ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Huber, Jürgen

Conference Paper

Collective Evidence on Behavioral Interventions Targeting Carbon Pricing Support: A Many-Designs Approach with 55 Studies

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2026: Behavioral Economics

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Huber, Jürgen (2026) : Collective Evidence on Behavioral Interventions Targeting Carbon Pricing Support: A Many-Designs Approach with 55 Studies, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2026: Behavioral Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:
https://hdl.handle.net/10419/343791

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Collective Evidence on Behavioral Interventions Targeting Carbon Pricing Support: A Many-Designs Approach with 55 Studies

Armando Holzknecht*[1,3] Rene Schwaiger*[1,3] Esther Blanco[2,3]

Juergen Huber[1,3] Michael Kirchler[1,3] co-authors[4]

February 25, 2026

[1] Department of Banking and Finance, University of Innsbruck, Innsbruck, Austria

[2] Department of Public Finance, University of Innsbruck, Innsbruck, Austria

[3] Members of the core organizing team that planned the project, pre-registered a pre-analysis plan, invited research teams, ran the experiments, analyzed the data, and wrote the paper.

[4]All members of the research teams (RTs) that submitted an experimental design and the corresponding software are full co-authors (see next page).

*Joint first authors

Corresponding Author: Jürgen Huber, juergen.huber@uibk.ac.at, Department of Banking and Finance, University of Innsbruck, Universitätsstraße 15, A-6020 Innsbruck, Austria.

Open Science: The study was preregistered and received approval by the Board for Ethical Questions in Science of the University of Innsbruck (Certificate of good standing, 30/2024). We provide the preregistered study protocol, data, and analysis files in an OSF repository at https://osf.io/ujvdn/?view_only=b46d7fab56d447c5ab516c8cbbf0602d. The project website was our main channel of communication https://manydesignscarbon.online.

Classification: Social Sciences, Environmental Social Science, Economics.

Keywords: Metascience, Behavioral Interventions, Climate Science.

Co-authors: Anna Lou Abatayo (Wageningen University and Research - Environmental Economics and Natural Resources Group), Magdalena Adamus (Centre of Social and Psychological Sciences, Slovak Academy of Sciences - Institute of Experimental Psychology), Lior Bakalo (Technion – Israel

Institute of Technology), Jesus Barreiro-Hurle (European Commission, Joint Research Centre (JRC)), Burcak Bas (KU Leuven - Department of Marketing), Piero Basaglia (Bordeaux School of Economics - Cluster of Excellence "Climate, Climatic Change, and Society" (CLICCS) and Center for Earth System Research and Sustainability (CEN)), Oindrila Bhattacharya (University of Canterbury - School of Psychology, Speech and Hearing), Ülkü Bıçakçı (ifo Institute & University of Erlangen-Nuremberg), Hayley Blunden (American University - Kogod School of Business), Michelle Blunt (King's College London), Sandra Bohmann (Ludwig-Maximilians-Universität München Department of Sociology), Pablo Brañas-Garza (Loyola Andalucia University - Department of Finance), Georgia Buckle (University of Portsmouth - School of Accounting, Economics and Finance), Lilith Burgstaller (Walter Eucken Institut), Romain Cadario (Erasmus University Rotterdam - Department of Marketing Management), Marina Carmela Orifici (FernUniversität in Hagen - Chair of Behavioral Economics and Intercultural Psychology), Juana Castro Santa (Universitat Autònoma de Barcelona - Institute of Environmental Science and Technology), Fadong Chen (Zhejiang University, School of Management, Neuromanagement Laboratory, The State Key Laboratory of Brain-Machine Intelligence), Doron Cohen (Carnegie Mellon University – Department of Engineering and Public Policy), Andreas Drichoutis (Agricultural University of Athens - Department of Agricultural Economics and Rural Development), Christoph Drobner (Central European University), Moritz Drupp (ETH Zurich - Department of Management, Technology, and Economics; University of Gothenburg - Department of Economics; University of Hamburg - Department of Economic), Elif E. Demiral (East Tennessee State University - Department of Management and Supply Chain), Charles Ebert (University of Wisconsin-Eau Claire - Marketing and Supply Chain Management Department), Laila Elina Nockur (Aarhus University - Department of Psychology and Behavioural Sciences), Hande Erkut (WZB Berlin Social Science Center), Filippos Exadaktylos (University of Barcelona - Department of Econometrics, Statistics and Applied Economics), Valeria Fanghella (EDHEC Business School, France), William Fawson (London Business School - Department of Organisational Behaviour), Lars Felder (DIW Berlin, Freie Universität Berlin), Kyle Fiore Law (Arizona State University - School of Sustainability), Hanna Fuhrmann-Riebel (German Institute of Development and Sustainability (IDOS)), Raphael Gottweis (WU Vienna University of Economics and Business - Department of Economics), Amélie Gourdon-Kanhukamwe (King's College London - Institute of Psychiatry, Psychology & Neuroscience), Manuel Grieder (UniDistance Suisse - Faculty of Business and Economics), Sven Gruener (Martin Luther University Halle-Wittenberg - Department of Environmental Sociology), Peter Haan (DIW Berlin, Freie Universität Berlin), Hooman Habibnia (Vienna University of Business and Economics - Institute for Cognition and Behavior), Sebastian Hafenbrädl (IESE Business School - Department of Managing People in Organizations), Emeric Henry (Sciences Po - Department of Economics), Atar Herziger (Technion - Israel Institute of Technology - Faculty of Data and Decision Sciences), Geraldine Holenweger (University of Bern - Institute of Marketing

and Management), Christoph Huber (Aalto University - Department of Finance), Sebastian J. Goerg (Technical University Munich - TUMCS for Biotechnology and Sustainability), Nicolas Jacquemet (University Paris 1 Panthéon-Sorbonne - Department of Economics), Tao Jin (University of Minnesota - Department of Psychology), Fabian Kalleitner (Ludwig-Maximilians-Universität München - Department of Sociology), Cankut Kuzlukluoglu (Humboldt University of Berlin), Kyle Fiore Law (Arizona State University - School of Sustainability), Kanchan Mukherjee (Indian Institute of Management Bangalore - Organizational Behavior & Human Resources Management), Matthias Kasper (Walter Eucken Institut), Austėja Kažemekaitytė (University of Trento - Department of Economics and Management), Tobias König (Linnaeus University - Department of Economics and Statistics), Christophe Lembregts (Erasmus University Rotterdam - Department of Marketing Management), Shuzhen Li (Utrecht University – Utrecht School of Economics), Kristian López Vargas (University of California Santa Cruz, USA, and Universidad del Pacífico, Perú), Wolfgang Luhan (University of Portsmouth - School of Accounting, Economics and Finance), Vincent Mak (University of Cambridge - Judge Business School), Aitor Marcos (University of the Basque Country UPV/EHU - Faculty of Economics and Business), Diego Marino Fages (Durham University - Department of Economics), Paul Matthias Gorny (Karlsruhe Institute of Technology (KIT) - Chair of Human Resource Management), Patrick Maus (University of Heidelberg - Alfred Weber Institute for Economics), Friederike Mengel (University of Essex - Department of Economics), Ilana Mermelstein (University of Michigan - School of Information), Arjun Mitra (Indian Institute of Technology Kanpur - Department of Cognitive Science), Andras Molnar (University of Michigan - Department of Psychology), Laura Moritz (University of Trento - Center Agriculture Food Environment - C3A), Irene Mussio (Leeds University Business School - Department of Economics and Centre for Decision Research), Oliver Musshoff (Humboldt University of Berlin, Farm Management Group), Sarah Necker (ifo Institute & University of Erlangen-Nuremberg), Petra Nieken (Karlsruhe Institute of Technology (KIT) - Chair of Human Resource Management), Marco Palma (Texas A&M University - Department of Agricultural Economics), Navya Pandit (University of Cologne - Department of Corporate Development), Tyler Parker-Rollins (Ohio University - Psychology Department), Julien Picard (Politecnico di Milano - Department of Management, Economics and Industrial Engineering), Ori Plonsky (Technion - Israel Institute of Technology - Faculty of Data and Decision Sciences), Raghabendra Pratap KC (Rollins College - Department of Business), Giulia Priolo (Copenhagen Business School - Department of Management, Society and Communication), Julian Quandt (Vienna University of Business and Economics - Institute for Cognition and Behavior), Todd R. Kaplan (University of Haifa – Department of Economics), Zahra Rahmani Azad (University of Basel – Department of Psychology), Rastislav Rehák (Max Planck Institute for Behavioral Economics), Andries Richter (Wageningen University and Research - Environmental Economics and Natural Resources Group), Anette Rohmann (FernUniversität in Hagen - Chair of Community Psychology),

