

RESEARCH ARTICLE OPEN ACCESS

Does Carbon Label Design Affect Flight Choices and Trustworthiness?

Henrik Wittwer¹  | Nick von Felten² | Andreas Wittmer¹

¹Center for Aviation and Space Competence, University of St. Gallen, St. Gallen, Switzerland | ²Institute of Computer Science, University of St. Gallen, St. Gallen, Switzerland

Correspondence: Henrik Wittwer (henrik.wittwer@unisg.ch)

Received: 18 February 2026 | **Revised:** 27 April 2026 | **Accepted:** 17 May 2026

Keywords: carbon labelling | flight emissions | traffic light colours | trustworthiness

ABSTRACT

Carbon emissions labels shown during flight booking may shift flight choices but also affect perceived trustworthiness of booking-page information. This study tests how carbon label design influences flight choices, booking propensity and trustworthiness. In a preregistered online experiment ($N=939$), participants completed a simulated flight booking task and were randomly assigned to one of five carbon label conditions: no label, numeric emissions only, traffic-light colour coding, colour coding plus a contextual benchmark or colour coding plus the European Union Flight Emissions Label. Flight choice was analysed with multinomial logistic regression, booking propensity with binary logistic regression and trustworthiness with multigroup confirmatory factor analysis. Traffic-light and enhanced carbon label formats increased selection of the lowest-emission option relative to no label but did not robustly change booking propensity. These formats also reduced perceived trustworthiness of booking-page information, even as they increased selection of the lowest-emission option.

1 | Introduction

As economic development and income levels rise, tourism has become an increasingly important source of greenhouse gas (GHG) emissions (UNWTO 2019). By 2019, tourism-related emissions were estimated at 8.8% of total global GHG emissions, and between 2009 and 2019 the tourism carbon footprint expanded at 3.5% per year—more than twice the growth rate of the global economy. This trend highlights tourism as a major challenge for climate mitigation and has been attributed to sustained growth in demand combined with only gradual technological efficiency gains, particularly in air travel (Sun et al. 2024). In this context, policy interest has increasingly turned to measures that can influence demand-side decisions in the near term, including information-based approaches such as carbon labelling (Vandenbergh et al. 2011).

Carbon labelling is a point-of-purchase information instrument that translates GHG emissions data into an indicator intended

to inform consumer decision-making (Taufique et al. 2022). By reducing information asymmetries between suppliers and consumers and making environmental externalities more salient, labels are intended to facilitate choices that lower emissions (Heinzle and Wüstenhagen 2012). In aviation, emissions information displayed during flight booking can enable passengers to compare the estimated per passenger GHG emissions of flight options at the moment of choice (Gössling and Dolnicar 2023). However, whether such disclosure influences behaviour depends on whether passengers regard the information as trustworthy (Baumeister et al. 2022).

This question of trustworthiness is particularly salient in aviation. Airlines have been criticised for communication interpreted as greenwashing (Baumeister and Onkila 2017; European Commission 2025). Related critiques argue that parts of the sector's climate-related claims have been untruthful or misleading (Guix et al. 2022; Peeters et al. 2016). In such a context, passengers may discount emissions information, limiting the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Business Strategy and the Environment* published by ERP Environment and John Wiley & Sons Ltd.

behavioural impact of carbon labels (Gössling and Buckley 2016; Thøgersen 2000). Aviation carbon labelling can be understood as an environmental information instrument whose effectiveness depends in part on whether passengers perceive the disclosed emissions information as trustworthy (Baumeister et al. 2022).

Given this, label effectiveness depends not only on whether emissions information is disclosed but also on how that information is designed and communicated at the point of choice (Feucht and Zander 2018). Design features are recommended to improve comprehension and comparability across options—for example, combining numerical CO₂ information with visual formats such as traffic-light colour schemes (Eijgelaar et al. 2016). Beyond supporting understanding, design may also shape perceived trustworthiness by influencing how credible the information appears—for instance, through cues such as independent third-party verification (Baumeister and Onkila 2017). Different carbon label formats may therefore affect not only how easily passengers interpret emissions information but also whether they consider it trustworthy (Baumeister et al. 2022).

Empirical research on aviation carbon labelling is growing, but much of the evidence comes from hypothetical choice settings (Atabekov et al. 2024; Carroll et al. 2022; Baumeister et al. 2022). While former research suggests carbon labels can shift stated preferences toward lower-emission flights, it remains unclear how different label formats shape perceived trustworthiness of the disclosed emissions information and whether this relates to booking behaviour. Accordingly, we pose the following research question: *How do different carbon label formats influence passengers' booking behaviour and their perceived trustworthiness of the flight-related information?*

To examine this question, a preregistered online experiment ($N=939$) randomly assigned participants to a simulated Google Flights-style booking task in which all attributes were held constant except the carbon-label display condition shown for each flight: no carbon-label information (Control); a numeric-only CO₂eq value (Neutral); the same numeric CO₂eq value with the number itself displayed in traffic-light colours (green/yellow/red) (Coloured); a traffic-light-coloured CO₂eq value plus a contextual benchmark showing the flight's CO₂eq as a percentage of the average EU resident's monthly CO₂eq footprint (benchmark source: European Union Aviation Safety Agency, n.d.) (Context); and a traffic-light-coloured CO₂eq value plus the EU Flight Emissions Label (FEL). Label format effects on flight choices were estimated using multinomial logistic regression, while perceived trustworthiness of the flight-related information was measured using a four-item semantic differential scale (Kirmani et al. 2017) and analysed as a latent construct via multigroup CFA to enable comparisons across conditions.

The study contributes to theory on carbon labelling in three ways. First, it provides evidence on which disclosure formats are most effective in shifting choices toward lower-emission flights. It shows that visually salient formats can make emissions behaviourally consequential and that combining numerical values with traffic-light colour cues may already be sufficient (Gössling and Buckley 2016). Second, it shows that some label formats can reduce perceived trustworthiness while still influencing choices. This pattern challenges the assumption

that trustworthiness and behavioural effectiveness necessarily move together, although the present design does not identify the mechanism underlying this divergence (Thøgersen 2000; Darnall et al. 2018; Taufique et al. 2017). Third, it helps bound claims about carbon label impacts beyond option selection by testing effects on booking propensity. The results suggest that carbon label disclosure can reweight preferences within a choice set without robustly increasing booking propensity (Feucht and Zander 2018; Atabekov et al. 2024; Leonelli et al. 2026).

2 | Theoretical Development

2.1 | Carbon Labelling in the Airline Industry

Carbon labelling is a point-of-purchase information instrument that translates GHG emissions data into an indicator intended to support consumer decision-making (Taufique et al. 2022). In the airline industry, its purpose is to enable consumers to compare flight options in terms of their environmental impacts at the point of booking. Without such disclosure, environmental consequences are difficult to assess in booking interfaces, making it unlikely that climate considerations become decision-relevant in flight choices (Baumeister and Onkila 2017). Emissions disclosure is therefore a necessary condition for environmentally informed flight decisions.

Whether disclosure affects choices, however, depends on whether consumers attend to emissions information and use it in their decision process at the point of booking (Gössling and Dolnicar 2023). For emissions information to become decision-relevant, the following conditions must therefore be met: it must be noticed, it must be understandable and it must be regarded as trustworthy (Thøgersen 2000; Gössling and Buckley 2016). The next section conceptualises carbon labels as decision-information nudges and derives these conditions for use, before turning to why trustworthiness is a particularly salient constraint in aviation (Baumeister and Onkila 2017).

2.2 | Carbon Labels as Decision Information Nudges

Carbon labels displayed during flight search can be conceptualised as decision-information nudges because they change how decision-relevant information is presented without altering the underlying choice set (Thaler and Sunstein 2008; Münscher et al. 2016). In the choice architecture literature, decision-information nudges facilitate the use of an attribute by making it more salient and easier to interpret at the moment of choice (Mertens et al. 2021). This framing is particularly relevant in air travel because flight-related GHG emissions are often absent from booking interfaces, making disclosure the primary way to introduce emissions as a comparable attribute for consumers (Gössling and Dolnicar 2023).

