Organizational system thinking as a cognitive framework to meet climate targets

Matteo Burato[a], Samuel Tang[b], Vincenzo Vastola[c], and Simone Cenci[a,1]

Edited by Stephen Polasky, University of Minnesota Twin Cities, St. Paul, MN; received June 6, 2023; accepted September 1, 2023

System thinking is a crucial cognitive framework to enable individual pro-environmental behavioral changes. Indeed, a large body of literature has shown a significant and positive association between individuals' system thinking capacities and perceptions of the threat posed by climate change. However, individual behavioral changes play a limited role in addressing climate change compared to large organizations involved in a significantly larger share of economic activities. Do organizations exhibit system thinking capacities? Here, we conjecture that system thinking is a cognitive framework observable at an aggregated group level and, therefore, organizations, not just individuals, can exhibit characteristic levels of system thinking. We conceptualize a definition of organizational system thinking and develop an empirical method to estimate it using a large body of textual data from business organizations. Then, we show that system thinking organizations are more likely to lower emissions and align them with the pathways required to meet the climate targets set by the Paris Agreement. Finally, we discussed the theoretical and policy implication of our study. Overall, our results suggest that system thinking is a relevant organization-level cognitive framework that can help organizations align their emissions with global climate targets.

system thinking | climate change | organizational behavior

Lowering greenhouse gas (GHG) emissions to a level compatible with the climate targets set by the Paris Agreement requires significant changes in behavior and attitudes toward environmental issues by both individuals and organizations (1, 2). A large body of research has shown that individuals' capacity to understand the effect of climatic changes and to change behaviors to address their root causes requires the development of specific cognitive abilities (e.g., logical reasoning, problem-solving, memory, information processing) (3–6). Among these cognitive abilities, system thinking has been shown to play a particularly relevant role in facilitating meaningful change toward sustainable pro-environmental behavior (7–12).

System thinking refers to a "cognitive paradigm that involves an implicit tendency to recognize various phenomena as a set of interconnected components that interact with one another to make a dynamic whole" (13). It is the capacity to explore and develop actions in complex contexts, enabling systems change. System thinkers recognize that their behavior is embedded in complex socioeconomic systems (12) and that natural and social phenomena result from constant dynamic and multiple interactions between the social, economic, and natural worlds as opposed to a sum of siloed processes (11, 13). That is, systems thinkers view the world as a set of dynamic and interconnected parts and processes.

The importance of system thinking in tackling wicked problems is becoming increasingly apparent (11). The US National Research Council (14) and the Next Generation Science Standards (15), for example, place systems thinking and integrated multidisciplinary science at the forefront of their agenda. Similarly, the UK government has put forward official guidance for civil servants to include system thinking in their toolkit to drive improved outcomes in complex situations. In the context of climate change, the growing emphasis placed on system thinking approaches is due to emerging theories and empirical evidence that illustrate a significant and positive association between individuals' capacity of system thinking and pro-environmental behaviors (7–11). For example, ref. 11 has shown that systems thinking is positively associated with an ecological worldview as defined by the New Ecological Paradigm of ref. 16, i.e., the belief that people should take care of the environment rather than exploit it. Similarly, ref. 10 has found that system thinkers ascribe more monetary and socioecological value to the natural world than individuals who score lower in system thinking assessment tests. Overall, several studies have shown that system thinking is associated with a greater perception of the threat posed by climate

Significance

System thinking is a cognitive framework associated with individuals' proenvironmental behavior and with their abilities to understand the threat posed by changing climates. In this study, we argue that large organizations, just like individuals, can exhibit system thinking capacities. We develop a theoretical framework for organizational system thinking and an empirical approach to estimate it from observational data in the context of business organizations. Then, we show that organizations with high levels of system thinking are more likely to lower greenhouse gas emissions and align them with the climate target set by the Paris Agreement. Our findings suggest that organizational system thinking is an essential cognitive framework for organizations to address climate change.