Jens Rommel (Swedish University of Agricultural Sciences - Department of Economics), Yefim Roth (University of Haifa - Department of Human Service), Umit Saglam (East Tennessee State University - Department of Management and Supply Chain), Melania Salazar-Ordoñez (Universidad de Córdoba – Water, Environmental and Agricultural Resources Economics (WEARE) Research Group), Renke Schmacker (Berlin Social Science Center (WZB) - Department "Market Behavior"), Jan Schmitz (Radboud University - Department of Economics and Business Economics), Claudia R. Schneider (University of Canterbury - School of Psychology, Speech and Hearing), Michael Schulte-Mecklenbeck (University of Bern - Institute of Marketing and Management), Garima Sharma (American University - Kogod School of Business), Christoph Siemroth (University of Essex - Department of Economics), Nisheeth Srivastava (Indian Institute of Technology Kanpur - Dept of Computer Science & Engineering), Jakub Šrol (Centre of Social and Psychological Sciences, Slovak Academy of Sciences - Institute of Experimental Psychology), Ivo Steimanis (Philipps University Marburg - Group of Sustainable Use of Natural Resources), Stylianos Syropoulos (Arizona State University - School of Sustainability), Iris Vilares (University of Minnesota - Department of Psychology), Vivek (independent researcher), Tobias Vorlaufer (Leibniz Centre for Agricultural Landscape Research (ZALF) - Ecosystem Service Governance), Matthew Walker (Newcastle University Business School - Newcastle University Business School, Economics Subject Group), Francisco Wang (University of Cambridge - Judge Business School), Elke Weidenholzer (University of Essex - Department of Economics), Simon Weidenholzer (University of Essex - Department of Economics), Yilong Xu (Utrecht University – Utrecht School of Economics), Han Yin (Zhejiang University, School of Management, Neuromanagement Laboratory, The State Key Laboratory of Brain-Machine Intelligence), Junhao Vincent Yu (Miami University – Farmer School of Business), Xian Zhao (Ohio University - Psychology Department)

Abstract

There is an urgent need to reduce carbon emissions globally to limit the damages caused by climate change. Most economists agree that a carbon price is an effective and cost-efficient policy to mitigate emissions, yet low public acceptance and limited political support remain major barriers to its widespread implementation. This crowdsourced "many-designs" project presents results from 55 behavioral interventions on real-world support for carbon pricing, independently developed by international research teams randomly selected from an initial pool of 135 applications. By implementing the interventions simultaneously with almost 20,000 U.S. residents, this pre-registered study ensures the comparability of results, accelerates scientific knowledge generation, and reduces the risk of scientific malpractices. The results show very small positive but statistically significant effects of behavioral interventions on real-world support, and stated support, including the willingness to endorse a carbon price that internalizes the social costs of \$120 per ton of CO_2 emissions (Cohen's d's: 0.04–0.08). Put differently, this entails an increase in support for carbon pricing across measures of around two percentage points. Furthermore, the results reveal low-to-medium between-study heterogeneity (τ: 0.07–0.12, I^2: 32%–57%). Lastly, we identify strong overconfidence among research teams regarding the expected effects of their interventions and those of their peers, indicating a potential miscalibration of community expectations.

Introduction

Behavioral interventions have been widely tested, and evidence shows that their effects on individual behavioral change are often small. Meta-analyses and large-scale "mega-studies" have raised concerns about publication bias (DellaVigna and Linos, 2022; Maier et al., 2022), questioned the overall impact of behavioral interventions (Mertens et al., 2022; Kaiser et al., 2025), and even documented instances of backfiring effects (DellaVigna and Linos, 2022) with unintended consequences. While these studies deepen our understanding of the effectiveness of behavioral interventions in promoting individual change via i-frame policies, considerably less focus has been placed on their capacity to influence s-frame policies and, in turn, systemic change (Connolly et al., 2025; Granulo et al., 2025). Thus, a more ambitious path is to test whether behavioral interventions foster public support for regulations that lead to behavioral change. This can be a relevant question for tackling social concerns, such as climate change, on which public opinion is divided (Umit and Schaffer, 2020; Andre et al., 2024; Carson et al., 2025). Minor changes in support for climate policies may be sufficient to change election outcomes or induce tipping points in public perceptions (Otto et al., 2020; IPCC, 2022). Recent innovative large-scale studies have advanced the field by

investigating interventions designed to increase support for a variety of climate policies (see, e.g., Vlasceanu et al., 2024; Dechezleprêtre et al., 2025). We focus on carbon pricing, which is widely regarded by scientists as effective (Hájek et al., 2019; Stechemesser et al., 2024) and cost-efficient for reducing carbon emissions (Nordhaus, 1992; Barrage and Nordhaus, 2024). For example, the European Union Emission Trading System reduced 3.8% of total EU-wide emissions compared to a world without carbon pricing, despite the low trading prices for CO_2 permits (Bayer and Aklin, 2020). Moreover, a worldwide analysis of climate policies showed that policy mixes that included price-based instruments were particularly effective in reducing emissions (Stechemesser et al., 2024). However, despite carbon pricing being actively discussed by various governments, only 28% of global carbon emissions are covered, with an average carbon price of \$19 per tCO_2e in 2025 (World Bank, 2025), much lower than estimates for a socially efficient carbon price (\$120 per tCO_2e, see, U.S. Environmental Protection Agency, 2023). Overall, political support appears to be insufficient for more widespread implementation (see, e.g., Dechezleprêtre et al., 2025).