Decision-information nudges affect behaviour only if the presented information is incorporated into the decision process (Münscher et al. 2016). In the context of emissions disclosure, two conditions follow from this logic: consumers must (i) notice the information and (ii) be able to understand what it implies

for the options they are comparing (Mertens et al. 2021). When either condition is not met, disclosure is unlikely to affect choices, which helps explain why carbon labels sometimes show weak effects when they are cognitively demanding to interpret (Thøgersen and Nielsen 2016). Accordingly, label format is important because it can strengthen these two conditions—salience and understandability—thereby increasing the likelihood that emissions information becomes decision-relevant.

Salience refers to whether emissions information attracts attention at the point of choice (Lohmann et al. 2024). Because attention is limited and typically allocated to only a subset of available attributes, information that stands out is more likely to be attended to and incorporated into judgment (Johnson et al. 2012; Beyer et al. 2024). Applied to carbon labels, this implies that consumers will not rely on emissions information if they do not notice it in the booking interface (Thøgersen 2000). Label formats that use visual cues—such as traffic-light colour coding—should therefore increase the probability that emissions information is salient during flight booking (Lehner et al. 2016; Thøgersen et al. 2024).

According to Gössling and Buckley (2016), understandability refers to whether consumers can interpret the disclosed emissions in a way that supports comparison across options. Even when a label is noticed, consumers may not use it if the information is difficult to translate into an evaluative judgment. Decision information nudges therefore aim not only to capture attention but also to reduce the cognitive costs of integrating an attribute into decision-making (Beyer et al. 2024). In the carbon labelling literature, limited understanding is identified as a key reason for inconsistent behavioural effects (Thøgersen and Nielsen 2016). Accordingly, design cues that communicate relative performance and provide reference points should make emissions information more intuitively interpretable, increasing the likelihood that it is used in booking decisions (Thøgersen and Nielsen 2016; Baumeister and Onkila 2017).

2.3 | Carbon Label Formats

Carbon label formats are design choices that add interpretive support to emissions disclosure, thereby enabling emissions information to become usable in decision environments (Taufique et al. 2022). Building on the two preconditions for use outlined above (salience and understandability), this section distinguishes three format types that vary in the degree of interpretive structure they provide: (i) quantitative labels, (ii) ordinal labels and (iii) contextual labels (Beyer et al. 2024). Because these formats may differ in how strongly they support salience and understandability, they are expected to differ in how likely emissions information is incorporated into decision-making (Gössling and Buckley 2016). These categories correspond to the label conditions examined in this study.

A baseline approach is a numerical-only emissions display (e.g., CO₂eq per passenger) (Gössling and Dolnicar 2023). While numerical information allows direct comparison across options, it presumes that consumers can interpret their magnitude—specifically, whether an amount such as 530 kg CO₂eq is ‘high’ or

‘low’. Evidence of generally low ‘carbon competence’ suggests that many individuals struggle to evaluate emissions figures in isolation, which may limit the behavioural impact of numeric-only disclosure and motivate formats that add evaluative cues and reduce interpretive effort at the point of choice (Johnson et al. 2024).

One widely recommended enhancement is traffic-light colour coding. Colour cues can increase visual salience and provide an intuitive signal of relative performance, improving label understanding compared with a colour-neutral display (Feucht and Zander 2018; Gössling and Buckley 2016). In flight booking environments, such ordinal cues should therefore make emissions differences easier to integrate into the booking trade-off than numeric-only disclosure (Thøgersen et al. 2024). However, while colour coding communicates relative performance, it does not provide an explicit reference frame for judging magnitude, leaving ambiguity about what the emissions level means in absolute terms (Beyer et al. 2024).

A further approach is to provide contextual benchmarks that anchor the displayed emissions in a reference frame. Environmental information may become easier to evaluate when mapped onto a benchmark resembling a personal carbon account (Hahnel et al. 2020). Consistent with this logic, Beyer et al. (2024) show that expressing food-related emissions as a share of a daily carbon budget can shift choices away from emissions-intensive options. Applied to aviation, framing flight emissions relative to a benchmark (e.g., as a share of an average monthly CO₂eq footprint) could support interpretation and increase the behavioural relevance of the information (Johnson 2021). By relating emissions to an explicit reference point, contextualisation may also increase personal relevance, further increasing the likelihood that emissions information is integrated into decisions (Mertens et al. 2021). However, even salient and interpretable emissions information may have limited influence in aviation if consumers doubt its credibility, making perceived trustworthiness a further constraint (Baumeister and Onkila 2017).

2.4 | Trustworthiness Constraints in Aviation

Even well-designed labels may have limited influence in contexts where environmental communication is met with scepticism (Taufique et al. 2017). Across consumer domains, perceived trustworthiness is a key determinant of whether individuals rely on environmental information when making choices (Darnall et al. 2018; Atkinson and Rosenthal 2014). When consumers doubt environmental claims, they are less likely to incorporate the information into decisions and may disengage from it altogether (Taufique et al. 2017). This constraint is particularly salient in aviation, where airlines’ environmental communications are frequently interpreted through a greenwashing lens (Baumeister and Onkila 2017; Gössling and Dolnicar 2023). Accordingly, perceived trustworthiness is treated in prior literature as a core condition for whether carbon labels can influence consumer behaviour (Thøgersen 2000).

One pathway shaping trustworthiness concerns the source of the information (Gorton et al. 2021). Environmental labels associated with governments or credible third parties tend to be

trusted more than labels perceived as corporate self-promotion (Darnall et al. 2018; Atkinson and Rosenthal 2014). In aviation, official or industry-wide labelling schemes may increase perceived trustworthiness by signalling common rules, verification and comparability across options, which may in turn make passengers more willing to rely on the information when evaluating flight options (Baumeister and Onkila 2017).

Beyond source cues, emissions disclosure itself may also affect trust perceptions by signalling transparency (Atkinson and Rosenthal 2014). Policy discussions frame voluntary emissions disclosure as a mechanism to enhance transparency and more informed decisions for airline passengers (European Commission 2024). If emissions information is interpreted as evidence of accountability, consumers may therefore perceive the booking information as more trustworthy than when no emissions information is provided. Together, these arguments motivate hypotheses not only about behavioural responses to carbon labels but also about perceived trustworthiness of the flight-related information.

2.5 | Hypothesis Development

Empirical research on aviation carbon labelling indicates that providing emissions information can shift stated preferences toward lower-emission flights, yet questions about how carbon labels operate in booking environments remain unresolved (Atabekov et al. 2024; Carroll et al. 2022; Baumeister et al. 2022). Three theoretical gaps motivate the present study. First, prior work rarely examines whether carbon label format affects the perceived trustworthiness of flight-related emissions information. Second, evidence is limited on behavioural effects of carbon label designs that add contextual benchmarks or third-party verification cues. Third, previous experimental designs tend to rely on forced choice, leaving unclear whether labels affect consumers' booking propensity. Booking propensity is defined as whether consumers indicate that they would book one of the displayed flights when asked, 'If you were currently preparing this trip, would you book one of these flights?'

Building on these gaps, this study derives hypotheses for two behavioural outcomes—(i) flight choice and (ii) booking propensity—and one perceptual outcome, namely the perceived trustworthiness of the flight-related information. Label-format hypotheses follow the theoretical ladder developed above. Adding interpretive support should increase the likelihood that emissions information is incorporated into consumer evaluation, and third-party verification cues should further increase reliance by enhancing perceived trustworthiness.