Author affiliations: [a]Leonardo Centre on Business for Society, Imperial College Business School, London SW7 2BX, UK; [b]Institute for Sustainable Resources, Bartlett School of Environment, Energy and Resources, University College London, London WC1H 0NN, UK; and [c]Department of Management, Strategy, and Entrepreneurship, Montpellier Business School, Montpellier 34080, France

Author contributions: M.B., S.T., V.V., and S.C. designed research; M.B. and S.C. performed research; S.C. supervised the research; M.B. and S.C. analyzed data; and M.B., S.T., V.V., and S.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: s.cenci@imperial.ac.uk.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2309510120/-/DCSupplemental.

Published October 2, 2023.

change and support for environmental policies to mitigate its effects.

Individual system thinking capacities can drive pro-environmental behavioral changes that are crucial to address the climate crisis. However, individuals' behavioral changes play a limited role in addressing climate change compared to the potential effect of behavioral changes in large organizations, such as business organizations and, in particular, publicly traded companies (hereafter referred to as companies). Indeed, a recent study from the International Energy Agency found that individuals, through their voluntary reduction in demand for consumption and other behavioral changes, account for approximately 11% of the emissions reductions required to reach global net zero by 2050 (17). On the other hand, organizational changes in companies are crucial for tackling climate change due to the share size of the economic activities companies are involved with (18) and the central role they play within modern intertwined societies. Indeed, they are not only responsible for a great portion of human polluting activities directly, but they also influence the worldview and further behavior of the collective of their stakeholders. Yet, despite increasing societal, financial, and regulatory pressure, there is little evidence of change at the scale or pace required for us to avoid catastrophe (18, 19). We lack clear mechanisms or guidance on how to change organizational behavior in general and companies' actions in particular.

To that end, we theorize that system thinking can be an effective cognitive framework to realign the behavior of organizations to meaningfully tackle climate change. We focus on for-profit business organizations, specifically on publicly traded companies; we investigate whether they can exhibit system thinking capacities and whether the presence of organizational system thinking capacities is associated with lower environmental impacts. Specifically, we build on the existing system thinking literature that studies individuals' cognitive abilities and their relation with pro-environmental behavior (7–11) as well as the literature that studies the emergence of group cognitive abilities (20–22). Building on these works, we conjecture that system thinking is a cognitive framework observable at an aggregated group level. Therefore, organizations, not just individuals, can exhibit characteristic levels of system thinking. Importantly, we propose and test the hypothesis that, just as individual system thinkers are more likely to exhibit pro-environmental behavior, system thinking organizations (companies) are able to achieve superior environmental outcomes (which can be seen as manifestations of pro-environmental behaviors). Importantly, here, we focus on GHG emissions to measure environmental outcomes due to data limitations in measuring the impact of business operations across other environmental dimensions, e.g., soil health, biodiversity (*Discussion*).

We start by developing a general definition of organizational system thinking as the capacity of organizations to recognize that their operations and multiple (often conflicting) goals affect and are affected by numerous societal actors and environmental factors (*Organizational System Thinking*). This conceptualization forms the basis for our empirical estimation of organizational system thinking, in the context of publicly traded companies, from observational data using human-in-the-loop natural language processes approaches* to analyze a large quantity of text from companies' disclosure of sustainable carbon management practices (*Materials and Methods*).

Using our empirical estimation of organizational system thinking, we explore its relationship with companies' GHG emissions. In particular, we analyze data from 615 large publicly traded companies distributed across 32 countries in the Energy, Industrial, Material, and Utilities sectors over the observation period 2012–2020. We hypothesize that system thinking organizations (companies) tend to have lower GHG emissions, compared to companies with similar asset characteristics. We expect that the level of organizational system thinking is associated with lower emissions due to the well-established relationship between climate change beliefs and actions and individual level of system thinking (11), i.e., the distinguishing features of individual-level capacities need to be preserved at the organization level. Moreover, we also explore the relationship between organizational system thinking and the capacity of companies to align their long-term projected emissions pathways with the required pathway to limit global warming well below 2 °C.

Overall, in this manuscript, we conjecture that organizations, like individuals, can exhibit system thinking skills, and we hypothesize that organizational system thinking provides an essential cognitive framework to address climate challenges. In the next section, we provide a detailed conceptualization of organizational system thinking. Then, we present our empirical estimation approach and analyze the relationship between our estimations and companies' emissions. Finally, we discuss our findings and their business and policy implications.