This study presents the results of the first "many-designs" crowdscience study on carbon pricing support (for the methodology, see, e.g., Huber et al., 2023). By employing a "many-designs" approach, it expands the knowledge frontier in several ways: First, it uses the expertise of 55 independent research teams (RTs), randomly selected from 135 international applicants, to design individual studies. Thus, this "many-designs" approach differs from the classical standard of "one research team, one design, one dataset" by benefiting from multiple design ideas of, among others, environmental economists, psychologists and behavioral scientists. Second, it compiles evidence from 55 independent randomized controlled trials. Unlike typical mega-studies (Harring et al., 2023; Andre et al., 2024; Vlasceanu et al., 2024; Woerner et al., 2024; Dechezleprêtre et al., 2025; Sinclair et al., 2025), each study design presented herein includes both a treatment *and* a control condition *and* a measure for real-world policy support as the outcome. This yields robust, high-powered meta-scientific evidence. Third, it speeds up the scientific knowledge generation process by years, with independent studies on the same research question being developed simultaneously rather than sequentially. Fourth, it ensures consistency by using a common experimental participant pool and a unified econometric analysis, as opposed to traditional meta-analyses. Therefore, this "many-designs" setup excludes population and analytical heterogeneity in evaluating the different interventions (Wicherts et al., 2016; Landy et al., 2020; Simonsohn et al., 2020; Menkveld et al., 2023; Holzmeister et al., 2024). Moreover, it transparently reports the results of all 55 designs regardless of their findings, mitigating publication bias, and does not exclude proposals ex-ante based on the subjective judgment of the principal investigators. We pre-registered an extensive pre-analysis plan in the OSF project page that outlines the entire structure, procedure, and main part of the data analysis. This setup prevents scientific malpractices such as p-hacking, HARKing, and selective reporting. Finally, we innovate by extending policy support outcomes to real-life

measures of policy support proposed by the RTs. Previous literature commonly drew on intentions and self-reports to measure policy support (Woerner et al., 2024; Sinclair et al., 2025), whereas real-life measures have only been used to capture pro-environmental individual behaviors (see, e.g., Andre et al., 2024; Vlasceanu et al., 2024; Dechezleprêtre et al., 2025).

Each RT's main task was to independently design a study featuring a policy treatment condition (TREATMENT) and a control condition (CONTROL) to answer the following research question: "What is the impact of behavioral interventions on the real-world support for a price of carbon?" This required the RTs to develop an outcome measure with real-world implications for the support of carbon pricing. Some examples from the set of studies are donating to NGOs that lobby for carbon pricing, signing petitions requesting the implementation of carbon pricing, writing letters to government officials asking to advance towards carbon pricing, or publishing advocacy-oriented posts for carbon pricing on social media. Detailed descriptions of each study, along with an overview of intervention categories and outcome variables, are provided in Table S.1 and Figures S.18 and S.19 in the Supplementary Information (SI); full details are available via the corresponding pre-registrations in our OSF repository. Each RT developed the software for their online study and assessed its expected effect size in Cohen's d units (self-assessment). Furthermore, RTs were randomly assigned to assess 10 other RTs' designs (peer assessment; see Methods). For an overview of the participating RTs' geographical regions and research fields, see Figures S.16 and S.17 in the SI. From January to March 2025, we simultaneously ran all 55 studies on a U.S. general population Prolific sample (Palan and Schitter, 2018) with 19,558 participants (i.e., approximately 356 participants per design—thus, 178 per condition—based on a pre-registered power simulation for the meta-effect; see SI, Figures S.20 and S.21). Each participant was randomly assigned to either the TREATMENT or CONTROL condition of one of the 55 studies (i.e., a random assignment to 1 of 110 [55 $\times$ 2] experimental conditions (see SI, Section 8 and Table S.15 for descriptive statistics on covariate balancing). We selected the United States because it is one of the world's largest emitters, lacks a federal carbon tax, and exhibits highly heterogeneous attitudes towards such policies (Baldwin and Lammers, 2016; Anderson et al., 2023).

In each study, participants' support for carbon pricing was first measured through the real-world support metric for a carbon price proposed by the RTs. Next, participants completed a follow-up questionnaire ("policy survey") based on a subset of questions in Vlasceanu et al., (2024) and Dechezleprêtre et al., (2025), including five measures of stated support for carbon pricing, which were constant across all studies. Altogether, we used (i) the observed real-world policy support variable of each RT ("General support [observed]"), (ii) a general stated support measure for carbon taxation ("General support [stated]"), (iii) a stated support measure for an increase of carbon tax ("General price increase"), (iv) a stated support measure for a carbon tax equal to the social cost of

carbon of \$120 ("Social price") based on the lower-bound estimates from the U.S. Environmental Protection Agency (2023), (v) a stated support measure for a carbon tax with carbon dividends ("Tax with rebates") and (vi) a stated support measure for a carbon tax equal to \$120 with carbon dividends ("Social price with rebates"; see the SI, Section 9, for the full policy survey). This set of six variables constitutes the primary outcomes for Hypothesis 1 and 2. Variables (ii), (iv), (v), and (vi) were measured on five point Likert scales (-2: strongly oppose, 2: strongly support). Variable (iii) was measured on a continuous scale of 0–100.

We formulated 10 pre-registered hypotheses (see SI, Section 4, for the exact formulation). Hypothesis 1 states that behavioral interventions increase support for a price on carbon, as measured by the primary outcome variables. Hypothesis 2 claims that there is heterogeneity in the estimated effect sizes for the primary outcome variables across RTs. Hypothesis 3 concerns the impact of behavioral interventions on participants' perceptions about carbon pricing (i.e., their fairness and effectiveness), and about climate change (i.e., adoption of climate friendly behavior, worries about impacts, technical feasibility of net-zero, and the potential harm on one's own life), which we pre-registered as "secondary" outcome variables. Hypotheses 4–9 revolve around the associations between the observed effects on real-world support and three peer-assessed variables of the studies as potential moderators: the informativeness of the intervention (Hypothesis 4), the corresponding intervention category (Hypothesis 6), and the expected effect size (Hypothesis 8). Hypotheses 5, 7, and 9 concern residual heterogeneity after controlling for each of the three moderators, respectively. Finally, Hypothesis 10 states that RTs' expectations of their own effect sizes do not significantly differ from the observed effect sizes.

Results

For all analyses, we pre-registered a significance threshold of $\alpha = 0.5\%$ (Benjamin et al., 2018). Accordingly, we refer to findings with a p-value between 0.5% and 5% as suggestive evidence. Unless stated otherwise, all analysis specifications in this paper were pre-registered, including the set of covariates—covering sociodemographic variables, private behavior, and trust in government—added to the multivariate ordinary least squares (OLS) regressions (see SI, Section 7, for details).

Panel (a) of Figure 1 addresses Hypotheses 1 and 2. It shows meta-analytical effect sizes (Cohen's d) and heterogeneity estimates (Cochran's Q, I^2, and τ) based on random effects meta-analyses (DerSimonian-Laird model) for all six primary outcome variables (i)–(vi). The effect sizes of the individual studies are obtained either via univariate OLS regressions with an intervention treatment dummy variable (green diamond) or a multivariate model additionally including control variables (orange square). We find positive yet very small (Cohen's d: 0.04–0.08) statistically significant