2.5.1 | Behavioural Hypotheses

Disclosing emissions information is expected to influence flight choice because it makes emissions an explicit, comparable attribute in the booking interface and can activate environmental considerations in the booking decision (Baumeister et al. 2022; Atabekov et al. 2024). Relative to no disclosure,

making emissions observable should therefore increase the likelihood that consumers select lower-emission options. Accordingly:

Hypothesis 1. *Consumers are more likely to choose a lower-emission flight when they are provided with any form of carbon label (vs. no label).*

Beyond shifting which option is selected, emissions disclosure may affect whether consumers proceed from evaluation to action in the booking interface. Prior evidence suggests that carbon labels can increase purchase probability (Feucht and Zander 2018; Atabekov et al. 2024), and related work shows higher purchase intent when consumers receive positive environmental signals (Leonelli et al. 2026). In the present study, we capture this transition as booking propensity, measured with a binary pre-choice question asked before participants selected among the flight options. This motivates the expectation that carbon emissions disclosure increases booking propensity relative to no disclosure. Accordingly:

Hypothesis 2. *Consumers exhibit higher booking propensity when they are provided with any form of carbon label (vs. no label).*

While any disclosure makes emissions observable, formats differ in the degree of salience and interpretive support they provide (Feucht and Zander 2018). Traffic-light colour coding can increase salience and convey relative performance through familiar evaluative cues, making emissions information easier to process at the point of choice (Gössling and Buckley 2016; Thøgersen et al. 2024). If colour coding reduces the cognitive effort required to incorporate emissions into the decision, it should strengthen both the shift toward lower-emission options and booking propensity. Therefore:

Hypothesis 1a. *Consumers are more likely to choose a lower-emission flight when they are provided with a colour-coded carbon label (vs. a colour-neutral label).*

Hypothesis 2a. *Consumers exhibit higher booking propensity when they are provided with a colour-coded carbon label (vs. a colour-neutral label).*

Contextual benchmarks can reduce ambiguity by anchoring numerical emissions information to a meaningful reference frame that supports magnitude judgments (Beyer et al. 2024). In this study, the contextualised label supplements the traffic-light-coloured CO₂eq value with a benchmark that expresses each flight option's emissions as a percentage of the average EU resident's monthly CO₂eq footprint (European Union Aviation Safety Agency, n.d.). By making emissions easier to interpret and more personally relevant, contextualisation should increase the likelihood that consumers use emissions information when deciding both which flight to select and whether to proceed to booking (Mertens et al. 2021). Accordingly:

Hypothesis 1b. *Consumers are more likely to choose a lower-emission flight when they are provided with contextualised carbon footprint information (vs. a colour-coded label).*

Hypothesis 2b. *Consumers exhibit higher booking propensity when they are provided with contextualised carbon footprint information (vs. a colour-coded label).*

Even salient labels may have limited impact in aviation if consumers question the trustworthiness of environmental information (Darnall et al. 2018; Atkinson and Rosenthal 2014; Thøgersen 2000). Third-party, government-backed formats may therefore matter by signalling verification. The EU Flight Emissions Label is designed to provide such third-party verification cues (European Commission 2024; Baumeister and Onkila 2017). If these cues increase willingness to rely on the information, the EU Flight Emissions Label format should outperform an otherwise similar contextualised display. Therefore:

Hypothesis 1c. *Consumers are more likely to choose a lower-emission flight when they are provided with carbon labels featuring the EU Flight Emissions Label (vs. a contextualised label).*

Hypothesis 2c. *Consumers exhibit higher booking propensity when they are provided with carbon labels featuring the EU Flight Emissions Label (vs. a contextualised label).*

2.5.2 | Perceptual Hypothesis

Trustworthiness has been described as a general condition for carbon labels to influence consumer responses (Thøgersen 2000). In aviation, where emissions figures are not possible for consumers to verify directly when they are not disclosed on the booking page, the provision of quantified footprint information may function as a cue of transparency and accountability relative to non-disclosure (Atkinson and Rosenthal 2014). More broadly, carbon labelling systems can serve as public signals of corporate movement toward environmental goals (Taufique et al. 2022). If disclosure is interpreted as an effort to make climate impacts visible at the point of choice, consumers should evaluate the flight-related information as more trustworthy than when no emissions information is provided. Accordingly:

Hypothesis 3. *Consumers who are provided with any form of carbon label perceive the flight-related information as more trustworthy than consumers who receive no carbon label.*

3 | Methods

3.1 | Study Design and Procedure

A preregistered online experiment was conducted to test the hypotheses. The study was implemented in Qualtrics, and participants were recruited via Prolific. The experiment was preregistered on the Open Science Framework, and deviations from the preregistration are documented in Supporting Information S1: Appendix A. Prior to the main study, a pilot with approximately 100 participants was conducted to check the survey instrument, confirm that the booking interface functioned as intended and assess preliminary response behaviour. The pilot employed the same experimental design, treatment conditions and outcome measures as the main study. Pilot responses were retained and

analysed as part of the final dataset. Because the pilot was completed before preregistration, we had partial access to the data prior to finalising the preregistered analysis plan. However, the preregistration was submitted before collecting the remainder of the sample and before conducting statistical analyses on the combined dataset.

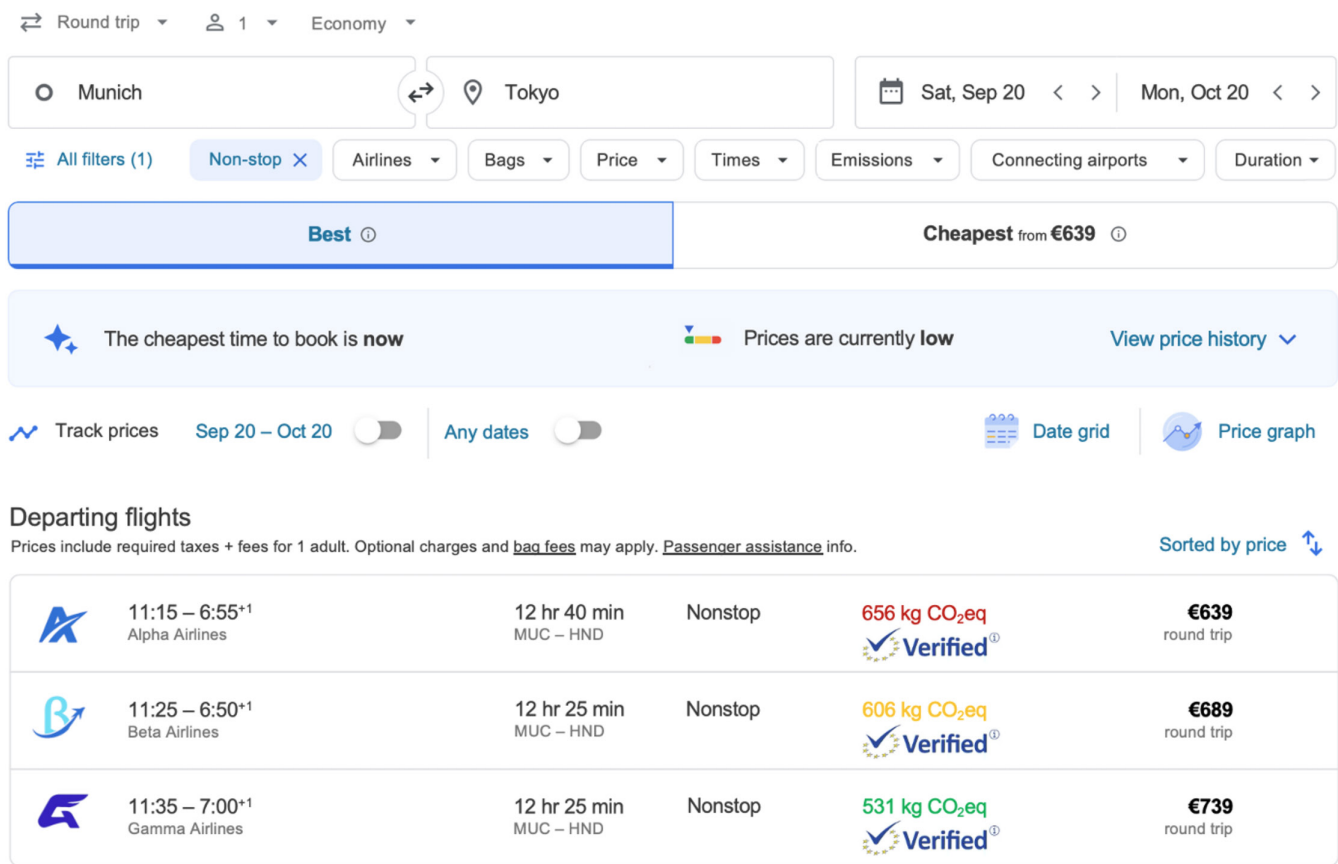
After providing informed consent, participants were introduced to a hypothetical flight booking task. They were asked to imagine planning a holiday trip to Japan with a friend and booking a nonstop economy-class flight from Munich to Tokyo. Participants then viewed a simulated booking interface modelled after Google Flights. The interface displayed three nonstop flight options (Alpha, Beta and Gamma Airlines) with nearly identical departure and arrival times and realistically priced economy-class fares. All interface elements were held constant across conditions, except for the presentation of emissions information, which varied by treatment.