Organizational System Thinking

System thinking has been traditionally studied as an individual-level trait that, although correlated with others, is independent and identifiable (9, 13, 24). It is best described as a cognitive paradigm that allows those who apply it to recognize and emphasize the interconnections between phenomena and how those interconnections affect the overall dynamic of a system (11, 25). Previous studies have shown that system thinkers can better engage in complex decision-making problems, encompass different perspectives, and understand resource accumulation dynamics (11, 26–29). Thus, as argued in several studies, a system thinking approach is paramount for effectively addressing "wicked" challenges, including climate change (30, 31). Due to their very nature, however, wicked challenges are beyond the reach of any individual system thinker and require extensive collective effort from organizations (e.g., businesses, NGOs). Hence, here, we argue that to address wicked challenges, and climate change in particular, organizations need to operate within this cognitive paradigm; i.e., to tackle climate change, organizations must be system thinkers.

To do so, we extend the notion of system thinking to a paradigm of the *collective cognition* of organizations. Cognition in organizations has been a crucial object of interest across many disciplines and multiple decades (32, 33). Extant research explored organizational cognition through the lenses of "shared causal maps," i.e., negotiated and symbolic representation of reality as sets of phenomena and relationships among phenomena (34–36). This literature generally describes organizational cognition as deeply intertwined with the functioning of the organization as a system.

In particular, we draw upon extensive research on cognition at the interindividual level, which identifies both the existence of shared understanding and specialized cognitive loads across individuals in organizations (37–41). From this perspective, organizational cognition does not overlap entirely with any individual cognition as it emerges from the collective process

*Human-in-the-loop are a series of machine learning approaches that leverage human knowledge in the training process to increase the accuracy of prediction and classification algorithms (23).

of information acquisition and processing. In other words, organizational cognition is an emergent property of the organization that arises from the interaction of its members (42).

This description of organizational cognition draws strong parallels with classic definitions of system thinking, which emphasize the recognition of relational complexity among phenomena. Following these considerations, we define organizational system thinking (O-ST) as a collective cognitive paradigm of organizations that recognizes that organizational processes and multiple (often conflicting) goals affect and are affected by multiple societal actors and environmental factors.

In particular, we expect that O-ST produces tangible effects in terms of organizational outcomes. Indeed, organizational modes of cognition have long been linked with the development of capabilities that organizations can employ to effectively interact with their surroundings (43–45), and the paradigm that underpins organizational cognition plays a pivotal role in shaping organizational capabilities and outcomes.

In the following section, we test this expectation by studying the relationship between characteristics level of system thinking in publicly traded companies and their capacity to lower their environmental impact to a level compatible with societal expectations. First, however, we introduce an operational framework to estimate these characteristic levels from observational data.

Overview of the Empirical Estimation Process. We now focus on for-profit organizations, particularly publicly traded companies, due to the large impact of their operations on long-term climate dynamics (18, 19). Following our conceptualization of O-ST, organizations (companies) with high levels of system thinking are those which recognize the system dynamics among 1) behavioral and decision-making processes, 2) their multiple (environmental, social, financial), often conflicting goals, and 3) multiple environmental factors and actors. In our setting, we focus on companies' efforts to tackle climate change. We investigate O-ST by looking into companies' disclosure of processes and policies across multiple functional domains and searching for cues that suggest alignment of organizational cognition to the system thinking paradigm.

Specifically, we estimate O-ST by analyzing companies' disclosure of sustainable carbon management processes to the Carbon Disclosure Project (CDP). CDP is a leading nonprofit international organization that systematically collects information on organizations' carbon management processes and outcomes and whose database is widely used by studies in this domain[†] (46, 47). Information is collected through surveys organized in closed-form and open-ended questions. Examples of questions include emissions targets, total GHG emissions, supply chain policies, product development, and responsibilities within the organization for managing and reporting emissions. Due to the extensive information required, answering the CDP survey implies collective inputs and significant interactions among different members of the organization, bridging diverse functional units (e.g., production, supply chain, marketing, top management, etc.). Therefore, the questionnaire provides a window into the core processes of the organizations, allowing the observation of O-ST as emergent from group interactions.

