidence (statistical, anecdotal, or both) that argued for or against six policies, such as meat taxes (climate change), banning e-cigarettes (public health), and 20 mph speed limits (community safety). We measured policy support and perceived effectiveness before and after exposure and explored participants' reasoning through open-text responses. Results showed that positive statistical and anecdotal evidence did not significantly increase perceived policy effectiveness or support, even when combined. However, negative anecdotes significantly reduced both, though this effect was sometimes mitigated when paired with statistical evidence. Qualitative results found that participants have broader concerns beyond policy effectiveness, such as fairness. Our findings suggest that communicating evidence on policy effectiveness alone may not increase support, as it does not address broader public concerns.

Keywords: belief updating; communication; narratives; policy; public acceptance; statistical evidence

Introduction

Many societal challenges, such as public health crises and climate change, require policymakers to implement interventions aimed at changing behaviour (Carroll *et al.*, 2019). However, these interventions often fail without public support, which is crucial

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

for policy implementation and compliance (Burstein, 2003, 2020; Reynolds *et al.*, 2020; Grelle and Hofmann, 2024; WHO Regional Office for Europe, 2024). The current research examines how communicating evidence of policy effectiveness influences public support. This strategy has increased the public's acceptance of interventions, such as vaccine uptake, by enhancing perceptions of policy effectiveness – a key predictor of support (Bigman *et al.*, 2010; Promberger *et al.*, 2012; Pechey *et al.*, 2014; Reynolds *et al.*, 2020). However, it remains unclear whether statistical evidence from scientific research, personal anecdotes from the public, or a combination of both is most effective in achieving this goal.

The persuasiveness of different types of evidence

Policy information may be communicated using statistical evidence (e.g., 'Research shows this policy could reduce greenhouse gas emissions by 22.1%') and qualitative, anecdotal evidence (e.g., 'I think this policy would help us reduce our carbon footprint'). While statistics are often regarded as the most effective way to change people's beliefs, attitudes and intentions (Graeber *et al.*, 2024), evidence on this issue is mixed. Recent meta-analyses (Freling *et al.*, 2020; Xu, 2023) found that while statistical and anecdotal evidence had similar effects on beliefs, attitudes, and intentions overall, anecdotal evidence was more persuasive when issues were personally relevant or contained high emotional stakes (Freling *et al.*, 2020). For example, anecdotes were more effective in public health messaging about personal health issues (Betsch *et al.*, 2011; Xu, 2023). This suggests that anecdotes may be more persuasive than statistics in certain contexts.

Combining statistical and anecdotal evidence may be more persuasive than either alone. Combined evidence improved U.S. public attitudes and intentions towards skin cancer interventions more than anecdotes or statistics alone (Zhou and Yang, 2024). However, its effectiveness may depend on anecdote type, with combined evidence using individual-based anecdotes being more effective than single pieces of evidence, while combined evidence using group-based anecdotes are not (Niederdeppe *et al.*, 2016; Hornikx, 2018; Zhou and Yang, 2024).

In real-world policy communication, the public often encounters conflicting evidence, including contradictory expert opinions and mixed messages from social media. For example, exposure to conflicting information on red meat and health risks reduces public trust in scientific consensus and confidence in the evidence (Iles *et al.*, 2022). Similarly, counterarguments using statistics in online comments from the public can weaken belief in climate crisis news (Hinnant *et al.*, 2016, 2023). Long-term exposure to conflicting health messages also increases resistance to health campaigns, even on unrelated health issues (Nagler *et al.*, 2022).

Overall, these mixed and sometimes limited research findings highlight the need for further study into the effect of anecdotal, statistical, combined, and conflicting evidence in policy communication to avoid eroding public trust and support (Reynolds *et al.*, 2020).

Overview of current research

We conducted three online experiments to examine how different types (statistical, anecdotal or combined) and directions of evidence (for or against a policy) influenced

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

UK adults' perceived policy effectiveness and support across various policy domains. Ongoing debates question whether people process political evidence through rational evaluation or motivated reasoning, emphasising the need to establish effects across various evidence types and policy domains before examining the underlying mechanisms. Recent meta-analyses have produced inconclusive findings on the impact of different evidence types (Freling *et al.*, 2020; Xu, 2023), highlighting the importance of first documenting these effects before delving into the mechanisms involved.

We focused on policies addressing pressing societal issues with typically low public support, including public health (e.g., banning e-cigarettes), climate change (e.g., taxing meat and seafood), and community safety (e.g., 20 mph speed limits). In all three studies, we used a within-subjects design. Participants saw various policies (e.g., meat/seafood tax, 20 mph speed limit) and rated the effectiveness of each policy and their support for it. They were then shown a specific type of evidence about each policy (different for each policy) and asked to rate each policy again. In Study 1, we used the following types and directions of evidence: statistical evidence supporting policy effectiveness, personal anecdotes covering a positive or negative opinion on effectiveness, a combination of positive statistical and anecdotal evidence, and a combination of positive statistical and negative anecdotal evidence. In Study 2, we added a stronger statistics condition for the statistical evidence, and in Study 3, we added a new condition combining positive and negative anecdotes. Our first research question was:

What effect (if any) does evidence condition have on (a) change in perceived policy effectiveness and (b) change in policy support?

A key limitation of research on communicating policy effectiveness is that studies only identify the effects of different evidence types, and offer limited insight into why these effects occur (Oeberst and Imhoff, 2023; Simon and Read, 2023), which is crucial for informing communication strategies (Bago *et al.*, 2023). To address this, we used qualitative methods in Studies 1 and 2 to examine participants' thought processes when evaluating policy evidence, identifying policy considerations beyond effectiveness that may inform changes in perceived policy effectiveness and support. Similar approaches have been applied in health and climate communication (Guan *et al.*, 2023; Song and So, 2023; Wolf *et al.*, 2024). Our second research question was:

What issues do people consider when reasoning with policy evidence and explaining their policy views?

Study 1

Method

All studies were approved by the UCL Psychology Ethics Committee

(EP/2018/005). The Open Science Framework project for this research (Rodger *et al.*, 2025; https://doi.org/10.17605/OSF.IO/3QTZJ) contains all pre-registered protocols (https://doi.org/10.17605/OSF.IO/M7X4R), supplementary materials and measures (https://doi.org/10.17605/OSF.IO/H6YGR), raw data (https://doi.org/10.17605/OSF.IO/Z4WEM) and analysis files (https://doi.org/10.17605/OSF.IO/YKCS4).

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Participants

We performed an a priori power analysis with SIMR (Green and MacLeod, 2016), which uses simulation-based power calculations appropriate for mixed-effects models, using pilot data ($N = 100$; see demo2_pilot2_analysis.Rmd on OSF). The effect size was defined by setting fixed effects at 4 points on the 100-point scale when comparing evidence conditions. We powered for 4% as previous work noted that small changes to public support, such as 4%, can impact policy outcomes (e.g., the EU referendum, which came down to 3.8%, Reynolds *et al.*, 2020). We calculated that a sample size of 250 participants would provide 90% power to detect this effect, based on power curve analysis for samples of 100, 150, 200, and 250 participants. We recruited 301 participants via Prolific (http://www.prolific.com), as 300 was the minimum required by Prolific for their representative sampling. After collecting Study 1 data, we conducted a further power analysis with the same method (see demo2_study_2_analysis.Rmd on OSF), which validated our sample size decisions for subsequent studies.

We recruited UK adults eligible to vote and representative of the UK voting population in terms of age, gender and political affiliation using UK census data. Participants were UK residents aged 18 and above, fluent in English and paid £4.11. Table A1 (Supplemental Methods) detail participant demographics. The Study 1 sample ($N = 301$) had a mean age of 46.2 years (SD = 15.92), 52% identified as women, 48% as men, 87% as white, 54% were employed full-time, and 59% held a Bachelor's degree or higher. Politically, 47% were affiliated with the Labour Party and 18% with the Conservative Party.

Design

We conducted a within-subjects experiment through Qualtrics (http://www.qualtrics.com), where participants were exposed to all six evidence conditions (Table A2), with each evidence condition randomly paired with a different policy (Table A3). This resulted in a repeated measures design with evidence condition as the fixed effect of interest and policy as a crossed random factor to enhance generalisability across different policy domains. Note that the 'Combined Congruent' condition includes two pieces of positive evidence (and not negative evidence), and the 'Combined Incongruent' condition in Study 1 included positive statistical evidence and negative anecdotal evidence. We did not test the full factorial of evidence types and directions, excluding negative statistical evidence. This decision was made for two reasons: (1) including negative statistical evidence would have significantly increased study length, potentially burdening participants, and (2) policymakers are unlikely to use statistics to highlight policy ineffectiveness in communication campaigns, making it slightly less relevant to current research.

Our dependent variables were changes in perceived policy effectiveness and policy support before versus after exposure to evidence. We also collected open-text data by asking participants to freely list their thoughts while viewing policy evidence ('thought-listing') and to explain their posterior policy judgements via open-text responses.

Materials and measures

Policies. We selected policies aimed at changing public behaviour that have varying levels of public support (<70%) and have not been rolled out nationwide across the UK.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

We reviewed academic research and reports from organisations such as the OECD to identify recommended policies for tackling public health, climate change, and community safety issues and to assess public support for each policy (see Table A3 for references used to source these policies). Support levels observed in our pilot study confirmed that these policies met our selection criteria (see OSF, demo2_pilot2_analysis.Rmd). This approach ensured we tested evidence effects across diverse policy contexts with varying levels of baseline support to enhance generalisability.

Policy evidence. Across all policies, we followed the same stimuli-generation protocol for each evidence condition (see OSF). For the 'Control' condition, we presented a brief rationale that the policy aimed to tackle a societal challenge. The 'Statistic' condition presented plausible UK-relevant statistics on policy effectiveness from existing research. For the 'Anecdote For' and 'Anecdote Against' conditions, we used ChatGPT to generate anecdotes based on real public opinions from existing qualitative research on each policy. The research team reviewed all anecdotes to ensure they accurately reflected authentic public concerns and maintained appropriate tone and realism (see OSF, demo2_stimuli_generation_protocol.docx). The credibility scores ($M > 60$ out of 100 across all conditions) from the pilot study and Study 1 also suggested that participants viewed the stimuli as credible and did not systematically interpret certain evidence types as less credible. Table A2 shows an overview of evidence conditions and example stimuli.

Measures. We used the following measures in the study:

Perceived effectiveness: Participants rated (dis)agreement with the statement: '[Policy *X*] is an effective solution for tackling [Issue *Y*] in the UK as it will change [Behaviour *Z*].' (1 = Strongly disagree, 100 = Strongly agree). For example: 'Introducing energy quotas is an effective solution for tackling climate change in the UK, as it will reduce household energy use.' We calculated a change score by subtracting the posterior from the prior judgment. This measure was adapted from prior research on communicating policy effectiveness evidence (Mantzari *et al.*, 2022; Reynolds *et al.*, 2022) and follows typical measurement formats in belief updating research (Vlasceanu *et al.*, 2023). In Study 1, participants provided open-text explanations for their posterior effectiveness ratings.