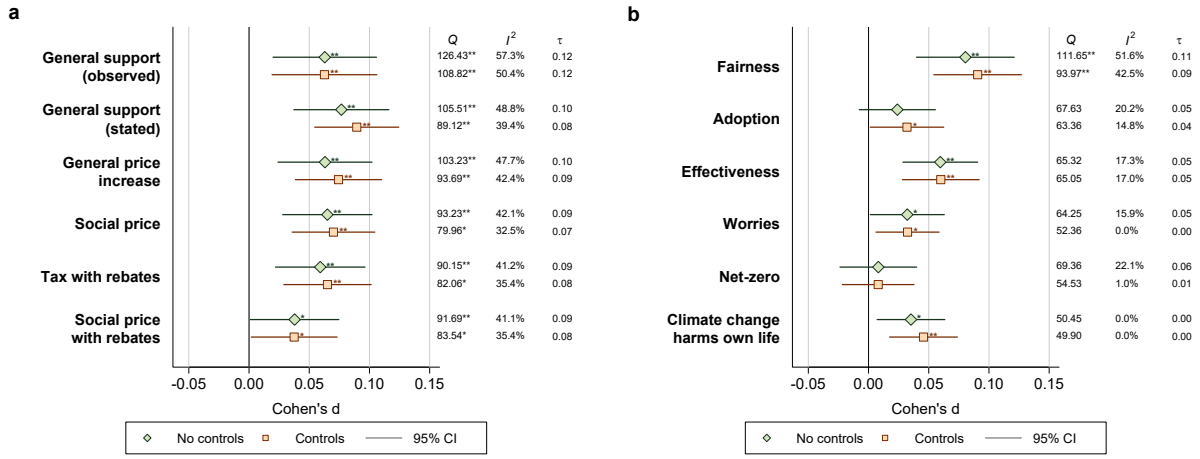

Figure 1 Panel (a): **Meta-analytical results across primary outcomes for Hypotheses 1 and 2.** Meta-analytical point estimates and the 95% CIs of random-effects meta-analyses based on effect sizes obtained from univariate (green diamonds) and multivariate (orange rectangles; see SI, Section 7, for the set of covariates) OLS regressions for each of the 55 studies (*Hypothesis 1*). "General support (observed)" denotes the outcome variables suggested by the RTs, reflecting real-world support for carbon pricing, while the other variables were part of a common survey at the end of the experiment measuring stated support for a price on carbon (see SI, Section 9). "General support (stated)" measures individual's general support for a carbon price. "Social price" denotes support for a carbon price equal to the social price of carbon of \$120/t CO_2. "Tax with rebates" measures support for a general carbon tax with cash transfers as redistribution to households. "Social price with rebates" denotes an individual's support for a carbon tax equal to the social price of carbon of \$120/t CO_2 with cash transfers as redistribution to households. All variables were measured on five point Likert scales (-2: strongly oppose, 2: strongly support), except the variable "General price increase", which was measured on a range slider from 0 (not at all) to 100 (very much). ** and * indicate statistical significance ($p < .005$) and suggestive evidence ($p < .05$), respectively, and are based on one-sided z-tests (due to the directionality of *Hypothesis 1*). Cochran's Q (stars indicate statistical significance for testing heterogeneity), I^2, and τ indicate effect size heterogeneity (*Hypothesis 2*). Detailed metrics are shown in Table S.4 in the SI. Panel (b): **Meta-analytical results across secondary outcomes for Hypothesis 3.** Meta-analytical point estimates and the 95% CIs of random-effects meta-analyses based on effect sizes obtained from univariate (green diamonds) and multivariate (orange rectangles; see SI, Section 7, for the set of covariates) OLS regressions for each of the 55 studies. All variables were part of a common survey battery at the end of each study measuring beliefs and perceptions about carbon pricing and climate change, as well as individuals' willingness to adopt climate friendly behaviors (see SI, Section 9). "Fairness" measures individual's agreement on whether a price on carbon is fair. "Adoption" is an index measuring an individual's willingness to adopt climate friendly behaviors, including the willingness to limit air travel, driving, beef consumption, heating or cooling and to get an e-vehicle. "Effectiveness" is an index of an individual's belief in the effectiveness of a carbon tax in addressing climate change and its economic impacts. This comprises the beliefs on whether such a tax would reduce driving, air pollution, and the use of fossil fuels, encourage building insulation, and whether such a tax would have a negative effect on the U.S. economy and its employment (reverse coded). "Worries" is an index of an individual's perceived likelihood of climate change leading to the following events: droughts and heatwaves, volcanic eruptions, rising sea levels, declines in agricultural production, reduced standards of living, increased migration flows, more armed conflict, and the extinction of humankind. "Net zero" is an individual's belief about the technical feasibility to stop greenhouse gas emissions by the end of the century while maintaining satisfactory standards of living in the United States. "Climate change harms own life" measures an individual's belief about whether climate change is already affecting, or will soon negatively affect, their personal life. ** and * indicate statistical significance ($p < .005$) and suggestive evidence ($p < .05$), respectively, and are based on two-sided z-tests (due to the non-directionality of *Hypothesis 3*). Cochran's Q (stars indicate statistical significance for testing heterogeneity), I^2, and τ indicate effect size heterogeneity (*Hypothesis 2*). Detailed metrics are shown in Table S.5 in the SI.

meta-effects of behavioral interventions across most of the measures. Specifically, results show statistically significant positive meta-effects on real-world support ("General support [observed]": $d = 0.063$, 99.5% CI: [0.001; 0.124], $z = 2.869$, $k = 55$, $p = 0.002$), stated general support for carbon pricing ("General support [stated]": $d = 0.077$, 99.5% CI: [0.020; 0.133], $z = 3.812$, $k = 55$, $p < 0.001$), stated support for an increase in the price of carbon ("General price increase": $d = 0.063$, 99.5% CI: [0.007; 0.119], $z = 3.168$, $k = 55$, $p = 0.001$), stated support for a carbon price internalizing the social cost of carbon at \$120 per ton of CO_2 emissions ("Social price": $d = 0.065$, 99.5% CI: [0.012; 0.118], $z = 3.439$, $k = 55$, $p < 0.001$), and stated support for a carbon price with

carbon dividends ("Tax with rebates": $d = 0.059$, 99.5% CI: [0.006; 0.112], $z = 3.096$, $k = 53$, $p = 0.001$). Suggestive evidence also points to a positive effect on the support for a carbon price at the social cost of carbon with carbon dividends ("Social price with rebates": $d = 0.038$, 99.5% CI: [-0.015; 0.090], $z = 2.018$, $k = 55$, $p = 0.023$). Controlling for the pre-registered covariates yields effect sizes ranging from $d = 0.04$ to $d = 0.09$, with qualitatively equivalent results. Placing these findings in perspective, the treatment conditions raise average support for carbon pricing by approximately two percentage points across all measured outcomes compared to the control (see Figures S.2 and S.3).

Panel (a) of Figure 1 reports several estimates of effect size heterogeneity relevant to Hypothesis 2. Cochran's Q is a test-statistic indicating potential heterogeneity in effect sizes. Based on this metric, we find statistically significant evidence of between-study heterogeneity for all primary outcome variables (i)—(vi). The pre-registered quantification measure of heterogeneity we use is τ, which is the estimated standard deviation of true effect sizes and represents the variation in effect sizes across RT designs other than sampling variation. The values of τ range from 0.074 for the general stated support of a social carbon price with control variables (99.5% CI [0.000, 0.138]) to 0.123 for the observed real-world support measure used by the RTs without control variables (99.5% CI [0.071, 0.193]). We find I^2 values between 32% and 57%. Since I^2 is the share of true between-study heterogeneity in total variance, this implies that roughly one-third to half of the observed variation reflects true heterogeneity rather than within-study sampling variance. This can be classified as small-to-medium between-study heterogeneity (see Holzmeister et al., 2024). All other pre-registered heterogeneity estimates are shown in Table S.4. Overall, the relatively similar level of τ between the observed and stated measure for general support suggests that the primary source of between-study heterogeneity is not related to differences in outcome measures across these two variables but rather differences in the effectiveness of the interventions compared to their respective control conditions.