Participants were randomly assigned to one of five carbon-label display conditions. In the Control condition, no emissions information was shown. In the Neutral condition, emissions were shown as an absolute CO₂eq value in kilograms in black text. In the Coloured condition, the same CO₂eq values were shown, but the numeral itself was colour-coded using a traffic-light scheme (green/yellow/red). In the Context condition, the display matched the Coloured condition and additionally reported each option's emissions as a percentage of the average EU resident's monthly CO₂eq footprint, using the benchmark provided by the European Union Aviation Safety Agency (n.d.). In the FEL condition, the coloured CO₂eq values were presented together with the EU Flight Emissions Label. Figure 1 shows the booking screen in the FEL condition, and Figure 2 summarises the carbon-label formats across conditions.

While interacting with the interface, participants completed two behavioural tasks. First, they reported their booking propensity in response to a binary question about whether they would book one of the displayed flights. Participants who indicated that they would book then selected the flight they would choose. Participants who indicated that they would not book briefly stated their reasons and then completed a forced-choice selection among the three options. After the booking decision task, participants ranked five booking attributes—price, flight time, airline, climate impact and other—on a scale from 1 (most important) to 5 (least important). The survey concluded with background questions on flight behaviour and demographics. Respondents reported the average number of flights taken annually for business and leisure, their typical travel class and demographic characteristics (gender, age, approximate annual gross household income and highest level of education).

3.2 | Measures

Two primary behavioural outcomes were assessed. Booking propensity captured whether participants indicated that they would book one of the presented options. Flight choice captured which of the three options participants selected. For participants who indicated that they would proceed with



The EU Flight Emissions Label helps passengers make more informed choices by providing standardized information on the estimated greenhouse gas emissions of their flight, expressed in kilograms of CO₂ equivalent (kg CO₂eq). Developed by the European Union Aviation Safety Agency, it uses a harmonized methodology that accounts for route, aircraft used, number of passengers, onboard freight and fuel consumption, among others.

FIGURE 1 | Simulated booking screen (European Union Flight Emissions Label condition). Screenshot of the Google Flights-style interface used in the experiment. In this condition, per-passenger flight emissions were displayed as a colour-coded CO₂eq value (traffic-light scheme) together with the EU Flight Emissions Label.

656 kg CO ₂ eq	656 kg CO ₂ eq	656 kg CO ₂ eq Equals 105% of the average monthly CO ₂ eq footprint of a person in the EU	656 kg CO ₂ eq
606 kg CO ₂ eq	606 kg CO ₂ eq	606 kg CO ₂ eq Equals 97% of the average monthly CO ₂ eq footprint of a person in the EU	606 kg CO ₂ eq
531 kg CO ₂ eq	531 kg CO ₂ eq	531 kg CO ₂ eq Equals 85% of the average monthly CO ₂ eq footprint of a person in the EU	531 kg CO ₂ eq

FIGURE 2 | Carbon label formats used across experimental conditions. Comparison of emissions displays shown in the booking interface. From left to right: Neutral (numeric CO₂eq value), Coloured (traffic-light colour-coded CO₂eq value), Context (colour-coded CO₂eq value with contextual benchmark), and FEL condition.

booking, flight choice was observed directly through their selection. For participants who indicated that they would not book at that moment, flight choice was elicited via a forced-choice selection among the three options after they provided a brief reason. Consistent with the hypothesis that emissions disclosure shifts choices toward lower-emission alternatives, the primary operationalisation of lower-emission choice in the airline-choice analyses was selection of Gamma Airlines, the lowest-emission option. In addition, the attribute-ranking

task was included to describe decision priorities across conditions and to support interpretation of observed behavioural patterns.

Perceived trustworthiness of the flight-related information was measured using a four-item, seven-point semantic differential scale adapted from Kirmani et al. (2017). While viewing the booking page, participants were instructed to indicate how trustworthy they considered the information presented on the

flight booking page. They rated the booking-page information on four bipolar adjective pairs—dishonest–honest, not trustworthy–trustworthy, insincere–sincere and manipulative–not manipulative—using a seven-point response format.

3.3 | Data Collection

Data were collected via Prolific and the survey was hosted in Qualtrics. Participants were eligible if they were at least 18 years old, resided in Germany, Austria or Switzerland, and reported fluency in English. Inclusion in the final dataset required meeting eligibility criteria, providing informed consent, completing the full survey and passing two attention checks (one instruction-based item and one forced-response item). The target sample size was determined a priori based on feasibility and the requirements of the preregistered models. Because the experiment comprised five conditions, we aimed to recruit 200 participants per condition (target total $N=1000$), with at least 160 participants per condition (minimum total $N=800$) deemed acceptable if recruitment constraints limited further sampling. To justify these targets, a Monte Carlo power and precision analysis was conducted in R for the most complex preregistered structural equation model, specified with 54 free parameters. For each candidate per-condition sample size, 1000 replications were simulated and evaluated using average absolute parameter bias and 95% confidence interval coverage following Muthén and Muthén (2002). Performance improved as group sizes approached 200, with parameter bias remaining within commonly used benchmarks ($\leq 5\%$) and confidence interval coverage approaching nominal levels. The final sample is summarised in Table 1.

3.4 | Evaluation and Analysis Methods

Analyses were conducted in R and statistical significance was evaluated at $\alpha=0.05$ (two-sided). For regression models, effects are reported as odds ratios with 95% confidence intervals. Post hoc pairwise comparisons were conducted using estimated marginal means with Holm correction to control the family-wise error rate.

Booking propensity was analysed using binary logistic regression with experimental condition as the predictor. Condition differences were examined using model-based predicted booking probabilities and pairwise contrasts, and booking proportions are reported descriptively by condition. Flight choice was analysed using multinomial logistic regression with experimental condition as the predictor and Alpha Airlines as the reference category. Condition differences were summarised using model-based predicted choice probabilities and pairwise contrasts. In line with the operationalisation of lower-emission choice, hypothesis tests for the choice outcome focus on selection of Gamma Airlines, the lowest-emission option. As a supplementary sensitivity analysis, both models were re-estimated with demographic covariates (collapsed gender, age, collapsed education, collapsed income and annual flight frequency) to assess the sensitivity of the main condition effects to observable variation across experimental groups. Minimum

cell counts were inspected, and the Hausman–McFadden test provided no evidence of violations of the independence of irrelevant alternatives (IIA) assumption. In addition to the preregistered hypotheses, we conducted analyses of participants' attribute rankings and the alignment between participants' top-ranked attribute and their selected flight option (preference matching) to support interpretation. Details are reported in the [Supporting Information](#).

The latent trustworthiness construct was analysed using multi-group confirmatory factor analysis with experimental condition as the grouping variable. Measurement invariance was tested sequentially (configural, metric and scalar) and model fit was evaluated using χ^2 , RMSEA, CFI and SRMR. Following Hu and Bentler (1999), model fit was judged against the preregistered criteria of RMSEA < 0.06 , SRMR ≤ 0.08 and CFI ≥ 0.95 , alongside a non-significant χ^2 test ($p > 0.05$) as an additional indicator of close fit.