Support: 'How much do you support or oppose the implementation of [Policy *X*] in the UK?' (1 = Strongly oppose, 100 = Strongly support). We calculated a change score by subtracting the posterior from the prior judgment. This measure was also adapted from prior research on communicating policy effectiveness evidence (Mantzari *et al.*, 2022; Reynolds *et al.*, 2022). We added a reference to policy 'implementation' to align with current definitions of policy support (Grelle and Hofmann, 2024) and broader policy support research (Petrescu *et al.*, 2016). In Study 1, participants provided open-text explanations for their posterior support ratings.

Thought listing: We asked participants to read the evidence and write their thoughts without filtering them in an open-text box.

Evidence credibility: Based on Appelman and Sundar (2016), participants rated how well different adjectives (e.g., accurate, authentic, believable) described the evidence (1 = Describes very poorly, 100 = Describes very well). We averaged responses ($\alpha = 0.96$, high reliability). More details and results (no credibility differences across conditions) are available on OSF.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Procedure

Study 1 took an average of 30 minutes. The full Qualtrics survey, available on OSF, included the following steps:

1. Instructions: Participants were informed that the study examined public opinions on UK government policies related to health, safety, and climate change, as well as how people evaluate policy-related information.
2. Baseline measures: Participants reported their prior perceived effectiveness and support for each of the six policies in a random order.
3. Evidence conditions. Participants were told they would read six excerpts from citizens' juries discussing various policies. The six evidence conditions were presented in random order, with the following process for each
   i. Each evidence condition was paired with a randomly selected policy (without replacement), meaning once a policy had been selected, it was not selected in any proceeding evidence conditions.
   ii. Participants read the stimulus related to the selected policy and evidence condition, then responded to the thought-listing task and message credibility questions.
   iii. After viewing the stimulus, participants reported their posterior perceived effectiveness and support, providing open-text justifications for their responses.
4. Final measures: Participants responded to demographic measures.

Data analysis

Quantitative data analysis. All analyses were conducted in R (Team, R Core, 2023), following the preregistered analysis plan on OSF. We used the lme4 package (Bates *et al.*, 2024) to run linear mixed-effects models assessing the effect of evidence condition on change in perceived effectiveness and support. The model was specified as: Outcome $\sim$ Evidence Condition + (1 | Participant) + (1 | Policy) + (1 | Stimuli)[1]. To test direct comparisons, we ran the model with two different reference groups: the 'Control' and the 'Statistic' condition. We hypothesised that in comparison to each reference condition, the 'Statistic' (compared to the control condition only), 'Anecdote For', and 'Combined Congruent' conditions would have a larger increase in both outcomes, the 'Anecdote Against' would have a larger decrease in both outcomes and the 'Combined Incongruent' would have different change in both outcomes. We had a non-directional hypothesis for the 'Combined Incongruent' condition due to limited prior research on the effect of this type of evidence. We applied a significance threshold of $p < .005$, using the Bonferroni correction for ten direct comparisons across two models.

Qualitative data analysis. We conducted a ChatGPT-assisted qualitative content analysis (Bijker *et al.*, 2024) using GPT-4o (OpenAI, 2025) to develop a coding scheme and code the dataset. GPT-4o was accessed via the API using the openai

[1]Random slopes were initially planned to assess variation in the effect of evidence condition across people and policies but were removed due to model convergence issues.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

package (Rudnytskyi, 2023) in R. The full protocol, including code, prompts, and the researcher-developed codebook, is available on OSF. The process is summarised below:

Data Preparation: Participants' open-text responses were processed by GPT-4o to generate discrete meaning units (1–4 per response). This was repeated three times, producing three datasets. Two authors reviewed the discrete meaning units to assess whether they accurately reflected content in the original response for 75 responses per dataset, resolving discrepancies through discussion. Precision, calculated as correct meaning units divided by total units, ranged from 0.87 to 0.91. The most precise dataset was selected for analysis.

Coding Scheme Development & Coding: GPT-4o generated three coding schemes from the most precise dataset based on patterns in the first 75 responses, removing one for redundancy. The full dataset was coded using the remaining two, each coded three times to simulate multiple coders. Inter-rater agreement (0.79 vs 0.82) was calculated, and the highest agreement scheme (0.82) was selected. R was used to compare category distributions across evidence conditions, policies, and participant groups. The final codebook (Table 1) includes example quotes. Unclassifiable responses were categorised as N/A. Table A4 reports N/A responses per policy.

Validity Assessment: GPT-4o's coding scheme was compared to a manually developed scheme based on 25% of the dataset. While GPT-4o identified similar themes (e.g., fairness), it produced fewer, higher-level codes. This aligned with the study's goal of providing a broad descriptive overview of issues raised by participants.

Results

Quantitative results

At baseline, participants generally opposed energy quotas and meat/seafood taxes, rating them as ineffective ($M < 50$). Most other policies received neutral ratings (~50%) for effectiveness and support, while the e-cigarette ban and alcohol health labels had the highest support ratings. Figure 1 shows similar baseline trends across studies.

Tables B1 and B2 (Supplemental Results) summarise the models testing the effects of evidence conditions on perceived effectiveness and support. The data underlying these models is visualised in Figure 2 (see Figure 3 for comparison across studies).

In contrast to our hypotheses, the models suggest the 'Statistic', 'Anecdote For' and 'Combined Congruent' conditions showed changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions. This suggests that exposure to evidence supporting policy effectiveness, regardless of type, did not affect perceived effectiveness or support.

In line with our hypotheses, the models suggest the 'Anecdote Against' condition (Mchange = − 8.8, SD = 26) decreased perceived effectiveness by 7.9 points ($t(22.645) = 3.44$, $p = .002$, $d = .35$) compared to the 'Control' (Mchange = −0.8, SD = 23.6) and 9.7 points ($t(22.647) = 4.23$, $p < .001$, $d = .43$) compared to the 'Statistic' condition (Mchange = 1.0, SD = 23.2). The 'Anecdote Against' condition (Mchange = − 6.9, SD = 23.6) also decreased support by 5.10 ($t(1495.036) = 3.01$, $p = .003$, $d = .25$) and 5.08 points ($t(1495.104) = 3.00$, $p = .003$, $d = .24$), respectively, compared to the 'Control' (Mchange = −1.7, SD = 22.9) and 'Statistic' (Mchange = −1.8, SD = 20.2) conditions. This suggests that exposure to negative

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Table 1. Study 1: GPT-developed codebook for qualitative data analysis

Category	Definition	Example Quote
Awareness and education for behaviour change	Focuses on the role of information, education, and health awareness in influencing personal choices.	'[…] there needs to be re-education as people will continue to either try and break these rules or smoke in private places as we have seen with actual cigarettes.' (Participant 231, ObvID 1387, E-cigarette ban policy)
Economic burden and class disparity	Highlights the impact of policies on lower-income individuals and fairness across economic classes.	'I agree with the idea and the excerpt. However, those on a lower income often buy these products because they are cheaper, and can't afford the healthier options.' (Participant 85, ObvID 507, Food merch restrictions policy)
Environmental and climate responsibility	Highlights support for environmental impact reduction, focusing on emissions, climate change, and sustainable practices.	'Because reducing the amount of these foods being produced would be positive in bringing down carbon emissions.' (Participant 210, ObvID 3063, Meat/seafood tax policy)
Freedom of choice and government control	Criticizes perceived government overreach in individual decision-making and lifestyle restrictions.	'It doesn't address the issue of careful driving but rather imposes yet further restrictions on everyone to mitigate the behaviour of the few.' (Participant 142, ObvID 4459, 20 mph speed limit policy)
Health equity and accessibility of alternatives	Addresses the need for affordable, healthy alternatives and considers policies that affect access to basic needs.	'Whilst it will make HCE food more expensive, people will simply pay more, whilst the underprivileged will eat less, creating possible health problems.' (Participant 249, ObvID 3300, Meat/seafood tax policy)
Market influence and consumer behaviour	Focuses on how discounts, promotions, and availability drive consumer purchasing habits and health impacts.	'People need to know the risks. People who drink too much probably won't care. And what about the businesses? This could hurt their sales.' (Participant 112, ObvID 667, Alcohol health label policy)

(*Continued*)

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Table 1. (*Continued.*)

Category	Definition	Example Quote
Policy enforcement and compliance concerns	Emphasises issues with policy adherence, enforcement, and likelihood of compliance.	'I think it is unrealistic and feel that this will not be monitored enough to make it effective.' (Participant 248, ObvID 3289, 20 mph speed limit policy)
Practicality and realism in policy implementation	Evaluates policy feasibility, considering practical challenges and unintended consequences.	'I think it will have an impact and make people think twice, but they will probably still purchase it anyway and the demand will still be there. Still worth a shot however.' (Participant 57, ObvID 2148, Meat/seafood tax policy)
Public health and safety prioritisation	Considers policies as tools to improve public health and safety, especially around high-risk activities.	'Alcohol causes many problems in society with health and also social problems. Not enough has been done to tackle the ever-increasing problems of alcohol use on health.' (Participant 49, ObvID 291, Alcohol health label policy)
Social acceptance and cultural influences	Examines how social norms and cultural views on behaviours influence attitudes toward policy effectiveness.	'[…] I just don't think people will care about the warning labels as most people use alcohol as a means of escapism from their daily life struggles and for socialising.' (Participant 54, ObvID 324, Alcohol health label policy)

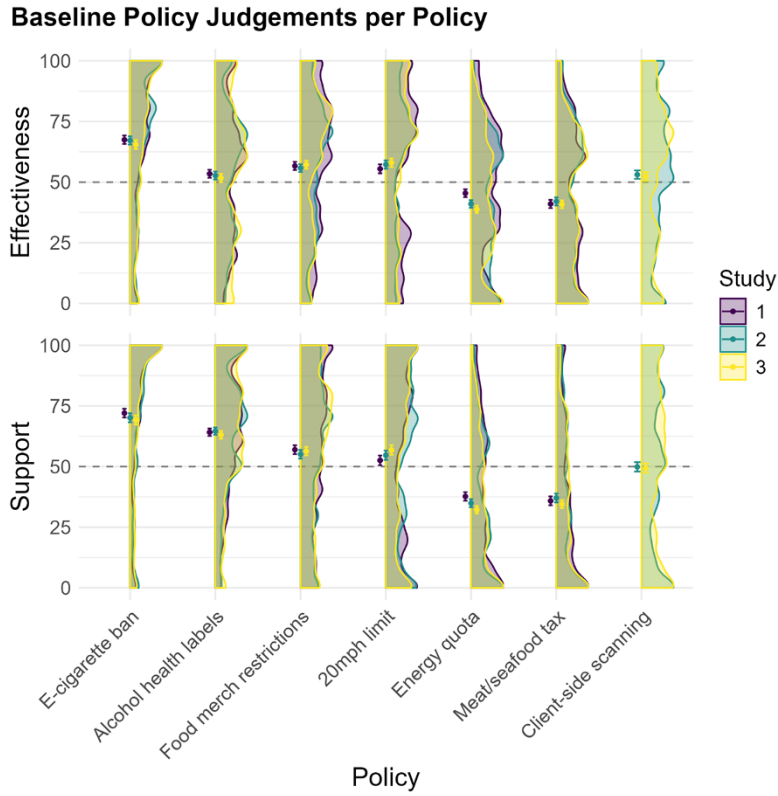


Figure 1. The plot shows the distribution (density plots), mean (points) and standard error of baseline policy judgements per policy for each study. The grey dashed line denotes a neutral policy judgement (i.e., Neither effective nor ineffective & neither support nor oppose). Each policy was rated by $N = 301$ (Study 1; 1806 observations), $N = 300$ (Study 2; 2100 observations) and $N = 307$ participants (Study 3; 2149 observations).

anecdotal evidence about policy effectiveness reduces perceived effectiveness and support.