Our estimation approach is described in detail in empirical estimation of organizational system thinking. We start with using our definition of O-ST (*Organizational System Thinking*) to create a template of characteristics that we expect to see in the CDP survey responses if those answers subsume a system thinking approach (*SI Appendix*, S1). Then, using the template and focusing on publicly traded companies, we manually classified approximately 2000 answers to the CDP questionnaires from 2012 to 2020 based on whether or not those answers meet the expected characteristics. Specifically, we assign a label of "one" to each answer that meets one or more of our expected characteristics and a label "zero" otherwise. *SI Appendix*, S2 in Supplementary information reports some examples and commentary on how this process works in practice. Then, we trained a transformer language model (BERT) on this manually annotated dataset. The model is then used to predict the class of the rest of the answers in the CDP dataset. Finally, we computed an average system thinking score for each company-year observation by averaging over the predicted classes of the answers of a given company in a given year. In the next section, we use this score as our measure of O-ST.

In order to appropriately compare the level of O-ST across entities in our sample, it is important to focus on companies with comparable business needs. Therefore, in the following analysis, we focus on publicly traded companies in the Energy, Industrial, Material, and Utilities sectors.[‡] The business needs of companies within these sectors are comparable in that production and revenues strongly depend on tangible assets and supply of fossil fuels, and carbon management processes are particularly relevant for continuing profitable business operations. Moreover, we expect that O-ST is integral for these industries due to their complex supply chains and their exposure to environmental risks, which include both physical and transition risks.

Results

First, we provide an overview of our population, the cross-sectional and temporal evolution of O-ST, and its distinguished features. Then, we test the hypothesis that O-ST is positively associated with key measures of environmental outcome, i.e., GHG emissions and alignment with climate targets. The summary statistics of all variables used in the analysis are shown in *SI Appendix*, Table S2.

Empirical Characterization of the Sample and Organizational System Thinking. Fig. 1 shows three important sets of characteristics of our population: corporate social responsibility (CSR) policies (Panel *A*), alignment with the target set by the Paris Agreement to limit global warming below 2 °C (Panel *B*), and GHG emissions (Panel *C*).

Established routines and structures, such as corporate policies, are crucial counterparts to shared cognitive frameworks such as O-ST. Fig. 1*A* summarizes the relative frequency of various CSR policies (*Data*) that might influence the extent and efficiency of companies' actions in achieving environmental outcomes. Generally, a significant proportion of our population had CSR policies in place during the sample period (2012–2020). Specifically, a majority of CDP respondents have set internal targets and been involved with stakeholder engagement processes, including establishing environmental partnerships. Most companies also have environmental material sourcing policies and have developed sustainable products. Of note, 50% to 60% of companies have board-level policies, including executive compensation and Equity, Diversity, and Inclusion (EDI) policies. However, broader EDI policies, such as the

[†] Importantly, CDP collects information from several forms of organizations, including companies, investors, and public authorities.

[‡] We use the Global Industry Classification Standard (GICS) to filter companies in these sectors.