4 | Results

4.1 | Flight Choice

Table 2 shows participants' flight choices across experimental conditions. In the control group, 93% chose Alpha Airlines, the cheapest and most carbon-intensive option. In all treatment groups, this share decreased with corresponding increases for Beta and Gamma Airlines. Notably, the coloured condition resulted in the lowest selection rate for Alpha Airlines and the highest selection rates for both Beta and Gamma Airlines.

To test Hypotheses 1–1c, we examined pairwise comparisons of estimated marginal means for Gamma Airlines, which represented the lowest-emission flight option (see Table 3). Hypothesis 1 predicted that any form of carbon footprint information would increase the likelihood of choosing Gamma Airlines compared to no information (Control). This was partially supported: compared to the Control, the Coloured ($p=0.019$), Context ($p=0.048$) and FEL ($p=0.023$) conditions all significantly increased selection of Gamma Airlines. However, the Neutral condition did not reach statistical significance ($p=0.081$).

Hypothesis 1a proposed that colour-coded information would lead to more selections of Gamma Airlines compared to the colour-neutral version. This hypothesis was not supported, as the difference between the Coloured and Neutral conditions was not significant ($p=0.819$). Hypothesis 1b suggested that contextualised carbon labels would increase Gamma choices relative to the colour-coded label, which was also not supported ($p=0.955$). Finally, Hypothesis 1c predicted that the EU Flight Emissions Label (FEL) would outperform contextualised information; however, the contrast was again not significant ($p=0.979$), and thus Hypothesis 1c was not supported. Supporting Information S1: Appendix B displays the full pairwise comparisons of estimated marginal means across experimental conditions for flight choice, including Alpha and Beta Airlines. A supplementary covariate-adjusted

TABLE 1 | Structure of the sample.

Category	Subcategory	Total sample (N = 939)	Control (n = 187)	Neutral (n = 184)	Coloured (n = 192)	Context (n = 187)	FEL (n = 189)
Age (average)		33.5	32.7	33.6	34.2	32.9	33.9
Gender	Woman	42%	46%	40%	34%	44%	46%
	Man	57%	52%	59%	65%	56%	53%
	Other	1%	2%	1%	1%	0%	1%
Education	Lower secondary	2%	3%	3%	3%	1%	3%
	Upper secondary	31%	29%	29%	36%	35%	27%
	Bachelor's degree	30%	30%	29%	27%	35%	32%
	Master's degree or equivalent	31%	30%	32%	31%	27%	33%
	Doctorate/PhD	4%	7%	4%	2%	3%	4%
	Prefer not to say	1%	1%	3%	0%	1%	1%
Annual household income (in €)	Below 20'000	18%	20%	15%	19%	20%	17%
	20'000–39'999	20%	19%	23%	17%	17%	24%
	40'000–59'999	20%	18%	20%	15%	24%	22%
	60'000–79'999	12%	13%	12%	13%	12%	13%
	80'000–99'999	9%	10%	5%	13%	9%	7%
	100'000–119'000	5%	5%	3%	5%	9%	5%
	120'000–139'000	4%	5%	4%	3%	3%	5%
	Above 140'000	4%	5%	5%	6%	3%	2%
	Prefer not to say	7%	5%	13%	9%	5%	3%
Flights per year (average)		2.9	2.7	2.8	2.9	2.9	3.3
Travel class private	Economy Class	84%	86%	83%	85%	78%	86%
	Business Class	2%	0%	2%	3%	3%	3%
	Premium Economy Class	6%	9%	5%	3%	8%	4%
	I do not fly for private reasons	8%	5%	9%	9%	11%	7%
Travel class business	Economy Class	38%	40%	34%	41%	34%	40%
	Business Class	7%	9%	8%	5%	9%	7%
	Premium Economy Class	5%	5%	5%	5%	4%	5%
	I do not fly for business purposes	50%	47%	53%	49%	52%	48%

sensitivity analysis is reported in Supporting Information S1: Appendix E. The adjusted multinomial model preserved the general directional pattern across conditions, particularly the lower selection of Alpha Airlines in the visually enhanced label conditions. However, the pairwise contrasts for selection of Gamma Airlines were attenuated and were no longer statistically significant after adjustment.

4.2 | Booking Propensity

Table 4 presents booking propensity across the experimental conditions. In the Control condition, 89% of participants indicated that they would book one of the displayed flights. This share was higher in all treatment conditions, with 93% in the Neutral condition, 95% in both the Coloured and Context conditions and

92% in the FEL condition. To test Hypothesis 2, we estimated a binary logistic regression predicting booking propensity from experimental condition (see Table 5).

The Coloured condition significantly increased booking propensity compared to the Control ($p=0.032$), with an odds ratio (OR) of 2.435 and a 95% confidence interval (CI) of [1.109, 5.762]. The Context, Neutral and FEL conditions were not significantly different from the Control. Predicted probabilities of booking propensity by experimental condition are reported in Table 6. However, the pairwise contrasts between conditions, reported in Table 7, indicated that none of the differences were statistically significant after Holm adjustment. Thus, the Coloured-versus-Control effect did not remain robust under pairwise comparisons. Therefore, Hypotheses 2–2c must be rejected.

The covariate-adjusted sensitivity analysis reported in Supporting Information S1: Appendix E did not change the substantive conclusion for booking propensity. Although the Coloured and Context coefficients remained positive in the

adjusted model, no pairwise differences between conditions were statistically significant after adjustment.

4.3 | Trustworthiness of the Flight-Related Information

For the trustworthiness construct, we first tested measurement invariance across the five experimental groups. The configural model demonstrated good fit ($\chi^2(10)=17.238$, $p=0.069$, RMSEA = 0.073, CFI = 0.995, SRMR = 0.017), indicating that the factor structure was consistent across conditions. The metric invariance model, which constrained factor loadings to be equal across groups, also showed good fit ($\chi^2(22)=32.231$, $p=0.074$, RMSEA = 0.052, CFI = 0.994, SRMR = 0.037) and did not differ significantly from the configural model ($\Delta\chi^2(12)=13.694$, $p=0.321$), supporting metric invariance. When testing for scalar invariance, model fit significantly worsened ($\Delta\chi^2(12)=27.059$, $p=0.008$), suggesting that full scalar invariance was not supported. Nevertheless, the scalar model still exhibited an acceptable approximate fit (RMSEA = 0.064, CFI = 0.986, SRMR = 0.051), allowing for a comparison of latent means across groups (Table 8).

TABLE 2 | Final flight choices by experimental condition.

Condition	Alpha Airlines n (%)	Beta Airlines n (%)	Gamma Airlines n (%)	Total (n)
Control	174 (93%)	12 (7%)	1 (0%)	187
Neutral	141 (77%)	32 (17%)	11 (6%)	184
Coloured	123 (64%)	52 (27%)	17 (9%)	192
Context	128 (68%)	46 (25%)	13 (7%)	187
FEL	129 (68%)	44 (23%)	16 (9%)	189

Note: Values are counts with row percentages in parentheses. Gamma Airlines represents the lowest-emission flight option.
Abbreviation: FEL = European Union Flight Emissions Label.

TABLE 4 | Booking propensity by experimental condition.

Condition	Would book n (%)	Total (n)
Control	167 (89%)	187
Neutral	171 (93%)	184
Coloured	183 (95%)	192
Context	177 (95%)	187
FEL	174 (92%)	189

Note: Values are counts with row percentages in parentheses.
Abbreviation: FEL = European Union Flight Emissions Label.

TABLE 3 | Pairwise comparisons of estimated marginal means for selection of the lowest-emission flight.

Contrast	Estimate (Δ probability)	SE	df	t	p
Neutral vs. Control	0.054	0.018	10	2.979	0.081
Coloured vs. Control	0.083	0.021	10	3.927	0.019*
Coloured vs. Neutral	0.028	0.027	10	1.067	0.819
Context vs. Control	0.064	0.019	10	3.317	0.048*
Context vs. Neutral	0.009	0.026	10	0.381	0.995
Context vs. Coloured	−0.019	0.028	10	−0.687	0.955
FEL vs. Control	0.079	0.021	10	3.788	0.023*
FEL vs. Neutral	0.024	0.027	10	0.930	0.879
FEL vs. Coloured	−0.003	0.029	10	−0.135	0.999
FEL vs. Context	0.015	0.028	10	0.550	0.979

Note: Estimated marginal mean differences in predicted probability of selecting Gamma Airlines from multinomial logistic regression (reference category = Alpha Airlines). Holm-adjusted p -values reported.
Abbreviations: df = degrees of freedom; SE = standard error; t = t -statistic. FEL = European Union Flight Emissions Label.
* $p < 0.05$.