In contrast to our hypotheses, the 'Combined Incongruent' condition showed changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions. This suggests that the effects of negative anecdotes on policy judgements may be mitigated by introducing positive statistical evidence.

Qualitative results

We identified various topics participants mentioned when reading new policy evidence and justifying posterior policy judgments. Figure 4 shows that most participants mentioned policy qualities such as fairness (i.e., economic burden), enforcement and compliance concerns, and more general beliefs such as freedom of choice. While these topics were mentioned across all policies, we observed some differences. For example,

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

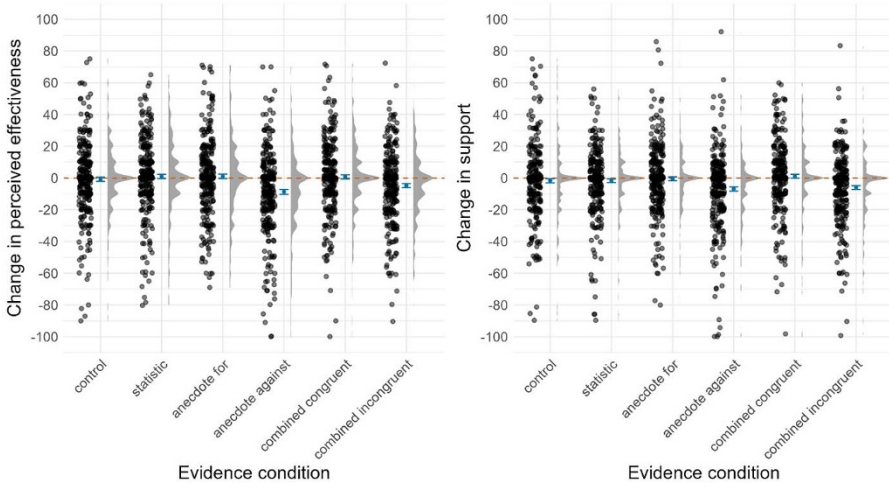


Figure 2. Study 1: The plot shows the change in perceived effectiveness and support per evidence condition. The points to the left of the blue mean and standard error bars show a participant's policy judgement for one of the six policies. The points are filled with a transparent grey, making overlapping points appear darker and distinguishing areas with high concentrations of data points. The density distribution to the right of the boxplot shows the data distribution. The Orange dashed line at zero represents no change in the policy judgements.

more participants mentioned issues with fairness for the meat/seafood tax, food merchandising restrictions, and energy quotas than the e-cigarette ban, 20 mph limit, and alcohol health labels (see Figure B3).

We also grouped participants into quantiles based on their baseline perceived effectiveness and support for each policy, exploring whether these topics were mentioned by more individuals in lower quantiles. Figure B4 shows that more participants with lower ratings for perceived effectiveness and support (Q1 and Q2) mentioned fairness, enforcement and compliance concerns, and freedom of choice compared to the other quantiles, while more participants with higher prior policy ratings (Q3 and Q4) emphasised public health and safety. Neutral topics, such as the need for awareness, education for behaviour change, social acceptance and cultural influences, were mentioned more evenly across quantiles.

Discussion

Unlike Reynolds *et al.* (2020), evidence supporting policy effectiveness, regardless of type, did not increase policies' perceived effectiveness or support. However, consistent with their findings, negative anecdotes reduced both. A key finding is that combining statistical evidence with negative anecdotal evidence mitigated the effect of negative anecdotes. Additionally, combining statistical and positive anecdotal evidence did not produce an additive effect, a finding consistent with Hornikx (2018) and Niederdeppe *et al.* (2016), but contrary to Zhou and Yang (2024). Our qualitative results indicate that participants considered factors beyond policy effectiveness, including fairness,

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

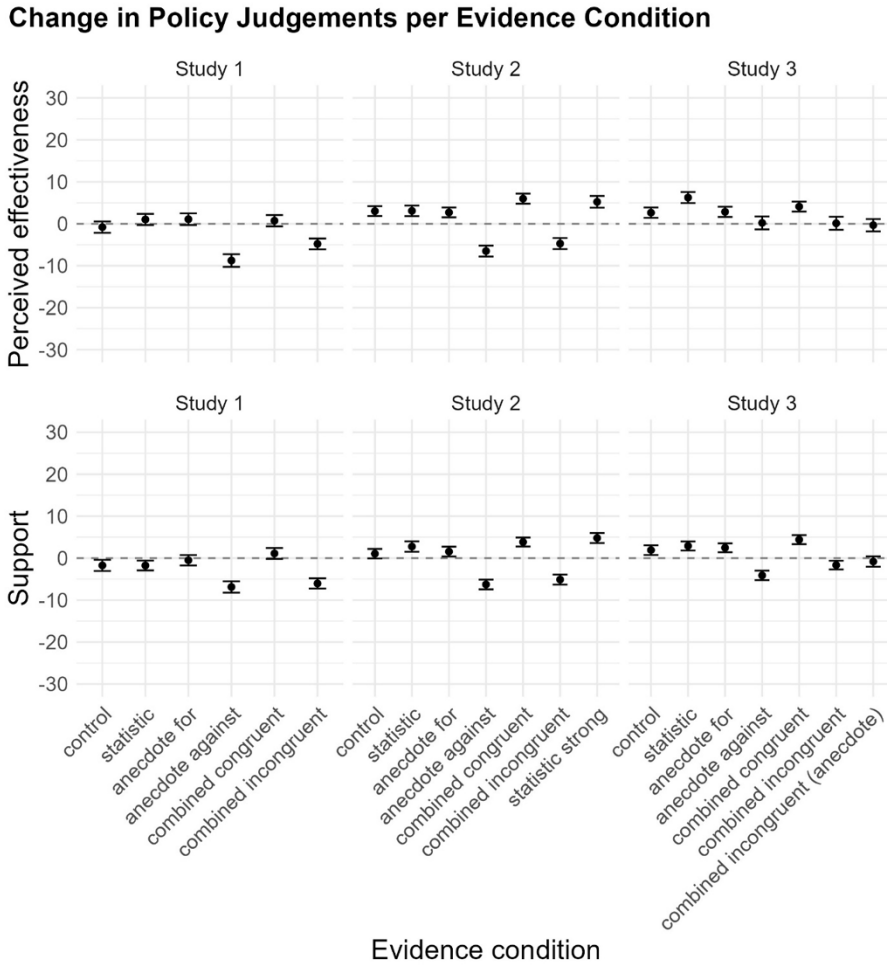


Figure 3. The plots show the mean and standard error of change in perceived effectiveness (left) and support (right) per evidence condition across policies for each study. The grey dashed line at zero represents no change in the policy judgement.

enforcement, compliance, and personal freedom, when evaluating evidence and justifying policy judgments. This supports a recent theory of public acceptance that suggests both general and policy-specific beliefs, beyond policy effectiveness, influence policy support (Grelle and Hofmann, 2024).

Study 2

In Study 2, we aimed to replicate Study 1's findings while addressing limitations. First, the 'Anecdote For' and 'Anecdote Against' conditions in Study 1 did not focus on the same aspects of policy effectiveness. For instance, in the 20 mph speed limit policy, the positive anecdote addressed changes in speeding behaviour, whereas the negative

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

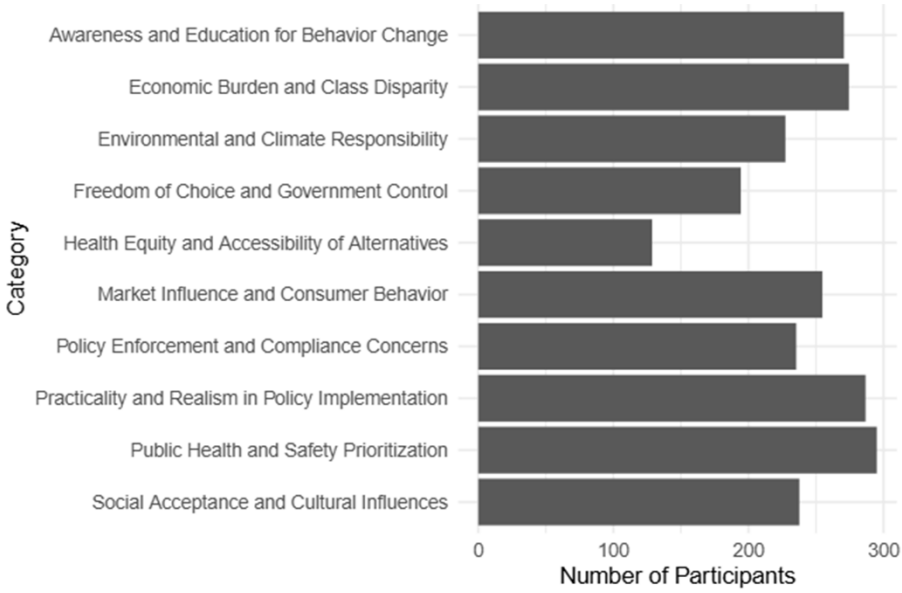


Figure 4. Number of participants mentioning each category in their responses.

anecdote discussed enforcement challenges. This imbalance may have contributed to the lack of effect from positive anecdotes. To address this, we generated more balanced anecdotes. Second, the lack of effect from statistical evidence in Study 1 may have resulted from the evidence's strength rather than an inherent ineffectiveness of statistics. We introduced a new evidence condition featuring stronger statistical evidence to test this. Finally, we added a more targeted open-text question asking why participants agreed or disagreed with the evidence, as the previous thought-listing task generated limited data directly related to the evidence being shown.

Method

Participants

We recruited 300 participants. The participants had a mean age of 46.44 years (SD = 15.99), 53% identified as women, 46% as men, 85% as white, 47% were employed full-time, and 52% held a Bachelor's degree or higher. Politically, 47% were affiliated with the Labour Party and 18% with the Conservative Party (see Table A1).

Materials and measures

We used the same design as Study 1, with the addition of the 'Strong Statistic's condition (see Table A2) and an online safety policy (see Table A3), resulting in a within-subjects experiment, where participants were exposed to all seven evidence conditions (Table A2), with each evidence condition randomly paired with a different policy. Differences from Study 1 are noted below.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Policy evidence. We revised the anecdotes in Study 1's 'Anecdote For' condition to ensure they argued for the same specific point about policy effectiveness as in the 'Anecdote Against' condition. For the 'Strong Statistic' condition, we doubled the values presented in the original 'Statistic' condition in Study 1.