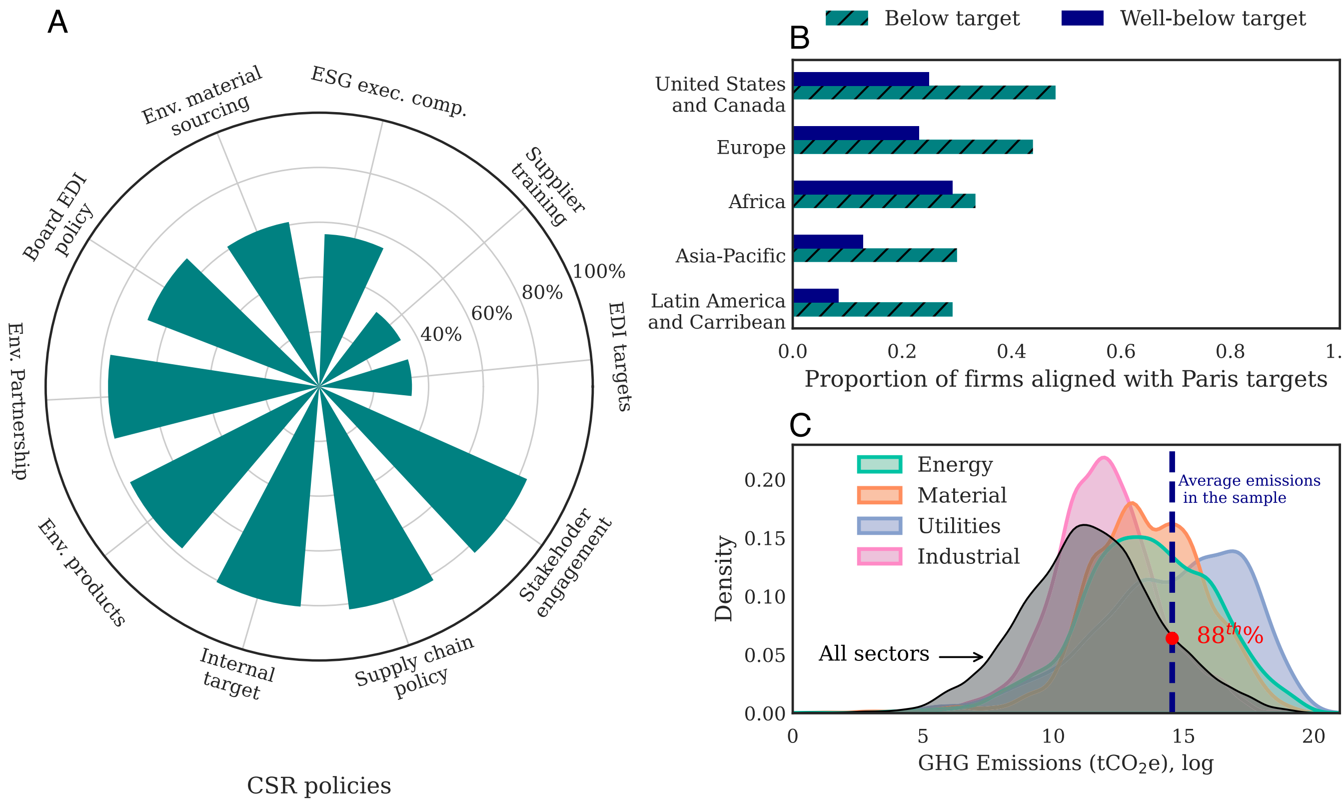

Fig. 1. Sample characteristics. Panel (*A*) shows the frequency of CSR policies in our sample. Panel (*B*) shows the proportion of companies with emissions pathways aligned with climate targets as of 2020. Panel (*C*) shows the average GHG emissions (log) of companies in our sample (blue vertical line), the emissions of all companies in the hard-to-abate and energy-intensive sectors (pastel distributions) and of all companies in the Trucost universe (~15,000 companies across all sectors, dark gray distribution). The red dot shows the 88th percentile of emissions in the whole universe distribution, corresponding to the average emissions of companies in our sample. Overall, the figure shows that, while companies in our sample adopt several environmentally related CSR policies, they fall short of meeting climate targets and are among the highest global emitters.

establishment of EDI targets, are rare. In terms of supplier relationships, we observe a broad diffusion of supply chain management practices (e.g., code of conduct, reviews), while active supplier training policies are significantly less common.

Despite the relatively high frequency of environmental CSR policies, companies in our sample are not on track to effectively contain their projected emissions within the boundaries necessary to address climate change. Global standards in this sense are set by the Paris Agreement, a binding international treaty signed by 196 countries in 2015, which puts forward as a key target the containment of increased average global temperatures to below 2 °C compared to preindustrial levels (48). Panel *B* of Fig. 1 shows the proportion of sample companies whose emission pathways, as of 2020, were aligned with Paris targets of limiting global warming below (teal) and well below (navy) 2 °C (see *Data* for details on the calculation of emission pathways). The panel shows that only 30% to approximately 50% of companies have emissions pathways aligned with the below 2 °C, and an even smaller fraction of companies (10% to 30%) are aligned with the well below 2 °C. Importantly, alignment with climate goals varies significantly across geographical regions. Notice that while our sample only comprises 615 companies, Panel **C** shows that their average emissions are on the top 88th percentile of the distribution of emissions of the largest 15,000 publicly traded companies (dark gray distribution in Panel *C*). Therefore, companies in our sample are global leading polluters, and they, together with comparable companies, play a crucial role in the achievement of the Paris targets. Further insights on the causes of their relatively low environmental performance and the apparent disconnect between the adoption of industry best practices and overall target alignment are therefore essential to bring back the private sector on track to meet global climate targets.