Panel (b) of Figure 1 provides results for Hypothesis 3 on the impact of behavioral interventions on secondary outcome variables measured via the policy survey. We calculate meta-analytical effect sizes based on the same econometric and visual approach as for Hypotheses 1 and 2. Due to the non-directionality of Hypothesis 3, we report results of two-sided z-tests (see SI, Figure S.4, for detailed results on each of the index items). We find significantly positive effects of behavioral interventions on perceptions about fairness of carbon pricing ($d = 0.08$, 99.5% CI: [0.022; 0.138], $z = 3.858$, $k = 55$, $p < 0.001$) and its effectiveness ($d = 0.06$, 99.5% CI: [0.015; 0.104], $z = 3.767$, $k = 55$, $p < 0.001$). Additionally, we report suggestive positive evidence for the effectiveness of interventions on worries about climate change ($d = 0.03$, 99.5% CI: [-0.012; 0.076], $z = 2.053$, $k = 55$, $p = 0.040$) and the belief in climate change harming one's own life ($d = 0.04$, 99.5% CI: [-0.005; 0.076], $z = 2.457$, $k = 55$, $p = 0.014$). However, we do not find statistically significant meta-effects

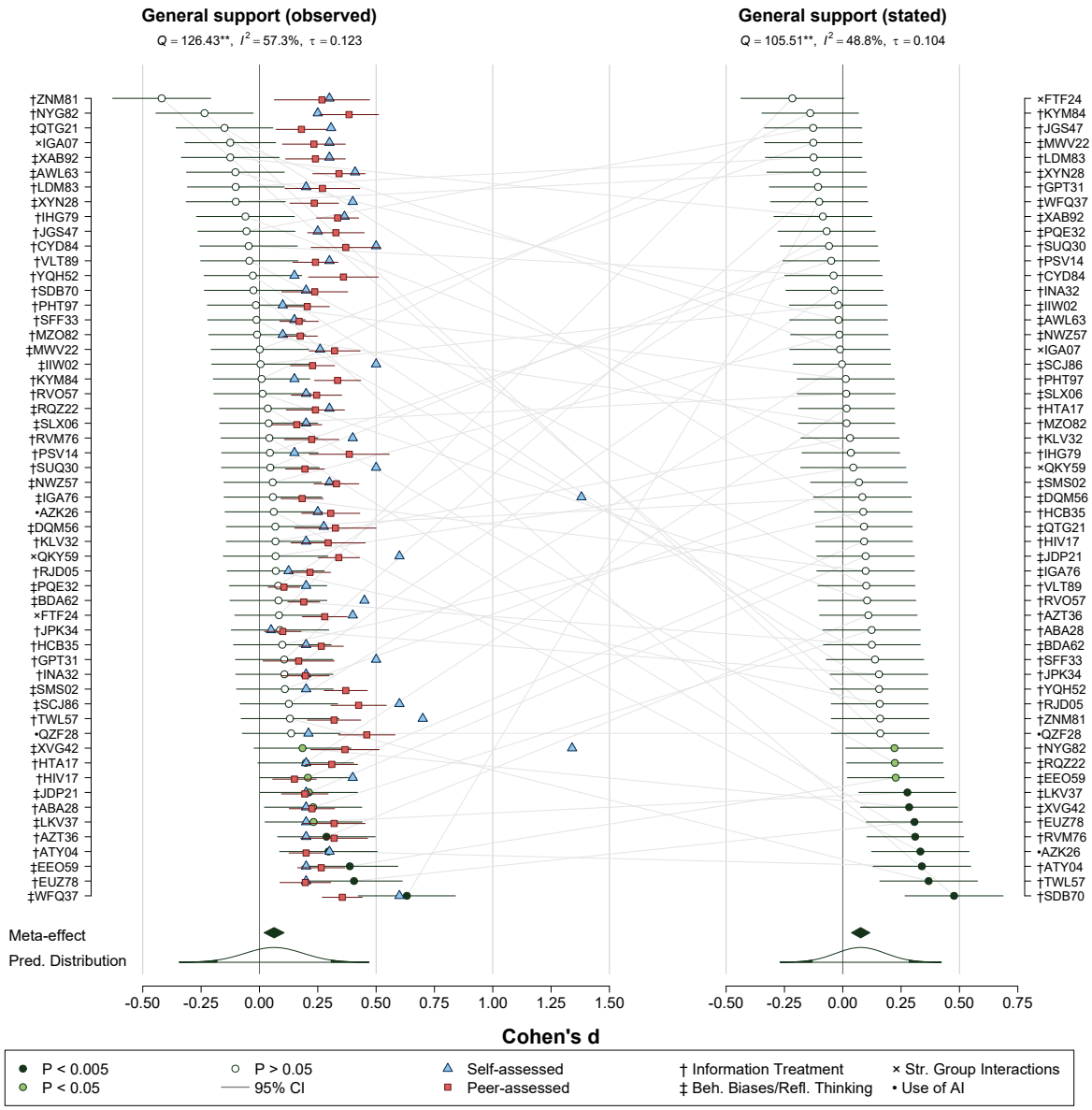

Figure 2 Forest plots of the meta-analytic results for real-world and general stated support for carbon pricing. The figure displays point estimates and 95% confidence intervals for the treatment effects on RTs' real-world measure ("General support [observed]", left) and the stated general support measure common to all teams ("General support [stated]", right) across all 55 designs. The color shading of the circles indicates statistical significance (dark green), suggestive evidence (light green), and non-significance (white), respectively, and are based on one-sided t-tests (due to the directionality of *Hypothesis 1*). Dark green diamonds represent the overall meta-analytic effects with the width of the diamond representing 95% confidence intervals. Furthermore, the plot shows the predictive distribution, also indicating the 95% prediction intervals by the white area under the curve. These metrics incorporate both the meta-effect's standard error and the between-study heterogeneity τ (all heterogeneity measures are shown at the top). For the "Observed" panel, blue triangles additionally indicate each RT's belief about its own effect size, and red rectangles (with 95% CIs) depict the average peer-belief estimates for the given study. Grey lines connect each RT's estimates across the two outcome variables.

for the self-reported willingness to adopt individual climate-friendly behavior ($d = 0.02$, 99.5% CI: [-0.021; 0.069], $z = 1.485$, $k = 55$, $p = 0.138$), nor for the belief in net-zero being technically feasible ($d = 0.01$, 99.5% CI: [-0.037; 0.054], $z = 0.617$, $k = 55$, $p = 0.500$). Controlling for pre-registered covariates yields effect sizes ranging from $d = 0.01$ to $d = 0.09$, with the effects on

willingness to adopt climate-friendly behavior becoming positive suggestive evidence, and effects on perceived harms of climate change to one's own life becoming statistically significantly positive. Moreover, we only find significant evidence of between-study heterogeneity in fairness perceptions. The estimated τ ranges from 0.00 to 0.11 and I^2 from 0% to 52%.

Next, we unpack the analysis and report estimated effects for each individual study. The left panel of Figure 2 displays a forest plot of effect sizes across all 55 designs on the RTs' chosen real-world support measure (circles with 95% confidence intervals), each RT's reported beliefs about the expected effect of their design ("self-assessment," blue triangles), and average reported beliefs by 10 other RTs ("peer assessment," red rectangles with 95% confidence intervals). Exploratory two-sided t-tests show that most (44) studies' effect sizes are non-significant at a p-value of 5%, three show a significantly positive effect ($p < 0.005$), six show a suggestive positive effect ($0.005 < p < 0.05$), one shows a suggestive negative effect ($p = 0.028$), and one exhibits a significantly negative effect ($p < 0.001$; see SI, Table S.2 for details). The predictive distribution of true effect sizes is shown below the meta-analytic effect, incorporating both within- and between-study heterogeneity. The predictive distribution extends across both the positive and negative domain, indicating that a newly implemented intervention from the same design space could plausibly yield effects in either direction, including potential backfiring effects. Specifically, while 69.2% of newly conducted studies are expected to yield positive effects (i.e., Cohen's $d > 0$), 30.8% of newly conducted studies from the same design space are expected to yield negative or null effects (i.e., Cohen's $d \leq 0$). Thus, the 95% prediction interval, which is more commonly reported in meta-analyses, spans negative and positive values, indicated by the left and right borders of the white area (95% PI: [-0.182; 0.308]).