Measures. In addition to the measures for perceived effectiveness and support used in Study 1, participants indicated their (dis)agreement with the point(s) being made in each evidence condition (1 = 'Strongly disagree' to 100 = 'Strongly agree'), then provided open-text explanations of their agreement rating.

Procedure

Study 2 took an average of 19 minutes, and participants were paid £3.75. The full Qualtrics survey for the study is available on OSF.

The procedure matched Study 1, except in Step 3, Part ii. While viewing the stimulus, participants rated agreement with the evidence and answered an open-text question but did not complete the thought-listing task, message credibility questions, or justify posterior policy judgments.

Data analysis

For the quantitative data, we followed the preregistered analysis plan in our study protocols (see OSF), the same as Study 1.

For the qualitative data, we conducted a ChatGPT-assisted qualitative content analysis, as outlined in Study 1. The final codebook, with example quotes, is shown in Table 2. It had an inter-rater agreement of 0.81.

Results

Quantitative results

Tables B3 and B4 summarise the models testing the effects of evidence condition on perceived effectiveness and support (see Figures 5 and 3).

In line with Study 1, the 'Statistic', 'Anecdote For', and 'Combined Congruent' conditions showed changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions. Similarly, the 'Strong Statistic' condition did not differ from the 'Statistic' condition, indicating that doubling statistical strength did not affect perceived effectiveness or support.

We also replicated the negative anecdote effect. The 'Anecdote Against' condition (Mchange = − 6.5, SD = 22.5) decreased perceived effectiveness by 9.52 points ($t(33.493) = 5.08$, $p < .001$, $d = .45$) compared to control (Mchange = 3.0, SD = 20.3) and 9.58 points ($t(33.485) = 5.11$, $p < .001$, $d = .45$) compared to the 'Statistic' condition (Mchange = 3.1, SD = 21.7). The 'Anecdote Against' condition also (Mchange = − 6.3, SD = 20.0) decreased support by 7.32 points ($t(34.293) = 3.61$, $p < .001$, $d = .38$) compared to control (Mchange = 1, SD = 19.8) and 9.05 points ($t(34.291) = 4.46$, $p < .001$, $d = .46$) compared to the 'Statistic' condition (Mchange = 2.8, SD = 21.5).

In contrast to Study 1, the 'Combined Incongruent' condition (Mchange = −4.7, SD = 22.9) decreased perceived effectiveness by 7.74 points ($t(33.487) = 4.13$,

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Table 2. Study 2: GPT-developed codebook for qualitative data analysis

Category	Definition	Example Quote
Policy Impact on Public Safety	Evaluates policy effectiveness in reducing accidents, injuries, or harm to public safety.	'Protection of the people is important knowing fully well the damage over speeding could lead to' (Participant 70, ObvID 5, 20mph limit)
Government Overreach and Regulation Concerns	Criticism of policies perceived as excessive government control or limiting personal freedoms.	'I am concerned that this could be abused and private information is being looked at/monitored by authorities.' (Participant 145, ObvID 276, Client-side scanning)
Policy Effectiveness and Implementation Challenges	Questions the real-world success of policies, their enforceability, and unintended consequences.	'I agree with the reason why the policy is implemented, but I don't think it will be effective if it is not enforced (e.g., with speed bumps or speed cameras)—I live in a built-up area with narrow roads and on-road parking and people still speed around without a care in the world, it is dangerous, but I don't think it is effective without enforcement.' (Participant 85, ObvID 9, 20mph limit)
Environmental and Climate Concerns	Discusses policies aimed at reducing carbon emissions, energy use, and environmental sustainability.	'It will put people off buying beef, and meat production produces high levels of carbon emissions.' (Participant 199, ObvID 98, Meat/seafood tax)
Economic and Social Equity	Focuses on fairness, affordability, and disproportionate impacts on lower-income individuals or vulnerable groups.	'I don't think this would work safely as some of the people who need to use more energy are vulnerable.' (Participant 119, ObvID 65, Energy quota)

(*Continued*)

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Table 2. (*Continued.*)

Category	Definition	Example Quote
Consumer Behaviour and Personal Choice	Examines how policies influence or restrict individual decisions regarding consumption, habits, and lifestyle.	'I really don't see this making that big a difference. It might stop a few impulse purchases but it isn't going to change their whole lifestyle or diet.' (Participant 191, ObvID 448, Food merch restrictions)
Scientific Credibility and Data Reliability	Questions the accuracy, validity, and sources of research supporting policies.	'Please explain how they extrapolate the saving of 100,000 hospitalisations from that statement. Made up number.' (Participant 27, ObvID 1976, Alcohol Health Labels)
Corporate and Systemic Responsibility	Argues that large corporations or systemic issues should be targeted rather than individuals for policy enforcement.	'The vast majority of carbon emissions are from big multi-national companies, not by the individual. This punishes the individual for the decisions they aren't making.' (Participant 101, ObvID 708, Meat/seafood tax)
Public Health and Well-being	Supports policies that improve public health, prevent harm, and reduce healthcare costs.	'I think this would address public health issues to some degree especially second-hand smoking/vaping, however, people would still vape discreetly and would not always follow the rules.' (Participant 93, ObvID 224, E-cigarette ban)
Scepticism and Distrust of Authorities	Expresses doubt about the government's intentions, policy motivations, and the influence of external entities.	'I agree that scanning messages could have an important role in tackling major online issues, but I'm not sure I trust technology companies to handle our data securely.' (Participant 116, ObvID 874, Client-side scanning)

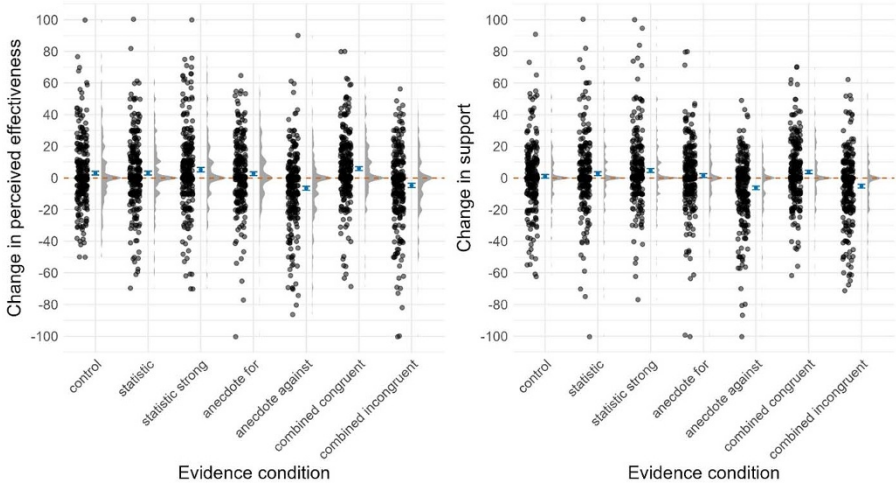


Figure 5. Study 2: The plot shows the change in perceived effectiveness and support per evidence condition. The points to the left of the blue mean and standard error bars show a participant's policy judgement for one of the six policies. The points are filled with a transparent grey, making overlapping points appear darker and distinguishing areas with high concentrations of data points. The density distribution to the right of the boxplot shows the data distribution. The Orange dashed line at zero represents no change in the policy judgements.

$p < .001$, $d = .29$) compared to control, and 7.80 points ($t(33.496) = 4.17$, $p < .001$, $d = .37$) compared to the 'Statistic' condition. The 'Combined Incongruent' condition (Mchange = −5.1, SD = 20.6) also decreased support by 6.12 points ($t(34.290) = 3.02$, $p = .005$, $d = .31$) compared to control and 7.84 points ($t(34.288) = 386$, $p < .001$, $d = .40$) compared to the 'Statistic' condition. This suggests that combining statistical and negative anecdotal evidence does not mitigate the effect of negative anecdotes alone on perceived effectiveness and support, which contrasts with results from Study 1.

Qualitative results

Similar to Study 1, Figure 6 shows that most participants discussed policy qualities such as fairness (e.g., economic and social equity), enforcement and compliance (e.g., implementation concerns), and freedom of choice (e.g., personal autonomy). The frequency of these topics varied by policy. For example, fairness concerns were more frequently mentioned for the meat/seafood tax, quotas, and food merchandising restrictions than for other policies. A new topic identified in Study 2 was concerns about government overreach, which was most frequently mentioned in relation to client-side scanning (a newly introduced policy) compared to all other policies (see Figure B5).

Similar to Study 1, Figure B6 shows that more participants in lower quantiles of baseline perceived effectiveness and support mentioned topics with negative sentiment towards policies, such as concerns with fairness, government overreach, and distrust of authorities, than higher quantiles. In contrast, more participants in higher quantiles mentioned positive topics, such as improvements to public safety and health. However,

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

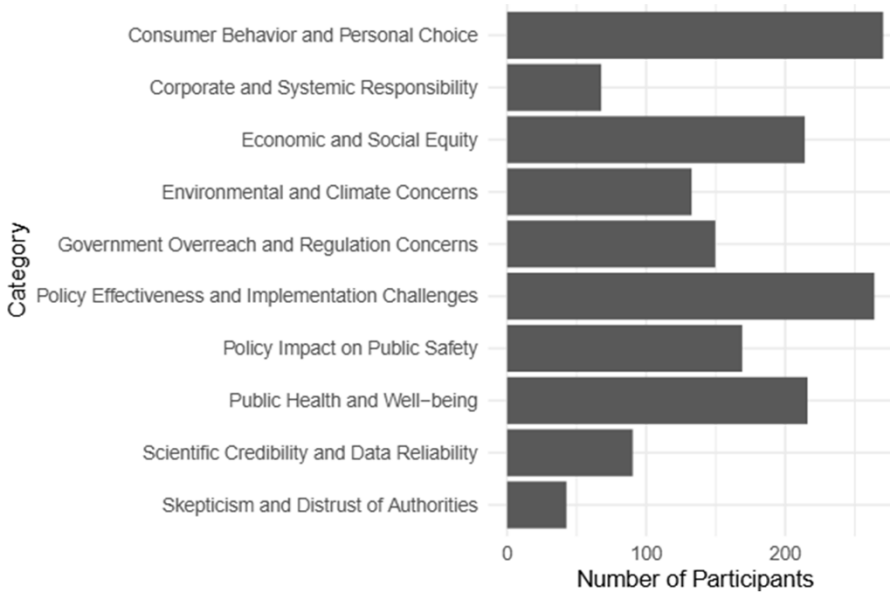


Figure 6. Number of participants mentioning each category in their responses (Study 2).

some topics, such as personal choice and implementation challenges, were mentioned consistently across support levels.

In Study 2, approximately 90 participants mentioned concerns about scientific credibility and data reliability. Figure B7 shows that these concerns were most common in conditions featuring statistical evidence (e.g., Statistic, Statistic Strong, and Combined). This scepticism may partly explain the lack of effect observed for statistical evidence in this study.