We now focus on the characterization of our derived measure of O-ST across the sample and the observation period (see Overview of the empirical estimation process and Empirical estimation of organizational system thinking). The black dotted line in Fig, 2*A* shows the temporal evolution of the average level of O-ST. We have found a substantial positive temporal trend, with a clear acceleration in recent years, starting from 2018. Indeed, the relative incidence of O-ST answers was approximately 10% in the early 2010s and reached approximately 20% in 2020. In Fig. 2*B*, we zoom in on the cross-sectional distribution of O-ST by year. Specifically, we split the yearly aggregate into quartiles and represent the distributions within each quartile in individual boxplots. The top quartile (blue) shows a greater dispersion, suggesting a significant presence of a few advanced companies that significantly outperform the pack. However, we observe a positive trend across all quartiles, which suggests a broader O-ST diffusion over time. Indeed, the bottom quartile (red) progressively detaches from the lower bound at 0. Together, the two panels show that system thinking is becoming more prevalent in our sample and that this trend encompasses the entirety of the companies we analyzed.

Fig. 2*C* shows the relative prevalence of different themes in answers that contain system thinking cues (*SI Appendix*, section S5). Specifically, we report the ten topics which showed

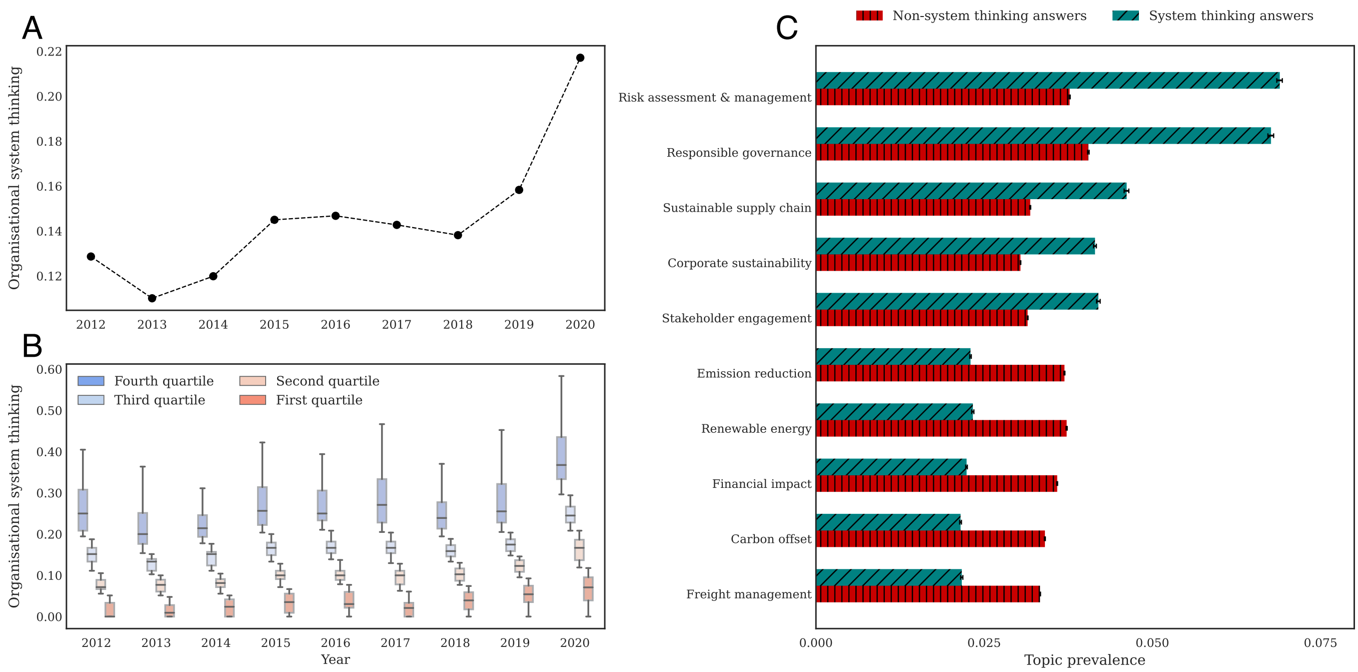

Fig. 2. Evolution and characterization of organizational system thinking. Panel (*A*) shows the temporal evolution of O-ST within our sample. Panel (*B*) shows the yearly distribution of O-ST by quartile. Panel (*C*) shows the relative prevalence of topics in the answers to the CDP questionnaire classified as system thinking (green) and non-system thinking (red). Here, we show the prevalence comparison only for the ten topics with the highest positive (first five) and negative (last five) differences between the two groups. In *SI Appendix*, Table S9, we show the full results from the topic analysis.

the starkest (top five positive and top five negative) difference between system thinking and non-system thinking answers. System thinking answers have a broader scope than the traditional objectives of emission reduction and are related to a wider interpretation of the company sustainability mandate (e.g., in areas like governance and supply chain management). They are also more concerned with information and assessment procedures, which suggest a greater understanding of the need to probe a complex, systemic environment. Conversely, non-system thinking answers were linked to themes more in line with the core of CDPs questions, with concrete and direct answers in terms of emissions, energy, and financial costs. Importantly, in *SI Appendix*, section S4 and Table S1, we show that, additionally to discuss topics with a broader scope, system thinking answers are also associated with greater text complexity.[§] Therefore, system thinking answers are not only different in content but also in structure. This suggests that the mechanisms and cognitive processes underlying the formulation of those answers, such as understanding the complexities involved in tackling climate change, are fundamentally different.