TABLE 5 | Binary logistic regression predicting booking propensity.

Predictor	Estimate (Logit)	SE	z	p	OR	95% CI
Intercept	2.122	0.237	8.969	<0.001***	8.350	[5.389, 13.693]
Neutral	0.455	0.373	1.220	0.223	1.575	[0.766, 3.343]
Coloured	0.890	0.415	2.143	0.032*	2.435	[1.109, 5.762]
Context	0.751	0.402	1.869	0.062	2.119	[0.985, 4.846]
FEL	0.329	0.358	0.917	0.359	1.389	[0.691, 2.848]

Note: Outcome = booking propensity (1 = would book). Reference category = Control condition.

Abbreviations: CI = confidence interval; FEL = European Union Flight Emissions Label; OR = odds ratio; SE = standard error; z = z-statistic.

* $p < 0.05$.

*** $p < 0.001$.

TABLE 6 | Predicted probabilities of booking propensity by experimental conditions.

Condition	Predicted probability	SE	95% CI
Control	0.893	0.023	[0.840, 0.930]
Neutral	0.929	0.019	[0.882, 0.959]
Coloured	0.953	0.015	[0.912, 0.975]
Context	0.947	0.017	[0.903, 0.971]
FEL	0.921	0.020	[0.873, 0.952]

Note: Model-based predicted probabilities from binary logistic regression.

Abbreviations: CI = confidence interval; FEL = European Union Flight Emissions Label; SE = standard error.

TABLE 7 | Pairwise contrasts of booking propensity between conditions.

Contrast	OR	SE	z	p
Control vs. Neutral	0.635	0.236	-1.22	0.740
Control vs. Coloured	0.411	0.171	-2.143	0.202
Control vs. Context	0.472	0.190	-1.869	0.334
Control vs. FEL	0.720	0.258	-0.917	0.890
Neutral vs. Coloured	0.647	0.289	-0.976	0.866
Neutral vs. Context	0.743	0.323	-0.684	0.960
Neutral vs. FEL	1.134	0.447	0.319	0.998
Coloured vs. Context	1.149	0.542	0.294	0.998
Coloured vs. FEL	1.753	0.762	1.291	0.697
Context vs. FEL	1.526	0.644	1.001	0.855

Note: Pairwise comparisons derived from binary logistic regression. Holm-adjusted p -values reported.

Abbreviations: FEL = European Union Flight Emissions Label; OR = Odds ratio; SE = standard error.

Compared to the Control group ($M=0$), participants in the Coloured ($M=-0.313$, $p=0.027$), Context ($M=-0.367$, $p=0.009$) and FEL ($M=-0.443$, $p=0.002$) conditions reported significantly lower trust in the flight-related information. The Neutral label had no significant effect ($p=1$). Contrary to our hypothesis Hypothesis 3, results showed that providing carbon

footprint information did not increase perceived trust in the flight-related information. In fact, participants exposed to Coloured, Context or FEL treatments expressed significantly lower trust than those in the Control or Neutral conditions. Therefore, the hypothesis Hypothesis 3 predicting higher trust in the flight-related information as a result of disclosing carbon labels must be rejected.

5 | Discussion and Implications

5.1 | Main Findings

This study tested how aviation carbon label formats affect flight choices, booking propensity and perceived trustworthiness of flight-related information. Three main findings emerge. First, compared to no carbon label disclosure, the Coloured, Context and FEL formats significantly increased selection of the lowest-emission flight option, whereas numeric-only disclosure did not. Supplementary covariate-adjusted models preserved the general directional pattern in flight choice, particularly the lower selection of Alpha Airlines in the visually enhanced label conditions, but the Gamma-specific contrasts were attenuated and no longer statistically significant after adjustment. Second, carbon labels did not robustly affect booking propensity. Any apparent increase did not hold under Holm-adjusted pairwise comparisons. This conclusion was unchanged in the covariate-adjusted sensitivity analysis. Third, contrary to Hypothesis 3, carbon labels did not increase perceived trustworthiness. Coloured, Context and FEL significantly reduced trustworthiness relative to Control and Neutral, while Neutral did not differ from Control. Overall, the results show that carbon label format can shift choices toward lower-emission options, but visually salient formats simultaneously reduced perceived trustworthiness of the flight-related information.

5.2 | Theoretical Implications

The findings extend aviation carbon labelling research by clarifying three gaps that remain underexplored in prior work: which carbon label formats are most effective in shifting booking choices toward lower-emission options, whether label format influences perceived trustworthiness of flight-related information and whether labels affect

TABLE 8 | Pairwise Wald tests of latent trustworthiness mean differences.

Contrast	Mean difference	SE	95% CI	Wald χ^2	df	p
Control vs. Neutral	−0.045	0.111	[−0.263, 0.173]	0.167	1	1.000
Control vs. Coloured	0.313	0.107	[0.103, 0.524]	7.764	1	0.027*
Control vs. Context	0.367	0.109	[0.153, 0.580]	10.194	1	0.009**
Control vs. FEL	0.443	0.115	[0.218, 0.669]	13.231	1	0.002**
Neutral vs. Coloured	0.358	0.106	[0.149, 0.567]	10.386	1	0.009**
Neutral vs. Context	0.412	0.108	[0.200, 0.623]	13.280	1	0.002**
Neutral vs. FEL	0.488	0.114	[0.265, 0.712]	16.716	1	<0.001***
Coloured vs. Context	0.054	0.104	[−0.151, 0.258]	0.247	1	1.000
Coloured vs. FEL	0.130	0.110	[−0.086, 0.347]	1.308	1	1.000
Context vs. FEL	0.077	0.112	[−0.143, 0.296]	0.445	1	1.000

Note: Latent mean differences derived from multigroup confirmatory factor analysis. Positive values indicate higher perceived trustworthiness in the first-listed condition. Holm-adjusted *p*-values reported.

Abbreviations: CI = confidence interval; FEL = European Union Flight Emissions Label; SE = standard error; Wald χ^2 = Wald chi-square test statistic.

**p* < 0.05.

***p* < 0.01.

****p* < 0.001.

booking propensity (Atabekov et al. 2024; Carroll et al. 2022; Baumeister et al. 2022).

Drawing on the choice architecture view of carbon labels as decision-information nudges (Münscher et al. 2016; Mertens et al. 2021), this study provides empirical support for the argument that label format features that increase salience can make emissions information behaviourally consequential. Compared to no carbon label disclosure, the more visually structured formats increased selection of the lowest-emission option. Importantly, the enhanced formats did not differ significantly from one another. This pattern is consistent with prior evidence that the most effective designs combine an absolute CO₂(eq) figure with intuitive traffic-light colour cues, pairing factual emissions information with a visual scale that supports comparison across differing levels of carbon capability (Feucht and Zander 2018; Gössling and Buckley 2016).

This study also advances theory by documenting a divergence between behavioural and perceptual responses to carbon label formats. Prior work treats trustworthiness as an important condition shaping whether consumers rely on environmental information in decision-making (Thøgersen 2000; Darnall et al. 2018; Taufique et al. 2017; Gössling and Buckley 2016). In the present study, however, the more salient formats reduced perceived trustworthiness relative to the no-label and numeric-only conditions while increasing selection of the lowest-emission flight. This pattern challenges the assumption that higher perceived trustworthiness is required for behavioural effectiveness in this setting. At the same time, because the study did not include process measures such as attention allocation, comprehension, perceived greenwashing, recall or response times, it cannot determine the mechanism underlying this divergence.