Discussion

In line with Study 1, evidence supporting policy effectiveness, regardless of type, did not increase perceived effectiveness and support, whereas negative anecdotes consistently reduced both. The asymmetry in the effect of anecdotal evidence persisted even after balancing the anecdotal arguments for and against policy effectiveness. The introduction of the 'Strong Statistic's condition had no additional effect on perceived effectiveness and support compared to the original 'Statistic' condition, aligning with prior research (Reynolds *et al.*, 2018). The limited effect of statistical evidence may not be due to insufficient strength but may reflect a broader challenge in leveraging statistical evidence to influence policy judgements. A key difference from Study 1 was that combining negative anecdotes with statistical evidence did not mitigate the negative anecdote effect. This finding warrants further investigation, as determining whether statistical evidence can reliably mitigate this effect is important for improving evidence-based policy communication. Finally, the qualitative results align with Study 1, indicating that participants consider both policy-specific issues (e.g., effectiveness,

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

fairness, compliance) and broader beliefs (e.g., personal choice, the role of government) when evaluating policy evidence.

Study 3

In Study 3, we again aimed to replicate the negative anecdote effect and examine the effect of 'Combined Incongruent' evidence, given the mixed findings across the previous studies. We also aimed to understand how people evaluate policy evidence by examining whether participants updated their policy judgments in line with Bayesian models of rational belief updating across the different evidence types.

Bayesian models predict that individuals adjust their beliefs in proportion to the diagnostic value of new evidence, systematically integrating information based on its credibility and relevance. To assess this, we asked participants to judge how likely different types of evidence would be if a given policy were truly effective. If participants believed that statistical evidence showing reductions in carbon emissions would be highly likely under an effective meat and seafood tax but unlikely under an ineffective one, this would indicate that statistical evidence holds high diagnostic value. Under a Bayesian model, such evidence should increase participants' perceived effectiveness of the policy. We compared participants' belief updates to Bayesian predictions.

While individuals generally update beliefs in the expected direction, they often do so more conservatively than Bayesian predictions (Kovach, 2021). Therefore, we examined whether participants' change in perceived effectiveness was qualitatively (i.e., in the right direction) and quantitatively (i.e., in the right direction and magnitude) consistent with Bayesian predictions.

Method

Participants

We recruited 307 participants. The participants had a mean age of 46.09 years (SD = 16.13). Of the sample, 51% identified as women, 48% male, 87% as white, 46% were employed full-time and 51% held a Bachelor's degree or higher. Politically, 48% were affiliated with the Labour Party and 18% with the Conservative Party. For more details, see Table A1.

Materials and measures

Policy evidence. We used the same stimuli as in Study 2, except we replaced the 'Strong Statistic' condition with a 'Combined Incongruent' (Anecdotes) condition, which presented excerpts from both 'Anecdote For' and 'Anecdote Against' conditions.

Measures. We used the same measures for perceived effectiveness and support as in previous studies. Participants also rated the evidence's diagnostic value, in all but the control condition, by providing two conditional probability judgements, with items adapted from (Shengelia and Lagnado, 2021):

1. The probability of the evidence (E) given a policy is effective (H): P(E|H).
2. The probability of the evidence given a policy is not effective (H): P(E|˜H).

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

In the 'Statistic' condition, participants were asked: 'If the policy IS effective, how likely is it that research would show these results?' followed by the same question but with 'If the policy IS NOT effective…' In both 'Anecdote For' and 'Anecdote Against' conditions, they were asked: 'If the policy IS effective, how likely is it that someone would make this claim about the policy?' and again followed by 'If the policy IS NOT effective…' In the Combined conditions, participants rated each piece of evidence separately using the same questions outlined above, depending on whether the evidence was a statistic or anecdote. These measures were adapted from research on Bayesian belief updating in legal domain.

Procedure

Study 3 took an average of 15 minutes, and participants were paid £2.55. The full Qualtrics survey is available on OSF.

The procedure matched Study 1 except in Step 3, Part ii. While viewing the evidence, participants completed the likelihood of evidence and posterior policy judgment measures. The order of these measures was randomised, with participants answering either the likelihood of evidence or posterior policy judgment measure first.

Data analysis

We followed the preregistered analysis plan in our study protocols (see OSF), using the same models for the effect of evidence condition on perceived effectiveness and support in the previous studies.

For the Bayesian belief updating analysis, we calculated the mean log-likelihood ratio for each evidence condition using conditional probability judgments. A log-likelihood ratio > 0 indicates evidence increases the probability that the policy is effective (H), < 0 indicates it decreases this probability, and $= 0$ indicates no diagnostic value. For the 'Statistic', 'Anecdote For', and 'Anecdote Against' conditions, the likelihood ratio was $P(E \mid H)/P(E \mid \tilde{}H)$. For the Combined conditions, it was $[P(E1 \mid H)/P(E1 \mid \tilde{}H)] \times [P(E2 \mid H)/P(E2 \mid \tilde{}H)]$. Log transformations were applied to assess the diagnostic value of different evidence types. We also calculated each participant's predicted posterior belief in policy effectiveness by inputting their conditional probability judgments and prior belief into Bayes' theorem and compared this to people's actual posterior beliefs.

Results

Tables B5 and B6 (SM) summarise the models testing the effects of evidence conditions on change in perceived effectiveness and support (see Figure 7 and Figure 3).

In line with the previous studies, the 'Statistic', 'Anecdote For', and 'Combined Congruent' conditions showed changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions.

We only partially replicated the negative anecdote effect observed in prior studies. The 'Anecdote Against' condition (Mchange = 0.2, SD = 27) decreased perceived effectiveness by 2.45 points compared to control (Mchange = 2.6, SD = 21.6) and 6.06 points compared to the 'Statistic' condition (Mchange = 6.3, SD = 23.2), but

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

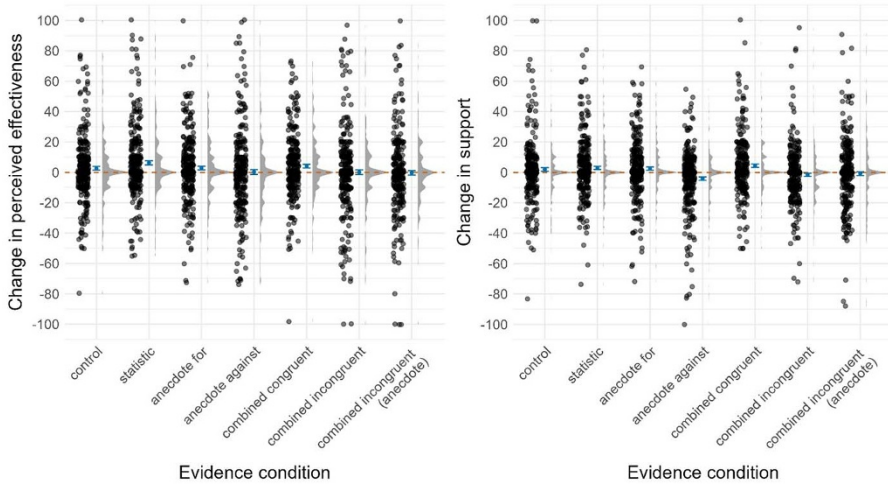


Figure 7. Study 3: The plot shows the change in perceived effectiveness and support per evidence condition. The points to the left of the blue mean and standard error bars show a participant's policy judgement for one of the six policies. The points are filled with a transparent grey, making overlapping points appear darker and distinguishing areas with high concentrations of data points. The density distribution to the right of the boxplot shows the data distribution. The Orange dashed line at zero represents no change in the policy judgements.

these differences were not significant given the adjusted significance threshold of $p < .005$ ($t(33.657) = 1.09$, $p = .282$, $d = .10$, and $t(33.657) = 2.71$, $p = .011$, $d = .26$, respectively). However, the 'Anecdote Against' condition (Mchange = – 4.1, SD = 19.6) significantly decreased support by 6.05 points ($t(33.812) = 3.30$, $p = .002$, $d = .32$) compared to control (Mchange = 1.9, SD = 20.3) and 7.04 points ($t(33.812) = 3.84$, $p = .001$, $d = .37$) compared to the 'Statistic' condition (Mchange = 2.9, SD = 18.8).

For the change in perceived effectiveness, we did not replicate the negative effect of the 'Combined Incongruent' condition in Study 2. This condition showed changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions. In light of the lack of negative anecdote effect, this result cannot inform whether combining statistical and negative anecdotal evidence either does (Study 1) or does not (Study 2) mitigate the effect of negative anecdotes alone on perceived effectiveness. However, it does seem to mitigate the negative anecdote effect for policy support.

The 'Combined Incongruent' (Anecdote) condition had similar results to the 'Combined Incongruent' condition with changes in perceived effectiveness and support comparable to the 'Control' and 'Statistic' conditions. This suggests either statistical or anecdotal evidence supporting policy effectiveness may mitigate the negative anecdote effect for policy support.

Figures B1 and B2 show a consistent effect of conditions across policies and studies, with some exceptions. Notably, perceived effectiveness and support did not decrease in the negative anecdote condition for the 20 mph speed limit and client-side scanning.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Bayesian belief updating

We assessed the diagnostic value of each evidence type (see Figure 8). Log-likelihood ratios for each condition suggest participants generally evaluated the evidence as expected based on prior research: statistical and anecdotal evidence supporting the policy should increase belief in its effectiveness, whereas combining these two types of evidence (in the 'Combined Congruent' condition) has a greater increase than either type alone. In contrast, anecdotal evidence against the policy should decrease belief in its effectiveness. For the 'Combined Incongruent' condition, the log-likelihood ratios should increase belief in policy effectiveness, though this increase was smaller than that of statistical and supportive anecdotal evidence. However, these results are harder to interpret due to limited research and mixed findings on the impact of conflicting evidence.

Figure 9 illustrates that belief updating did not fully align with Bayesian predictions. While the 'Statistic', 'Anecdote For' and 'Combined Congruent' conditions showed changes in the expected direction, these changes were consistently smaller than predicted. The 'Anecdote Against' and 'Combined Incongruent' conditions showed belief changes around zero, deviating from the predicted changes. This pattern is known as conservatism (i.e., participants updating less than expected or not updating at all) and is a common result in this type of modelling (Corner, 2010). Table B7 in the supplemental materials shows that despite discrepancies, predicted and actual posterior beliefs were moderately positively correlated in each condition, with $_r$ ranging from 0.47 to 0.67.

To assess whether the discrepancies between actual and predicted changes were driven by conservatism, we classified participants based on the magnitude and direction of their actual belief update compared to the predicted. Figure 10 shows that while most participants updated as predicted or in the expected direction, conservatively, a notable proportion updated in the opposite direction to their likelihood ratio. Additionally, a small proportion of participants changed their belief even though they did not rate the evidence as diagnostic of the policy's effectiveness or ineffectiveness.