Concurrently, expectations about the effect sizes for one's own study, as well as the expectations by other RTs, are overly optimistic. Each RT expected their intervention, as well as their peers' interventions, to have a positive effect. Their average expectations significantly exceed the materialized effects of behavioral interventions (effect size in Mann-Whitney-U-tests observed vs. self-assessed belief: $r = 0.668$, 95% CI [0.540, 0.770], $p < 0.005$; observed vs. peer-assessed beliefs: $r = 0.672$, 95% CI [0.540, 0.790], $p < 0.005$; see SI, Figure S.5), contrary to Hypothesis 10.

As a counterpart, the right panel of Figure 2 shows an equivalent forest plot of effect sizes across designs on the stated general support measure ("Do you support or oppose a carbon tax in general?"). We rank the studies from lowest (top) to largest effect sizes (bottom; see SI, Table S.3 for details). Similarly, the predictive distributions spans both negative and positive values, although more leaned towards positive effects: 76% of newly conducted studies from the same design space are expected to yield, positive effects and 24% are expected to yield negative or null effects. The 95% prediction interval confirms this (95% PI: [-0.131; 0.285]).

We also connect the RT's effects across measures with gray lines. Perfect correlation in effect sizes

across measures would yield horizontal lines, which is not what we observe in general (Pearson's correlation coefficient between real-world and stated support: $r = 0.14$, 95% CI [-0.094, 0.421], $p = 0.202$). For example, the study with the largest observed real-world effect ("WFQ37") shows one of the lowest stated effects, while some studies with larger stated effects entail smaller and non-significant real-world effects (e.g., "SDB70" or "RVM76"). Nevertheless, some studies show consistent positive effects across both outcome measures (e.g., "EUZ78", "EE059", and "ATY04"). Figure S.1 in the SI presents the comparisons of individual and meta-analytical effect sizes for all primary outcomes (i)–(vi).

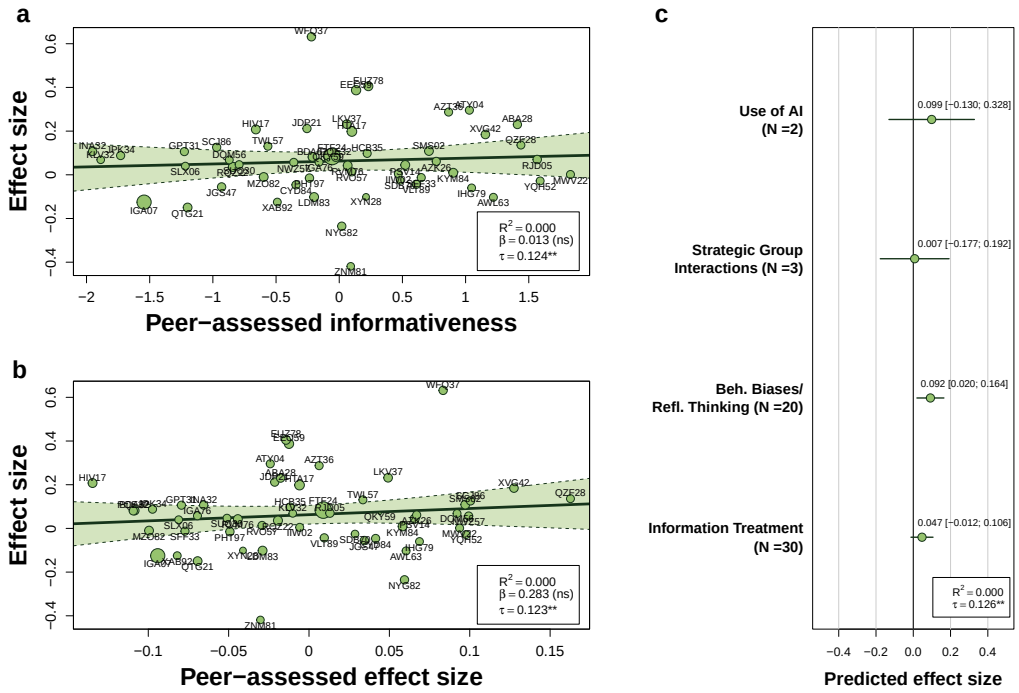

Figure 3 Panel (a): **Relationship between observed effect sizes and peer-assessed informativeness**. Plotted are the 55 estimated univariate effect sizes on real-world support (Cohen's d) over the average peer-assessed informativeness rating (demeaned) of the experimental designs. The green shaded area indicates 95% CIs of the regression line. Bubbles indicate the effects of the individual studies, with the size being inversely related to the standard errors of the effects. Additionally, residual between-study heterogeneity τ is added to the legend with stars indicating significance based on a test for heterogeneity (*Hypotheses 4 and 5*). Panel (b): **Relationship between observed effect sizes and peer-assessed effect size.** Plotted are the 55 estimated univariate effect sizes (Cohen's d) over the average peer-assessed effect (demeaned) of the experimental designs. The green shaded area shows 95% CIs of the regression line. Bubbles indicate the effects of the individual studies, with the size being inversely related to the standard errors of the effects. Additionally, residual between-study heterogeneity τ is added to the legend with stars indicating significance based on a test for heterogeneity (*Hypotheses 6 and 7*). Panel (c): **Relationship between observed effect sizes and behavioral intervention categories (peer-assessed and AI classified).** Plotted are the predicted values and their 95% CIs of effect sizes for each of the four categories, i.e., "Use of AI," "Strategic Group Interactions," "Behavioral Biases/ Reflective Thinking," and "Information Treatment." The estimates are based on a meta-regression using the uncontrolled OLS estimates of general observed policy support (*Hypotheses 8 and 9*).

Next, we investigate which characteristics of studies explain differences in observed real-world support across studies. We begin with the pre-registered moderators based on the peer assessments. Figure 3 presents the results of three meta-regressions, each testing one potential moderator (see Huber et al., 2023): Panel (a) concerns the perceived informativeness of the intervention, Panel (b) deals with the expected effect size by peers, and Panel (c) focuses on the intervention category (see

SI, Section 6, for the classification procedure; complete regression outputs are provided in Table S.7). None of the three moderators significantly predicts observed treatment effects on real-world support (Hypotheses 4, 6, and 8). Consequently, significant residual heterogeneity remains (τ: 0.124–0.129; Hypotheses 5, 7, and 9). In Panel (c), we report predicted meta-effects based on the meta-regression. Although the effect sizes on observed support do not differ significantly between categories, we find suggestive evidence that interventions targeting behavioral biases and reflective thinking tend to produce positive effects on average ($d = 0.09$, 95% CI: [0.010; 0.164]).

Examining the studies that yield consistent positive results in at least one of our primary outcome measures, we noted the following: A first block of studies, namely "ATY04," "AZT36," "EUZ78," "SDB70," "XVG42," which considers the effect of information interventions and considers a control condition where participants are provided information about climate change and carbon pricing, with different degrees of detail across studies. A second block of studies that lets participants actively interact with new information; specifically "AZK26" considers interaction with a chatbot, and "EE059" considers writing essays. Two additional studies, "TWL57" and "WFQ37," consider experimental decisions that are economically incentivized and vary the in-game institutions between control and treatment. Lastly, "LKV37" considers a package of interventions (see SI, Table S.1, for details on the designs).