Third, this study increases understanding of whether carbon labels influence whether consumers proceed from evaluation to a

booking decision. Prior work suggests that carbon labels can increase purchase intentions (Feucht and Zander 2018; Atabekov et al. 2024; Leonelli et al. 2026), motivating expectations of higher booking propensity under carbon label disclosure. In this study, however, differences in booking propensity were not robust under pairwise comparisons. This finding helps bound claims about carbon label effectiveness by suggesting that, in a simplified booking context with high baseline booking propensity, emissions disclosure primarily reweights preferences within the choice set rather than increasing the incidence of booking.

5.3 | Managerial Implications

The findings have practical implications for booking platforms, airlines and policymakers seeking to reduce aviation emissions through demand-side measures. The key message is that carbon labels can shift flight choice, but implementation must balance behavioural effectiveness with trustworthiness.

First, platforms and policymakers should prioritise simple, standardised and visually structured disclosure. The findings indicate that emissions information becomes more actionable when it is easy to spot and compare at the point of booking. Pairing an absolute CO₂(eq) value with intuitive colour cues is therefore a feasible and scalable way to introduce emissions as a decision attribute under real-world cognitive constraints (Johnson et al. 2012). Practically, this implies displaying emissions information in the primary choice set, using consistent formatting across flight options to support comparability (Gössling and Dolnicar 2023).

Second, implementation strategies should address the observed trustworthiness deficit associated with salient formats. Prior work suggests that third-party, industry-wide and

airline-independent schemes can reduce concerns about self-serving corporate communication (Baumeister and Onkila 2017; Darnall et al. 2018). However, third-party cues will only function as verification signals if passengers understand what they represent. This requires transparent specifications of how emissions values are calculated, who governs the methodology and what is excluded from the metric (e.g., offsetting or compensation claims). Without such transparency, “official” or “verified” cues risk being interpreted as untrustworthy.

5.4 | Limitations and Future Research

This study provides empirical insights into how aviation carbon label formats affect flight choice, booking propensity and perceived trustworthiness, but it has several limitations that suggest avenues for future research. First, the study relied on a hypothetical booking scenario conducted online. Although the experimental setting increased internal validity by holding flight attributes constant and varying only the label format, the absence of actual choice consequences limits behavioural realism. Future research should examine whether the observed choice shifts replicate in field settings (e.g., A/B tests on booking platforms) where choices entail real costs, time constraints and competing interface features.

Second, the experiment employed a stylised choice environment: a single long-haul route (Munich–Tokyo), three fictional airlines, nearly identical schedules and only non-stop economy flights. Real booking environments typically present broader route and carrier sets, including connecting flights, large price dispersion, airline loyalty considerations and heterogeneous time constraints. These factors may interact with emissions disclosure and potentially change both the magnitude and direction of label effects. Subsequent studies should therefore test carbon label formats across multiple routes (short-haul vs. long-haul), ticket classes and choice sets that include connections and greater trade-offs, to assess when labels remain decision-relevant and when conventional attributes dominate.

Third, the sample was restricted to English-fluent residents of Germany, Austria and Switzerland, and the booking task was administered in English. This likely introduced systematic self-selection by favouring a subgroup that is more accustomed to English-language booking environments than the broader population of air travellers in these countries. Such participants may also differ in education, travel experience and environmental awareness, which may influence responses to carbon labelling. The findings should therefore not be assumed to generalise directly to all travellers in German-speaking countries or to booking interfaces presented in users’ native language. Future research should replicate the experiment in native-language booking contexts and in broader samples to assess the generalisability of the observed effects.

Fourth, the study documents a divergence between behaviour and perceived trustworthiness, but it does not identify the process underlying this pattern. Since the experimental conditions differed only in carbon label design, the observed trustworthiness differences can be attributed to label format. However,

the study did not include process measures such as attention allocation, comprehension, perceived greenwashing, recall or response times. As a result, the present findings do not allow us to determine whether trustworthiness is a precondition for behavioural effectiveness. Rather, they show that, in this study, some visually salient label formats increased selection of the lowest-emission flight while also lowering reported trustworthiness. Future work should study the mechanisms behind this pattern by measuring attention and processing (e.g., eye-tracking, response times and recall) and testing whether label familiarity or explanatory information moderates trustworthiness responses. Qualitative work (e.g., focus groups or interviews with frequent flyers) could further clarify why some formats are perceived as less trustworthy and what design features would increase trustworthiness.

Author Contributions

Henrik Wittwer: conceptualization, methodology, writing – original draft, writing – review and editing, project administration, investigation, visualization, validation, formal analysis. **Nick von Felten:** conceptualization, methodology, software, formal analysis, data curation, writing – review and editing, validation, investigation, visualization. **Andreas Wittmer:** writing – review and editing.

Acknowledgements

The authors used an artificial intelligence-based language model (ChatGPT) to assist with language editing and clarity improvements. Open access publishing facilitated by University of St Gallen, as part of the Wiley - University of St Gallen agreement via the Consortium Of Swiss Academic Libraries.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Atabekov, M., V. Bilotkach, K. Kawata, et al. 2024. “Double-Edged Impacts of Carbon Footprint Information on International Air Travel Demand.” *Journal of Air Transport Management* 117: 102570. <https://doi.org/10.1016/j.jairtraman.2024.102570>.
- Atkinson, L., and S. Rosenthal. 2014. “Signaling the Green Sell: The Influence of Eco-Label Source, Argument Specificity, and Product Involvement on Consumer Trust.” *Journal of Advertising* 43, no. 1: 33–45. <https://doi.org/10.1080/00913367.2013.834803>.
- Baumeister, S., and T. Onkila. 2017. “An Eco-Label for the Airline Industry?” *Journal of Cleaner Production* 142: 1368–1376. <https://doi.org/10.1016/j.jclepro.2016.11.170>.
- Baumeister, S., C. Zeng, and A. Hoffendahl. 2022. “The Effect of an Eco-Label on the Booking Decisions of Air Passengers.” *Transport Policy* 124: 175–182. <https://doi.org/10.1016/j.tranpol.2020.07.009>.
- Beyer, B., R. Chaskel, S. Euler, J. Gassen, A. Grosskopf, and T. Sellhorn. 2024. “How Does Carbon Footprint Information Affect Consumer Choice? A Field Experiment.” *Journal of Accounting Research* 62, no. 1: 101–136. <https://doi.org/10.1111/1475-679X.12505>.
- Carroll, J., W. Brazil, M. Howard, and E. Denny. 2022. “Imperfect Emissions Information During Flight Choices and the Role of CO₂ Labelling.” *Renewable and Sustainable Energy Reviews* 165: 112508. <https://doi.org/10.1016/j.rser.2022.112508>.
- Darnall, N., H. Ji, and D. A. Vázquez-Brust. 2018. “Third-Party Certification, Sponsorship, and Consumers’ Ecolabel Use.” *Journal*