Discussion

We conducted three online experiments with representative UK samples to test whether evidence on effectiveness changes perceived policy effectiveness and support. The findings suggest there was no evidence that evidence supporting policy effectiveness – whether statistical, anecdotal or both – increased perceived effectiveness or support compared to simply stating the policy aimed to address a societal issue. In contrast, negative anecdotes significantly reduced perceived effectiveness (though this effect was not replicated in Study 3) and support (replicated in all three studies). These results challenge an earlier meta-analysis (Reynolds *et al.*, 2020) suggesting that evidence supporting policy effectiveness increases support, but aligns with its finding that evidence against policy effectiveness can lower support. Our results also align with recent research showing evidence supporting public health food policy does not increase support (e.g., reducing portion sizes, Reynolds *et al.*, 2022), suggesting that evidence of policy effectiveness may not always be an effective communication strategy.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

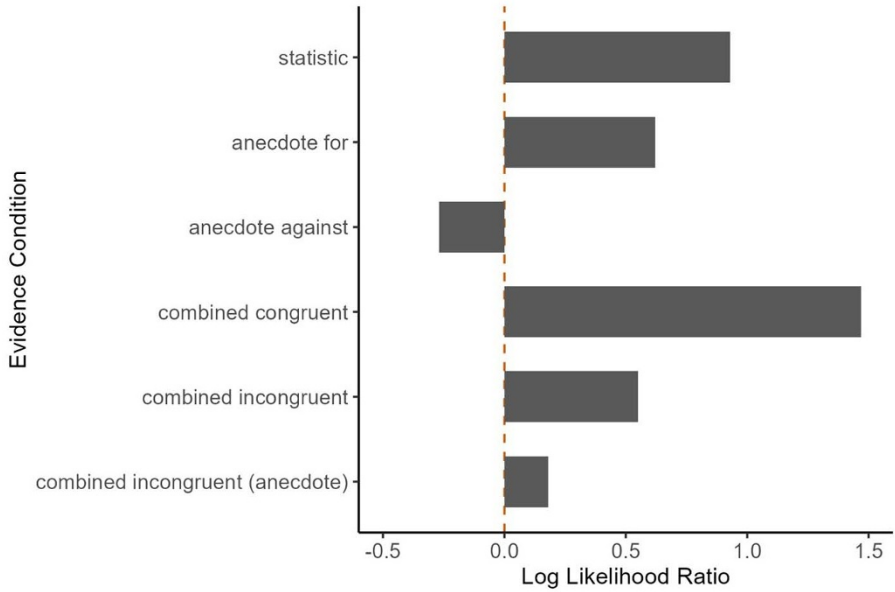


Figure 8. Log Likelihood Ratio per Evidence Condition. Note. The plot shows the mean log-likelihood ratio for each evidence condition. A log-likelihood ratio > 0 indicates that the evidence (E) increases the posterior probability that the policy is effective (H); < 0 indicates that the evidence decreases the posterior probability of H; and 0 indicates no probative value.

The negative anecdote results suggest that anecdotal evidence may be more persuasive than statistical evidence in policy communication. However, we found no evidence that positive anecdotes increased outcomes more than statistical evidence, even on health-related topics where anecdotes are typically more impactful (Freling *et al.*, 2020; Xue and Taylor, 2023). This asymmetry suggests that while negative anecdotes may undermine policy support, positive anecdotes may not have the opposite effect. This finding reflects a common psychological principle: negative events, cues and memories influence outcomes more than positive ones, affecting cognition, motivation and social interactions (Baumeister *et al.*, 2001). While theories suggest negative information is more diagnostic or salient, they don't fully explain the phenomenon. Baumeister et al. suggest that this may have adaptive benefits, helping individuals avoid threats and increase survival. Recent studies on framing reinforce this pattern in policy decisions, revealing that negative framing triggers stronger emotional responses and makes problems seem more urgent and deserving of government intervention, even when information is logically equivalent (Sahn *et al.*, 2025). This finding also presents a notable challenge for effective policy communication. If the effects of negative anecdotes on public support are replicated in real-world settings, the impact could be substantial. Compared to statistical evidence, misinformation presented with anecdotal evidence is perceived as less false and less likely to prompt verification behaviours (Zhao and Tsang, 2024). Negative news is also more widely read and shared online (Oh *et al.*, 2023).

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

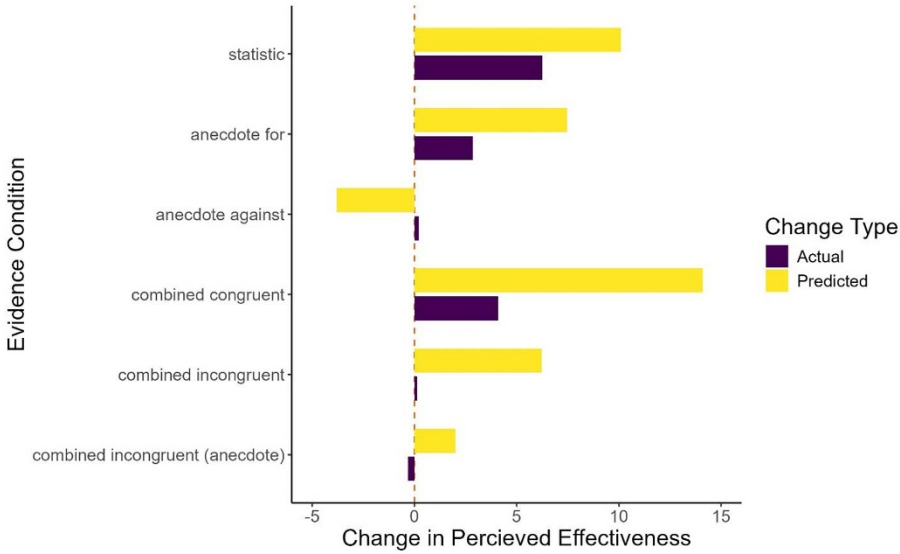


Figure 9. Predicted versus Actual Change in Belief in Policy Effectiveness per Evidence Condition. Note. The plot shows the mean change in belief in policy effectiveness for each evidence condition. The bars are coloured to indicate whether the change represents the actual change observed in the sample (yellow) or the predicted change.

In most studies, there was no evidence that combining statistical and negative anecdotal evidence altered perceived effectiveness or support. This suggests that while statistical evidence alone may not increase these outcomes, it could help mitigate the negative anecdotal effect. Yet, findings were mixed: study 2 showed that the negative anecdote effect persisted even when paired with statistical evidence supporting policy effectiveness. This aligns with research on public decision-making about medical treatments, where negative anecdotes reduced belief in effectiveness and willingness to adopt treatments despite statistical evidence supporting the treatments (Line *et al.*, 2024).

Although we did not directly investigate the mechanisms behind the effects, related research offers some explanations, but no single theory fully explains the findings. We highlight key mechanisms with limitations and recommend further research to understand why the negative anecdote effect occurs.

One possible explanation is that emotional language or stories can increase emotional engagement and reduce counterarguments (Ecker *et al.*, 2020; Line *et al.*, 2024). However, post-hoc sentiment analysis (see Appendix B; Figures B8 and B9) indicated that, although anecdotes varied in emotional tone as expected (i.e., anecdotes against the policy having a negative sentiment and anecdotes for the policy having a positive one), most language within all of the anecdotes was classified as neutral. Additionally, the positive sentiment in supportive anecdotes had a higher compound score than the negative sentiment in opposing anecdotes, while both had compound scores close to zero, which also indicates a neutral overall framing. This suggests that emotional intensity alone doesn't explain the asymmetry effects.

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

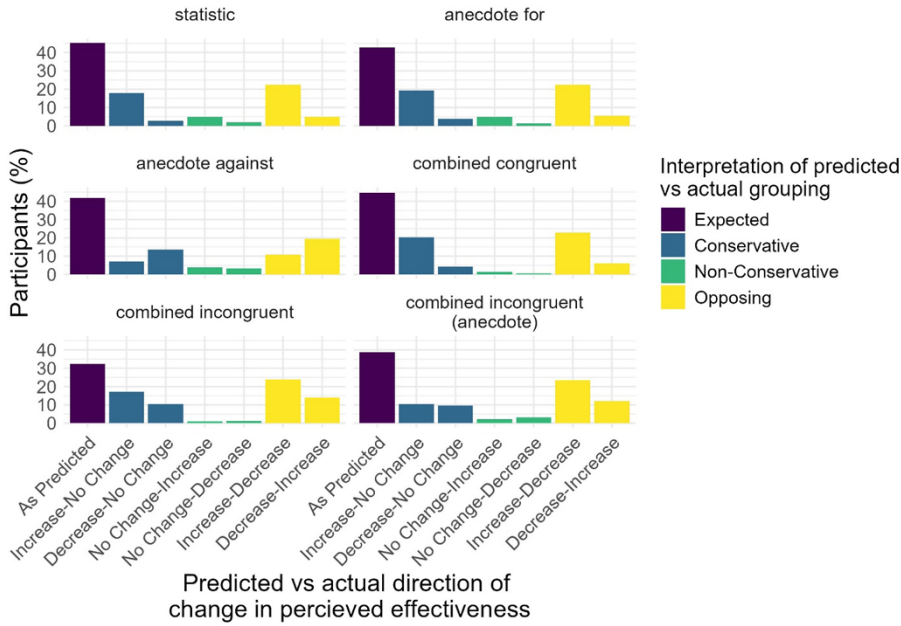


Figure 10. Predicted versus actual direction and magnitude of Change in Perceived Effectiveness per Evidence Condition. Note. The plot displays all possible combinations of predicted versus actual direction of change in belief in policy effectiveness on the x-axis. These combinations are named with the predicted direction followed by the actual direction. For example, 'Increase-No Change' represents a predicted increase in belief but no actual change observed. 'As Predicted' represents that the actual change aligned with the predicted direction and magnitude. The bars indicate the percentage of participants within each combination, and their colours reflect the interpretation of each grouping.

An explanation based on motivated reasoning suggests participants were more influenced by negative evidence because it aligned with their beliefs and identities (Taber and Lodge, 2006; Kahan, 2017; Druckman and McGrath, 2019). For instance, the negative anecdote effect may reflect confirmation bias (Oeberst and Imhoff, 2023), where participants were more influenced by negative evidence because it aligned with their prior beliefs. This explanation also corresponds with research indicating that decisions on educational interventions are often driven more by prior beliefs than evidence quality (Michal *et al.*, 2021), and corrective information rarely changes political beliefs (Nyhan and Reifler, 2010). However, this has limitations; not all policies had low initial support (e.g., banning e-cigarettes), yet negative anecdote effects appeared across policies and participants in the top quantiles of prior policy support.

Alternatively, these patterns could be explained by framing effects, where anecdotal conditions activate different cognitive frames influencing policy considerations like personal freedom versus collective responsibility or government effectiveness versus overreach (Chong and Druckman, 2007). These frames may influence which beliefs are accessible and relevant when participants assess the evidence. However, Study 2's results challenge this: even with balanced anecdotes addressing the same aspects of policy effectiveness, asymmetrical effects persisted. This suggests differences arise from

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

how positive and negative information is processed, not from different framing due to content.

Finally, from a rational updating perspective – where people objectively weigh new evidence against their prior beliefs to form updated opinions – our Bayesian modelling in Study 3 suggests that most participants adjusted their policy beliefs in the expected direction but conservatively, based on how strongly they believed the evidence indicated policy effectiveness. This conservatism in belief updating is common and often linked to scepticism about evidence reliability in experimental settings (Corner, 2010; Howe *et al.*, 2022). However, our simple Bayesian model did not incorporate parameters such as prior belief strength (Druckman and McGrath, 2019), which may have explained the conservatism and other deviations from the predictions, limiting conclusions about rational updating processes and limiting our ability to rule out motivated reasoning.