Additionally, after data collection, we conducted an exploratory survey among RTs to gather more information about the study's design and researchers' attributes (see SI, Section 1 for details on the questionnaire). First, using data clustering (see SI, Section 1), we identified two distinct clusters of RTs based on observable characteristics elicited via the questionnaire (see SI, Figure S.14). The first cluster is defined by the prevalence of present-focused interventions, interventions not yet established in previous literature, information treatments, belief and knowledge updating as the main mechanisms, and economists as researchers. Cluster 2 is mainly characterized by interventions targeting behavioral biases and reflective thinking. Interestingly, in exploratory meta-regressions, we found suggestive evidence of interventions belonging to the second cluster to generate larger effect sizes regarding real-world support ($\beta = 0.114$, 95% CI [0.016; 0.212], $z = 2.279$, $p = 0.023$; see Table S.13). Second, using exploratory factor analysis for mixed data, we identified two factors (see SI, Section 1). Factor 1 captures motivation-based interventions, future-focused interventions and, as a main mechanism, behavioral biases and reflective thinking. Factor one is positively associated with real-world support ($\beta = 0.025$, 95% CI [0.001; 0.049], $p = 0.023$; see SI, Table S.13). No significant associations were found for Factor 2 or for stated support.

Finally, we conducted several robustness checks (see SI for details). First, we ran a multiverse analysis with 32 pre-registered analytical specifications on the studies' individual effect sizes of observed and stated (exploratory) carbon pricing support (see SI, Table S.8). Across all specifications,

results regarding meta-effects on observed real-world and stated support and the associated between-study heterogeneity remain robust and qualitatively identical to the ones reported in the main analysis (median z-statistic observed: 2.636; stated: 4.256; see SI, Figure S.6 and Figure S.8). Additionally, we depict noteworthy but unsystematic variations in obtained Cohen's d estimates for each RT throughout all analysis specifications in the multiverse. For example, the inclusion of control variables increases effects in absolute terms for some RTs, while for other RTs, it attenuates the effects obtained by univariate OLS regressions (see SI, Figure S.7 and Figure S.9). Second, we applied robustness checks for meta-research (i.e., the leave-one-out method) to the results regarding real-world and stated support (see SI, Figure S.10 and Figure S.12), and explored the variation of meta-effects and heterogeneity values across one million different random subsets of study combinations, also known as *GOSH* plots (see SI, Figure S.11, Figure S.13, Table S.9, and Table S.12). Results throughout all these robustness diagnostics support the main conclusions reported within the results section.

Discussion and Conclusion

Carbon taxes are widely discussed political instruments for reducing carbon emissions to mitigate global warming. This project represents the first effort to systematically leverage the creativity of the international research community to design behavioral interventions to foster support for carbon pricing. Relying on a "many-designs" methodology, we presented results of 55 experimental studies independently proposed by international RTs, without any pre-selection of studies or outcomes. Compared to relying on isolated, uncoordinated studies, this approach fosters the development of diverse ideas and accelerates the generation of scientific knowledge on this critical topic. In this study, impact assessment includes real-world measures of policy support, advancing toward the external validity of findings. Combined with our pre-registered analysis plan, the knowledge generated is transparently presented, free from selective reporting, p-hacking, HARKing, and forking (e.g., Gelman and Loken, 2014; Munafò et al., 2017; Rubin, 2017; Simonsohn et al., 2020).

Overall, we found positive and statistically significant meta-effects of behavioral interventions on both real-world outcomes (Cohen's d: 0.04–0.08) and a battery of variables on stated support for carbon pricing. The latter includes general support for carbon pricing, support for an *increase* in carbon pricing, a carbon price with *rebates*, and support for a carbon price internalizing the *social cost of carbon* at \$120 per ton of CO_2 emissions (with and without rebates). Across these variables, the estimated meta-effects correspond to an increase in support of about two percentage points (see SI, Figure S.2 and Figure S.3). Thus, the effects, despite being small, are non-negligible. Indeed, even modest attitudinal shifts can substantially affect social norms, tipping points, and,

ultimately, policy outcomes, in a polarized context (Otto et al., 2020; IPCC, 2022; Connolly et al., 2025; Granulo et al., 2025). This result contrasts with interventions targeting individual behaviors, where no comparable collective threshold or direct policy leverage exists (see, e.g., DellaVigna and Linos, 2022; Kaiser et al., 2025).

A second key set of findings concerns the quantification of between-study heterogeneity that arises solely from differences in experimental designs. We observed low-to-medium levels of between-study design heterogeneity (I^2 and τ), which are smaller than documented in other many-designs projects (see Holzmeister et al., 2024). This advances the field in providing a better understanding of interventions that appear to work best in increasing carbon tax support. Specifically, most effects are statistically insignificant (52 out of 55 at the 0.5% level; and 44 out of 55 at the 5% level), mirroring the much lower share of significant results observed in registered reports compared to standard studies (Scheel et al., 2021). Nine studies showed significant or suggestive positive evidence, and two studies showed significant or suggestive negative evidence. Peer- and AI-based classification of intervention categories reveal suggestive evidence that interventions drawing on bias correction or reflective thinking tend to yield positive effects. Although behavioral interventions show positive effects on support overall, policymakers should exercise caution, as backfiring effects cannot be ruled out: prediction intervals for the primary outcomes clearly extend into the negative domain.

When comparing the observed effect sizes with researchers' expectations for their own study and those of their peers', we found clear evidence of overconfidence, as documented for other domains different from climate action (Botvinik-Nezer et al., 2020; Menkveld et al., 2023; König-Kersting et al., 2025). RTs anticipated significantly larger positive treatment effects than those actually observed. A possible explanation for this discrepancy could be researchers' distorted beliefs due to widespread publication bias. As a reference, the pooled meta-effect restricting only to individually statistically significant studies is Cohen's d = 0.25 (see SI, Columns (2) of Table S.10 and Table S.12 for stated support), approximately four times larger than the main estimate for this study. This much larger estimate is what a reader might encounter across the average of published individual studies, subject to publication bias. This further supports the need to moderate the scientific community's expectations of the potential of behavioral interventions.

Despite its strengths, this "many-designs" project also has limitations. First, the requirement for RTs to develop a real-world outcome measure may have prompted some to artificially link experimental behavior (e.g., behavior in economic games) to real-world outcomes (such as donations), as proxies for support. This may have inadvertently produced mechanical drivers on observed outcomes that do not necessarily reflect genuine shifts in underlying preferences. More generally, we found a low and non-significant correlation between real-life and stated support outcome measures. This might highlight a broader challenge for research relying on proxy outcomes. Conservatively

excluding seven studies that may suffer from this issue or are not designed to distinguish genuine preference shifts from mechanisms such as decoy or default effects, an exploratory robustness check reveals an attenuated but detectable meta-effect (Cohen's $d = 0.049$, $k = 49$, $p = 0.012$; see SI, Columns (3) of Table S.10 and Table S.12 for stated support), consistent with reduced precision under the more restrictive sample. Second, results of individual studies need to be treated with caution, as, for each single study, we were underpowered to detect small effects with high probability (i.e., Cohen's $d < 0.2$ with power $\geq 80\%$). As such, ceteris paribus, insignificant results reported for single studies might become significant when increasing the sample size. In any case, caution is always advisable against drawing policy implications from single lower-powered studies. Nevertheless, the main conclusions of our meta-analyses remain reliable, given the high power for very small effects based on a sample size of over 19,000 participants. Third, we randomly selected "only" 55 RTs from over 135 applicant teams, all of whom met our inclusion criteria and represented the broader research community. We considered this random selection approach the most appropriate way to avoid systematic selection effects and to meta-scientifically capture the creativity of the field in an unbiased manner. Finally, our study focuses on a single country, the US, and further studies could be planned to investigate whether the presented results generalize to other cultures and societies.