- of *Business Ethics* 150, no. 4: 953–969. <https://doi.org/10.1007/s10551-016-3138-2>.
- Eijgelaar, E., J. Nawijn, C. Barten, L. Okuhn, and L. Dijkstra. 2016. “Consumer Attitudes and Preferences on Holiday Carbon Footprint Information in the Netherlands.” *Journal of Sustainable Tourism* 24, no. 3: 398–411. <https://doi.org/10.1080/09669582.2015.1101129>.
- European Commission. 2024. “EU Introduces Flight Emissions Label for Moreinformed and Sustainable travelling.” https://transport.ec.europa.eu/news-events/news/eu-introduces-flight-emissions-label-more-informed-and-sustainable-travelling-2024-12-18_en.
- European Commission. 2025. “Twenty-One European Airlines Agree to Modify Their Practices Regarding Environmental Claims.” https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2608.
- European Union Aviation Safety Agency. n.d. “Empowering Passengers to Make Informed Choices About Their Flights.” Retrieved March 4, 2025. <https://www.flightemissions.eu/en>.
- Feucht, Y., and K. Zander. 2018. “Consumers’ Preferences for Carbon Labels and the Underlying Reasoning. A Mixed Methods Approach in 6 European Countries.” *Journal of Cleaner Production* 178: 740–748. <https://doi.org/10.1016/j.jclepro.2017.12.236>.
- Gorton, M., B. Tocco, C.-H. Yeh, and M. Hartmann. 2021. “What Determines Consumers’ Use of Eco-Labels? Taking a Close Look at Label Trust.” *Ecological Economics* 189: 107173. <https://doi.org/10.1016/j.ecolecon.2021.107173>.
- Gössling, S., and R. Buckley. 2016. “Carbon Labels in Tourism: Persuasive Communication?” *Journal of Cleaner Production* 111: 358–369. <https://doi.org/10.1016/j.jclepro.2014.08.067>.
- Gössling, S., and S. Dolnicar. 2023. “A Review of Air Travel Behavior and Climate Change.” *WIREs Climate Change* 14, no. 1: e802. <https://doi.org/10.1002/wcc.802>.
- Guix, M., C. Ollé, and X. Font. 2022. “Trustworthy or Misleading Communication of Voluntary Carbon Offsets in the Aviation Industry.” *Tourism Management* 88: 104430. <https://doi.org/10.1016/j.tourman.2021.104430>.
- Hahnel, U. J. J., G. Chatelain, B. Conte, V. Piana, and T. Brosch. 2020. “Mental Accounting Mechanisms in Energy Decision-Making and Behaviour.” *Nature Energy* 5, no. 12: 952–958. <https://doi.org/10.1038/s41560-020-00704-6>.
- Heinze, S. L., and R. Wüstenhagen. 2012. “Dynamic Adjustment of Ecolabeling Schemes and Consumer Choice—The Revision of the EU Energy Label as a Missed Opportunity?” *Business Strategy and the Environment* 21: 60–70. <https://doi.org/10.1002/bse.722>.
- Hu, L.-t., and P. M. Bentler. 1999. “Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives.” *Structural Equation Modeling* 6, no. 1: 1–55. <https://doi.org/10.1080/10705519909540118>.
- Johnson, E. J. 2021. *The Elements of Choice: Why the Way We Decide Matters*. Riverhead Books.
- Johnson, E. J., S. B. Shu, B. G. C. Dellaert, et al. 2012. “Beyond Nudges: Tools of a Choice Architecture.” *Marketing Letters* 23, no. 2: 487–504. <https://doi.org/10.1007/s11002-012-9186-1>.
- Johnson, E. J., E. R. Sugerman, V. G. Morwitz, G. V. Johar, and M. W. Morris. 2024. “Widespread Misestimates of Greenhouse Gas Emissions Suggest Low Carbon Competence.” *Nature Climate Change* 14: 707–714. <https://doi.org/10.1038/s41558-024-02032-z>.
- Kirmani, A., R. W. Hamilton, D. V. Thompson, and S. Lantzy. 2017. “Doing Well Versus Doing Good: The Differential Effect of Underdog Positioning on Moral and Competent Service Providers.” *Journal of Marketing* 81, no. 1: 103–117. <https://doi.org/10.1509/jm.15.0369>.
- Lehner, M., O. Mont, and E. Heiskanen. 2016. “Nudging – A Promising Tool for Sustainable Consumption Behaviour?” *Journal of Cleaner Production* 134: 166–177. <https://doi.org/10.1016/j.jclepro.2015.11.086>.
- Leonelli, S., M. Muhn, T. Rauter, and G. Sran. 2026. “How Do Consumers Use ESG Disclosure? Evidence From a Randomized Field Experiment With Everyday Product Purchases.” *Journal of Accounting and Economics* 81, no. 1: 101811. <https://doi.org/10.1016/j.jacceco.2025.101811>.
- Lohmann, P. M., E. Gsottbauer, J. Farrington, S. Human, and L. A. Reisch. 2024. “Choice Architecture Promotes Sustainable Choices in Online Food-Delivery Apps.” *PNAS Nexus* 3, no. 10. <https://doi.org/10.1093/pnasnexus/pgae422>.
- Mertens, S., M. Herberz, U. J. J. Hahnel, and T. Brosch. 2021. “The Effectiveness of Nudging: A Meta-Analysis of Choice Architecture Interventions Across Behavioral Domains.” *Proceedings of the National Academy of Sciences* 119, no. 1: e2107346118. <https://doi.org/10.1073/pnas.2107346118>.
- Münscher, R., M. Vetter, and T. Scheuerle. 2016. “A Review and Taxonomy of Choice Architecture Techniques.” *Journal of Behavioral Decision Making* 29, no. 5: 511–524. <https://doi.org/10.1002/bdm.1897>.
- Muthén, L. K., and B. O. Muthén. 2002. “How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power.” *Structural Equation Modeling: A Multidisciplinary Journal* 9, no. 4: 599–620. https://doi.org/10.1207/S15328007SEM0904_8.
- Peeters, P., J. E. S. Higham, D. Kutzner, S. A. Cohen, and S. Gössling. 2016. “Are Technology Myths Stalling Aviation Climate Policy?” *Transportation Research Part D: Transport and Environment* 44: 30–42. <https://doi.org/10.1016/j.trd.2016.02.004>.
- Sun, Y.-Y., F. Faturay, M. Lenzen, S. Gössling, and J. Higham. 2024. “Drivers of Global Tourism Carbon Emissions.” *Nature Communications* 15, no. 1: 10384. <https://doi.org/10.1038/s41467-024-54582-7>.
- Taufique, K. M. R., K. S. Nielsen, T. Dietz, R. Shwom, P. C. Stern, and M. P. Vandenbergh. 2022. “Revisiting the Promise of Carbon Labelling.” *Nature Climate Change* 12, no. 2: 132–140. <https://doi.org/10.1038/s41558-021-01271-8>.
- Taufique, K. M. R., A. Vocino, and M. J. Polonsky. 2017. “The Influence of Eco-Label Knowledge and Trust on Pro-Environmental Consumer Behaviour in an Emerging Market.” *Journal of Strategic Marketing* 25, no. 7: 511–529. <https://doi.org/10.1080/0965254X.2016.1240219>.
- Thaler, R. H., and C. R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.
- Thøgersen, J. 2000. “Psychological Determinants of Paying Attention to Eco-Labels in Purchase Decisions: Model Development and Multinational Validation.” *Journal of Consumer Policy* 23, no. 3: 285–313. <https://doi.org/10.1023/A:1007122319675>.
- Thøgersen, J., F. J. Dessart, G. Marandola, and S. L. Hille. 2024. “Positive, Negative or Graded Sustainability Labelling? Which Is Most Effective at Promoting a Shift Towards More Sustainable Product Choices?” *Business Strategy and the Environment* 33, no. 7: 6795–6813. <https://doi.org/10.1002/bse.3838>.
- Thøgersen, J., and K. S. Nielsen. 2016. “A Better Carbon Footprint Label.” *Journal of Cleaner Production* 125: 86–94. <https://doi.org/10.1016/j.jclepro.2016.03.098>.
- UNWTO. 2019. “Transport-Related CO2 Emissions of the Tourism Sector: Modelling Results.” World Tourism Organization. <https://doi.org/10.18111/9789284416660>.
- Vandenbergh, M. P., T. Dietz, and P. C. Stern. 2011. “Time to Try Carbon Labelling.” *Nature Climate Change* 1, no. 1: 4–6. <https://doi.org/10.1038/nclimate1071>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table B1.** Full pairwise comparisons of estimated marginal means by flight choice. **Table C1.** Participants ranking each attribute as most important (rank = 1). **Table C2.** Cumulative link model predicting attribute rankings. **Table D1.** Preference matching between top-ranked attribute and chosen airline. **Table E1.** Covariate-adjusted multinomial logistic regression predicting flight choice. **Table E2.** Full pairwise comparisons of covariate-adjusted estimated marginal means by flight choice. **Table E3.** Covariate-adjusted binary logistic regression predicting booking propensity. **Table E4.** Covariate-adjusted pairwise contrasts of booking propensity between conditions.