Overall, our quantitative findings can be explained by multiple and sometimes competing theoretical perspectives, highlighting the need for future research to robustly test proposed mediators through experimental manipulation and establish causal pathways.

Our qualitative results suggest that participants integrate broader beliefs (e.g., regarding personal freedom and government's role) and policy-specific concerns (e.g., fairness and enforcement) when evaluating policy evidence. These issues were more frequently cited by participants with lower baseline perceived effectiveness and support, varying by policy (e.g., fairness was mentioned by more participants for taxes than for alcohol health labels). While aligning with recent theories on how general and policy-specific beliefs shape policy support (Huber *et al.*, 2020; Grelle and Hofmann, 2024), our findings also highlight additional considerations, such as implementation and enforcement concerns, which may have been overlooked. These results also highlight the potential complexity of reasoning and belief updating in policy communication, where evidence evaluation is likely influenced by factors such as prior beliefs, emotions and knowledge (Sommer *et al.*, 2023).

Policy implications

Consistent with prior surveys, there was limited support for environmental policy (OECD, 2023). However, public health policies, such as banning e-cigarettes in public places, received greater support, suggesting these types of measures may be currently acceptable to the UK public (Laverty *et al.*, 2017). For online safety, findings indicate a neutral to slightly positive stance on client-side scanning under the Online Safety Act, consistent with studies showing public support for such measures in the USA and Germany to tackle online child sexual exploitation and abuse despite privacy concerns (Geierhaas *et al.*, 2023; Deldari *et al.*, 2024).

To improve public support, communicating policy effectiveness alone is likely insufficient, though it may help counteract negative narratives. Our qualitative data suggest that strategies should also address policy qualities like fairness and enforcement, as well as broader beliefs about government intervention infringing personal freedom (Reynolds *et al.*, 2018). While these factors were mentioned across policies and individuals, policy-specific approaches informed by public engagement may help develop

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

effective communication strategies, particularly for non-supporters. This is evident in attitudes towards client-side scanning, where support exists despite privacy concerns only when framed as a tool against child abuse or terrorism, but not other illegal activities (Geierhaas *et al.*, 2023).

Limitations, strengths & future directions

We tested different evidence types using a highly powered, representative UK sample, focusing on policies addressing major societal challenges. Our within-subjects design extended prior between-subjects research by exposing participants to multiple evidence types across various policies, better reflecting real-world contexts where people encounter diverse policy issues and evidence. We also expanded conditions beyond statistical and anecdotal evidence to include combined evidence, enabling direct comparisons.

However, our anecdotal stimuli differed not only in argument direction but potentially in emotional tone, with anecdotes opposing the policy likely containing more negative language than neutral statistical evidence. While this confounds evidence type with emotional content, it reflects the ecological validity of real-world policy communication where supportive and opposing arguments naturally employ different emotional registers. Post-hoc sentiment analysis confirmed directional differences but showed these were moderate rather than extreme (see Figure B8). Additionally, the anecdotal evidence conditions may have inadvertently contained content that activated broader ideological beliefs (e.g., personal freedom, government overreach, fairness), creating a confound between evidence type and ideological content that we did not fully control for. While we included a control condition providing only policy rationale, separating evidence format effects from ideological priming effects would require either a neutral anecdote condition with matched ideological content across evidence types or a more complex experimental design manipulating these factors independently.

Future research can improve the validity of this design, for example, by replicating our study in more naturalistic settings such as online social media environments to assess the broader applicability of our findings, as well as evaluating the potential confounds between argument direction and other factors such as emotional language and ideological priming across evidence types. Additionally, since we omitted a negative statistic condition, it remains unclear whether negative statistics would have similar effects as negative anecdotes. Future research should test whether evidence type (statistic vs anecdote) or sentiment (supporting vs opposing policy effectiveness) plays a greater role in shaping policy support. Finally, improvements can be made to how to perceived policy effectiveness and policy support are measured. For instance, our perceived policy effectiveness measure combined policy efficacy assessments ('X is effective for Y') and mechanism explanations ('as X will change Z') in single items, potentially introducing measurement error if participants held different views about these components.

While our samples were representative of the UK voting population during the study period in terms of age, gender and political affiliation, other demographic measures were not representative. Our sample was majority white, had higher education levels than the general population, and included limited representation from devolved

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

nations (Scotland, Wales, Northern Ireland), which may limit generalisability to the broader UK voting population. Additionally, the sample was mostly affiliated with the Labour Party, reflecting the political landscape after the last election, which is already changing. We note that attrition rates for participants who started but did not complete the survey and were therefore not compensated or included in the analysis, as per Prolific's guidelines, ranged from 4% to 8% across studies. While this did not affect the representative quotas for age, gender and political affiliation, we cannot determine whether attrition was non-random by assessing whether attrition correlated with other demographics or study characteristics.

Replicating these results is crucial for understanding their robustness and broader applicability, particularly given their divergence from recent meta-analyses. While some trends were consistent (e.g., negative anecdotal evidence reducing policy support), other effects (e.g., positive statistical evidence mitigating the negative anecdote effect) were not. Further research should continue to test the effect of different combinations of evidence on policy support using a wider variety of materials. Beyond replication, it is essential to investigate the underlying mechanisms driving these effects, for example by building and testing causal models to examine theorised mechanisms such as confirmation bias (Sunstein *et al.*, 2016; Michal *et al.*, 2021) or emotional engagement (Ecker *et al.*, 2020; Line *et al.*, 2024).

ChatGPT-assisted content analysis proved reliable for capturing descriptive qualitative results on how people evaluate policy interventions. Given the brief nature of the responses (averaging 17–30 words), deeper analysis of how participants reasoned about each policy and evidence type was impractical. Most participants provided single-sentence responses, making it difficult to identify complex reasoning patterns. Future research should employ more in-depth qualitative methods, such as think-aloud studies, to better understand the reasoning processes underlying policy evidence evaluation and whether observed effects reflect motivated reasoning, emotional processing or other mechanisms. Additionally, future research should apply Bayesian modelling more robustly to assess whether belief updating is rational or motivated, informing strategies to influence policy support (Druckman and McGrath, 2019).

Conclusion

Communicating policy effectiveness may not always increase public support. While negative anecdotes consistently reduced support, there was no evidence that positive evidence – statistical or anecdotal – affected support. However, statistical evidence may help counter negative narratives. To improve public acceptance, policy communication should address broader public concerns beyond effectiveness, such as fairness.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/bpp.2025.10028

Funding statement. This work was supported by the Economic and Social Research Council [Grant number ES/Y001044/1].

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

References

Appelman, A. and S. S. Sundar (2016), 'Measuring message credibility: construction and validation of an exclusive scale', *Journalism & Mass Communication Quarterly*, **93**(1): 59–79. SAGE Publications Inc. doi:10.1177/1077699015606057

Bago, B., D. G. Rand and G. Pennycook (2023), 'Reasoning about climate change', *PNAS Nexus*, **2**(5): pgad100. doi:10.1093/pnasnexus/pgad100

Bates, D., M. Maechler, B. Bolker and S. Walker (2024) *Lme4: Linear mixed-effects models using Eigen and S4*. Version 1.1–35.3. Available at: https://cran.r-project.org/web/packages/lme4/index.html [Accessed 23 April 2024].

Baumeister, R. F., E. Bratslavsky, C. Finkenauer and K. D. Vohs (2001), 'Bad is Stronger than Good', *Review of General Psychology*, **5**(4): 323–370. doi:10.1037/1089-2680.5.4.323

Betsch, C., F. Renkewitz, T. Betsch and C. Ulshöfer (2011), 'The influence of narrative vs. statistical information on perceiving vaccination risks', *Medical Decision Making*, **31**(5): 742–753. SAGE Publications Inc STM. doi:10.1177/0272989X11400419

Bigman, C. A., J. N. Cappella and R. C. Hornik (2010), 'Effective or ineffective: attribute framing and the human papillomavirus (HPV) vaccine', *Patient Education and Counseling*, **81**(Suppl): S70–S76. doi:10.1016/j.pec.2010.08.014

Bijker, R., S. Van de Walle and B. Voorn (2024), 'ChatGPT for automated qualitative research: content analysis', *Journal of Medical Internet Research*, **26**(1): e59050. doi:10.2196/59050

Burstein, P. (2003), 'The impact of public opinion on public policy: a review and an agenda', *Political Research Quarterly*, **56**(1): 29–40. doi:10.2307/3219881

Burstein, P. (2020), 'The determinants of public policy: what matters and how much', *Policy Studies Journal*, **48**(1): 87–110. doi:10.1111/psj.12243

Carroll, P., B. Caulfield and A. Ahern (2019), 'Measuring the potential emission reductions from a shift towards public transport', *Transportation Research Part D: Transport and Environment*, **73**: 338–351. doi:10.1016/j.trd.2019.07.010

Chong, D. and J. N. Druckman (2007), 'Framing theory', *Annual Review of Political Science*, **10**: 103–126. doi:10.1146/annurev.polisci.10.072805.103054

Corner, A. (2010) 'Conservatism in belief revision and participant skepticism', in *Cognition in Flux: Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, 32.

Deldari, E., P. Thakkar and Y. Yao (2024), 'Users' perceptions of online child abuse detection mechanisms', *Proceedings of the ACM on Human-Computer Interaction*, **8**(CSCW1): 1–26. 147:1–147:26. doi:10.1145/3637424

Druckman, J. N. and M. C. McGrath (2019), 'The evidence for motivated reasoning in climate change preference formation', *Nature Climate Change*, **9**(2): 111–119. Nature Publishing Group. doi:10.1038/s41558-018-0360-1

Ecker, U. K. H., L. H. Butler and A. Hamby (2020), 'You don't have to tell a story! A registered report testing the effectiveness of narrative versus non-narrative misinformation corrections', *Cognitive Research: Principles and Implications*, **5**(1): 64–90. doi:10.1186/s41235-020-00266-x

Freling, T. H., Z. Yang, R. Saini and O. S. Itani (2020), 'When poignant stories outweigh cold hard facts: a meta-analysis of the anecdotal bias', *Organizational Behavior and Human Decision Processes*, **160**: 51–67. doi:10.1016/j.obhdp.2020.01.006

Geierhaas, L., C. Hammer, H. Hermanns and K. Kuehne (2023), 'Attitudes towards client-side scanning for CSAM, terrorism, drug trafficking, drug use and tax evasion in Germany', *2023 IEEE Symposium on Security and Privacy (SP)*, 217–233. doi:10.1109/SP46215.2023.10179417

Graeber, T., C. Roth and F. Zimmermann (2024), 'Stories, Statistics, and Memory', *The Quarterly Journal of Economics*, **139**(4). doi:10.1093/qje/qjae020.