Methods

Experimental Design and Methods

We pre-registered an extensive pre-analysis plan on OSF (https://doi.org/10.17605/OSF.IO/35KRQ) outlining the entire structure, procedure and the main part of the data analysis. Furthermore, the pre-analysis plan included protocols for communication with the teams and the participation requirements.

We then began an open call for collaboration (scientific email lists such as the Economic Science Association, Society of Experimental Finance, American and European Associations of Environmental and Resource Economists, and Judgment and Decision Making, as well as social media posts) while simultaneously launching the project website (https://manydesignscarbon.online) which was the main communication channel. The general research question that was presented on the project webpage and clearly communicated to the RTs was, "What is the impact of behavioral interventions on real-world support for a price of carbon?". The spirit of the project was to tackle the introduction of a carbon price or the substantial increases in the price of carbon, aiming to advance toward carbon pricing incorporating external damages from CO_2 emissions. This is referred to as the social price, and we refer to the lower-bound estimate by the U.S. Environmental Protection Agency (2023) of \$120 per ton of CO_2. RTs were not required to tailor their interventions specifically to this price, but it was mentioned as a reference value, explicitly included on the project website.

Overall, 135 international RTs indicated interest in the project, all of which complied with the prerequisites (maximum of two members, at least one with a PhD, and previous experience with experimental methods—at least one team member needed one experimental study in the social sciences that was either published or a working paper). We originally chose 42 random RTs with a pre-registered STATA script to be fully funded by the project coordinators, based on the ex-ante power analysis. In addition, due to the large number of applications by RTs (i.e., 135 was much higher than our initial expectations) and the offer of some RTs to fund their design themselves, in addition to the already randomly selected RTs, we pre-registered an Addendum A to the pre-analysis plan, specifying the random selection of 25 additional RTs using the same pre-registered STATA. In the few cases where these additional RTs could not self-fund their study, the core team covered the funding, and the availability of fund played no role in being included in the study, thus, avoiding selection bias. Next to funding constraints, the upper bound of 67 invited RTs was defined by the technical feasibility on Prolific to simultaneously run the different treatments from all designs. In the end, 55 of the selected 67 RTs completed the full study protocol on time. Researchers from RTs whose designs are included in the project are full co-authors of this paper. RTs cannot be connected to the individual studies mentioned in the paper.

The RTs had several tasks to complete. First, they had to submit their design through the website. RTs described the intervention, control condition, suggested real-world outcome measure, and details related to their study, such as expected duration and the incentives (see the pre-registration template in the pre-analysis on the OSF repository). Additionally, RTs were required to complete a self-assessment survey, in which we asked them about the effect size that they expected from their proposals, in Cohens' d units. The project coordinators carefully assessed all proposed designs to evaluate their fit to the project guidelines (i.e., IRB criteria for fast-track approval, use of economic incentives as compensation, design qualifying as behavioral intervention and the suggested outcome having a real-world carbon pricing support). We notified the RTs that satisfied all criteria and asked them to start preparing the software (either in oTree or Qualtrics). In case proposals deviated from the project guidelines in any dimension, we contacted the RTs requesting revisions before developing their software. In all cases, we requested RTs to add a pre-programmed template of the "policy survey" by the project coordinators (either an *.otreezip or *.qsf file) at the end of their software (see SI, Section 5, for the full policy survey). In the meantime, RTs were also randomly assigned 10 other designs to assess: (a) the informativeness for answering the research question (Likert scale from 0 to 10), (b) the category that best described, in their view, the intervention, and (c) the expected effect sizes as measured in Cohen's d units. For the main data collection, we used Prolific with the following sample screening criteria: U.S. based, English fluency, and a 95% approval rate on the platform. We used an oTree-based balancer software (building on Huber et al., 2023) to randomly assign the participants to one of the 110 experimental conditions (see the OSF repository for the pre-registered balancer software code). Every participant from Prolific was directed to a common landing page with a general introduction and informed consent form, as well as a captcha task to prevent bots. Then, batches of four people were formed to allow for group experiments of up to four people.

References

Anderson, S., Marinescu, I., & Shor, B. (2023). Can pigou at the polls stop us melting the poles? *Journal of the Association of Environmental and Resource Economists*, *10*(4), 903–945.

Andre, P., Boneva, T., Chopra, F., & Falk, A. (2024). Globally representative evidence on the actual and perceived support for climate action. *Nature Climate Change*, *14*(3), 253–259.

Baldwin, M., & Lammers, J. (2016). Past-focused environmental comparisons promote proenvironmental outcomes for conservatives. *Proceedings of the National Academy of Sciences*, *113*(52), 14953–14957. https://doi.org/https://doi.org/10.1073/pnas.1610834113

Barrage, L., & Nordhaus, W. (2024). Policies, projections, and the social cost of carbon: Results from the dice-2023 model. *Proceedings of the National Academy of Sciences*, *121*(13), e2312030121. https://doi.org/10.1073/pnas.2312030121

Bayer, P., & Aklin, M. (2020). The european union emissions trading system reduced co₂ emissions despite low prices. *Proceedings of the National Academy of Sciences*, *117*(16), 8804–8812. https://doi.org/10.1073/pnas.1918128117

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., et al., (2018). Redefine statistical significance. *Nature human behaviour*, *2*(1), 6–10.

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Carson, R. T., Lu, J., Khossravi, E. A., Köhlin, G., Sterner, E., Sterner, T., & Whittington, D. (2025). The public's views on climate policies in seven large global south countries. *Nature Climate Change*, *15*(9), 954–962. https://doi.org/10.1038/s41558-025-02389-9

Connolly, D. J., Loewenstein, G., & Chater, N. (2025). An s-frame agenda for behavioral public policy research. *Behavioural Public Policy*, *9*(3), 593–613. https://doi.org/10.1017/bpp.2024.58

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Sanchez Chico, A., & Stantcheva, S. (2025). Fighting climate change: International attitudes toward climate policies. *American Economic Review*, *115*(4), 1258–1300. https://doi.org/10.1257/aer.20230501

DellaVigna, S., & Linos, E. (2022). Rcts to scale: Comprehensive evidence from two nudge units. *Econometrica*, *90*(1), 81–116. https://doi.org/https://doi.org/10.3982/ECTA18709

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *6*(102), 460–465. https://doi.org/https://doi.org/10.1511/2014.111.460

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, *27*(4), 857–871. https://doi.org/10.2307/2528823

Granulo, A., Fuchs, C., & Böhm, R. (2025). Psychological reactance to system-level policies before and after their implementation. *Proceedings of the National Academy of Sciences*, *122*(18), e2409907122. https://doi.org/10.1073/pnas.2409907122

Hájek, M., Zimmermannová, J., Helman, K., & Rozenský, L. (2019). Analysis of carbon tax efficiency in energy industries of selected eu countries. *Energy Policy*, *134*, 110955. https://doi.org/https://doi.org/10.1016/j.enpol.2019.110955