Green, P. and C. J. MacLeod (2016), 'SIMR: an R package for power analysis of generalized linear mixed models by simulation', *Methods in Ecology and Evolution*, **7**(4): 493–498. doi:10.1111/2041-210X.12504

Grelle, S. and W. Hofmann (2024), 'When and why do people accept public-policy interventions? An integrative public-policy-acceptance framework', *Perspectives on Psychological Science*, **19**(1): 258–279. SAGE Publications Inc. doi:10.1177/17456916231180580

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Guan, M., S. He, J. Z. Yang and C. Zeng (2023), 'COVID-19 message fatigue: how does it predict preventive behavioral intentions and what types of information are people tired of hearing about?', *Health Communication*, **38**(8): 1631–1640. Routledge. doi:10.1080/10410236.2021.2023385

Hinnant, A., S. Hu and E. Ash (2023), 'Contested certainty and credibility: the effect of personal stories and scientific evidence in user comments on news story evaluation and relevance', *Science Communication*, **45**(1): 65–94. SAGE Publications Inc. doi:10.1177/10755470221150503

Hinnant, A., R. Subramanian and R. Young (2016), 'User comments on climate stories: impacts of anecdotal vs. scientific evidence', *Climatic Change*, **138**(3): 411–424. doi:10.1007/s10584-016-1759-1

Hornikx, J. (2018), 'Combining anecdotal and statistical evidence in real-life discourse: comprehension and persuasiveness', *Discourse Processes*, **55**(3): 324–336. Routledge. doi:10.1080/0163853X.2017.1312195

Howe, P. D. L., A. Tadmor, N. Gurney, J. Baxter, J. C. Jackson and S. Lewandowsky (2022), 'Base rate neglect and conservatism in probabilistic reasoning: insights from eliciting full distributions', *Judgment and Decision Making*, **17**(5): 962–987. doi:10.1017/S1930297500009281

Huber, R. A., M. L. Wicki and T. Bernauer (2020), 'Public support for environmental policy depends on beliefs concerning effectiveness, intrusiveness, and fairness', *Environmental Politics*, **29**(4): 649–673. Routledge. doi:10.1080/09644016.2019.1629171

Iles, I. A., L. Feldman, B. L. Quick, S. S. Ho and M. D. Barnes (2022), 'Understanding responses to different types of conflicting information about cancer prevention', *Social Science & Medicine*, **311**: 115292. doi:10.1016/j.socscimed.2022.115292

Kahan, D. M. (2017). Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition (SSRN Scholarly Paper 2973067). Social Science Research Network. doi:10.2139/ssrn.2973067

Kovach, M. (2021) 'Conservative updating', *arXiv preprint arXiv:2102.00152*. Available at: https://arxiv.org/abs/2102.00152.

Laverty, A. A., F. T. Filippidis, C. I. Vardavas, E. Patelarou, C. Gratziou and G. N. Connolly (2017), 'E-cigarette use and support for banning e-cigarette use in public places in the European Union', *Preventive Medicine*, **105**: 10–14. doi:10.1016/j.ypmed.2017.08.007

Line, E. N., J. Chen, S. Patel and T. Johnson (2024), 'Anecdotes impact medical decisions even when presented with statistical information or decision aids', *Cognitive Research: Principles and Implications*, **9**(1): 51–74. doi:10.1186/s41235-024-00577-3

Mantzari, E., J. P. Reynolds, S. A. Jebb, G. J. Hollands, M. A. Pilling and T. M. Marteau (2022), 'Public support for policies to improve population and planetary health: a population-based online experiment assessing impact of communicating evidence of multiple versus single benefits', *Social Science & Medicine*, **296**: 114726. doi:10.1016/j.socscimed.2022.114726

Michal, A. L., Y. Zhong and P. Shah (2021), 'When and why do people act on flawed science? Effects of anecdotes and prior beliefs on evidence-based decision-making', *Cognitive Research: Principles and Implications*, **6**(1): 28–51. doi:10.1186/s41235-021-00293-2

Nagler, R. H., C. A. Bigman, J. N. Cappella and R. C. Hornik (2022), 'Effects of prior exposure to conflicting health information on responses to subsequent unrelated health messages: results from a population-based longitudinal experiment', *Annals of Behavioral Medicine*, **56**(5): 498–511. doi:10.1093/abm/kaab069

Niederdeppe, J., S. Roh and C. Dreisbach (2016), 'How narrative focus and a statistical map shape health policy support among state legislators', *Health Communication*, **31**(2): 242–255. Routledge. doi:10.1080/10410236.2014.998913

Nyhan, B. and J. Reifler (2010), 'When corrections fail: the persistence of political misperceptions', *Political Behavior*, **32**(2): 303–330. doi:10.1007/s11109-010-9112-2

Oeberst, A. and R. Imhoff (2023), 'Toward parsimony in bias research: a proposed common framework of belief-consistent information processing for a set of biases', *Perspectives on Psychological Science*, **18**(6): 1464–1487. SAGE Publications Inc. doi:10.1177/17456916221148147

OECD (2023), *How Green Is Household Behaviour?: Sustainable Choices in a Time of Interlocking Crises. OECD Studies on Environmental Policy and Household Behaviour*, Paris: OECD Publishing. doi:10.1787/2bbbb663-en

Oh, H., K.-Y. Goh and T. Q. Phan (2023), 'Are you what you tweet? The impact of sentiment on digital news consumption and social media sharing', *Information Systems Research*, **34**(1): 111–136. doi:10.1287/isre.2022.1112

OpenAI (2025) *ChatGPT-4: Language model developed by OpenAI*. Available at: https://openai.com/chatgpt [Accessed 13 January 2025].

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Pechey, R., P. Burge, E. Mentzakis, M. Suhrcke and T. M. Marteau (2014), 'Public acceptability of population-level interventions to reduce alcohol consumption: a discrete choice experiment', *Social Science & Medicine (1982)*, **113**: 104–109. doi:10.1016/j.socscimed.2014.05.010

Petrescu, D. C., Hollands, G. J., Couturier, D.-L., Ng, Y.-L., & Marteau, T. M. (2016). Public Acceptability in the UK and USA of Nudging to Reduce Obesity: The Example of Reducing Sugar-Sweetened Beverages Consumption. *PLOS ONE*, **11**(6), e0155995. https://doi.org/10.1371/journal.pone.0155995

Promberger, M., P. Dolan and T. M. Marteau (2012), '"Pay them if it works": discrete choice experiments on the acceptability of financial incentives to change health-related behaviour', *Social Science & Medicine (1982)*, **75**(12): 2509–2514. doi:10.1016/j.socscimed.2012.09.033

Reynolds, J. P., A. Hobson, T. M. Marteau and M. Pilling (2022), 'Effect of visualising and re-expressing evidence of policy effectiveness on perceived effectiveness: a population-based survey experiment', *Behavioural Public Policy*, 1–19. doi:10.1017/bpp.2022.32

Reynolds, J. P., M. Pilling and T. M. Marteau (2018), 'Communicating quantitative evidence of policy effectiveness and support for the policy: three experimental studies', *Social Science & Medicine*, **218**: 1–12. doi:10.1016/j.socscimed.2018.09.037

Reynolds, J. P., K. Stautz, M. Pilling and T. M. Marteau (2020), 'Communicating the effectiveness and ineffectiveness of government policies and their impact on public support: a systematic review with meta-analysis', *Royal Society Open Science*, **7**(1): 190522. Royal Society. doi:10.1098/rsos.190522

Rodger, A. D., G. A. Sanna, V. Cheung, N. Raihani and D. Lagnado (2025 March 25), 'Demo project 2: negative anecdotes reduce policy support: evidence from three experimental studies on communicating policy (in)effectiveness', doi:10.17605/OSF.IO/3QTZJ

Rudnytskyi, I. (2023) *openai: R Wrapper for OpenAI API*. R package version 0.4.1. Available at: https://CRAN.R-project.org/package=openai.

Sahn, A., L. Stoker and A. E. Lerman (2025). Equivalency Framing of Problems and Policy Solutions. Political Behavior. doi:10.1007/s11109-025-10023-9

Shengelia, T. and D. Lagnado (2021), 'Are jurors intuitive statisticians? Bayesian causal reasoning in legal contexts', *Frontiers in Psychology*, **11**. doi:10.3389/fpsyg.2020.519262 [Accessed 11 December 2024].

Simon, D. and S. J. Read (2023), 'Toward a general framework of biased reasoning: coherence-based reasoning", *Perspectives on Psychological Science*. 17456916231204579. SAGE Publications Inc. doi:10.1177/17456916231204579

Sommer, J., J. Musolino and P. Hemmer (2023), 'Updating, evidence evaluation, and operator availability: a theoretical framework for understanding belief', *Psychological Review*, **131**(2): 373–401. doi:10.1037/rev0000444.

Song, H. and J. So (2023), 'Message fatigue beyond the health message context: a replication and further extension of So et al. (2017)', *Human Communication Research*, **49**(3): 339–344. doi:10.1093/hcr/hqad021

Sunstein, C. R., D. Schkade, L. A. Reisch and D. Kahneman (2016), 'How people update beliefs about climate change: good news and bad news', *Cornell Law Review*, **102**: 1431–1743.

Taber, C. S. and M. Lodge (2006), 'Motivated skepticism in the evaluation of political beliefs', *American Journal of Political Science*, **50**(3): 755–769. doi:10.1111/j.1540-5907.2006.00214.x

Team, R Core (2023) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Vlasceanu, M., C. E. McMahon, J. J. Van Bavel and A. Coman (2023), 'Political and nonpolitical belief change elicits behavioral change', *Journal of Experimental Psychology: Applied*, **29**(3): 467–476. doi:10.1037/xap0000455

Wolf, L. J., M. P. White, W. Poortinga, S. Fisher and L. Whitmarsh (2024), 'What about the children? The effectiveness of including children in environmental appeals', *Journal of Environmental Psychology*, **94**: 102195. doi:10.1016/j.jenvp.2023.102195

WHO Regional Office for Europe (2024), *Measuring and Maximizing Public Support for Health Policies: Behavioural and Cultural Insights Policy Brief Series*, Copenhagen: WHO Regional Office for Europe. 2024. Licence: CC BY-NC-SA 3.0 IGO

Xu, J. (2023), 'A meta-analysis comparing the effectiveness of narrative vs. statistical evidence: health vs. non-health contexts', *Health Communication*, **38**(14): 3113–3123. Routledge. doi:10.1080/10410236.2022.2137750

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press

Xue, H. and L. Taylor (2023), 'When do people believe, check, and share health rumors on social media? Effects of evidence type, health literacy, and health knowledge', *Journal of Health Psychology*, **28**(7): 607–619. SAGE Publications Ltd. doi:10.1177/13591053221125992

Zhao, X. and S. J. Tsang (2024), 'How people process different types of health misinformation: roles of content falsity and evidence type', *Health Communication*, **39**(4): 741–753. doi:10.1080/10410236.2023.2184452

Zhou, C. and Q. Yang (2024), 'When arguments meet a story: an experiment testing message design strategies for skin cancer prevention and detection', *Health Communication*, (in press), pp. 1–12. Routledge, doi:10.1080/10410236.2024.2315320

Cite this article: Rodger, A., G.A. Sanna, V. Cheung, N. Raihani and D. Lagnado (2026), 'Negative anecdotes reduce policy support: evidence from three experimental studies on communicating policy (in)effectiveness', *Behavioural Public Policy*, 1–32. https://doi.org/10.1017/bpp.2025.10028

https://doi.org/10.1017/bpp.2025.10028 Published online by Cambridge University Press