Organizational System Thinking, GHG Emissions, and Alignment with Climate Targets. We now estimate the relationship between O-ST and the environmental impact of business operations. Specifically, we focus on two impact measures (outcomes): GHG emissions and their alignment with climate targets. Emission data are from Trucost (*Data*) (49), and here, we focus exclusively on all the emissions that are under the control of the company and can therefore be directly related to management choices and practices. This includes GHG protocol scope 1 emissions, plus any other emissions derived from a wider range of GHGs relevant to a company's operations, plus GHG protocol scope 2 emissions, plus the company's first-tier upstream supply chain. In the analysis, GHG emissions are measured cumulatively two years ahead. Data on alignment with climate targets are also from Trucost (50). Alignment with climate targets is defined as the difference between the projected emission pathway of an organization and its required pathway to limit global warming below and well below 2 °C (*Data*). We use these two outcome measures as a proxy for short-term (GHG emissions) and long-term (alignment) impact. In section Discussion, we discuss the applicability of our framework to other environmental impact measures.

The estimation approach is divided into three steps (further details, including strategies to address endogeneity issues, are provided in *Empirical Specifications*). First, we identify factors associated with characteristic levels of O-ST. Specifically, we run a linear model with O-ST as the dependent variable and a series of asset characteristics as independent variables. Importantly, we focus on asset characteristics, such as Size and proportion of tangible assets, that are often studied in relation to GHG emissions (51). In conjunction with the asset characteristics, we also control for fixed effects, self-selectivity, and a series of complexity measures of the CDP answers for every observation in the panel (see *Empirical Specifications* for further details on the model). The estimated regression coefficients of the asset characteristics are shown in Fig. 3*A*. The error bars denote bootstrapped 95% CIs. We have found that companies with high levels of system thinking tend to be small, have a large presence in the market (measured as total invested capital), have a high value of growth opportunities (market to book), and have a low proportion of tangible assets in their books.

Second, we estimate the relationship between O-ST and several CSR policies. Specifically, we run several Probit models, each explaining the presence or absence of a CSR policy. In the models, we control for O-ST, asset characteristics, fixed effects, self-selectivity, and text complexity measures. The CSR policies

[§]Text complexity is an important factor to account for in the analysis presented in the following section to account for endogeneity issues, as discussed in *Empirical Specifications*.

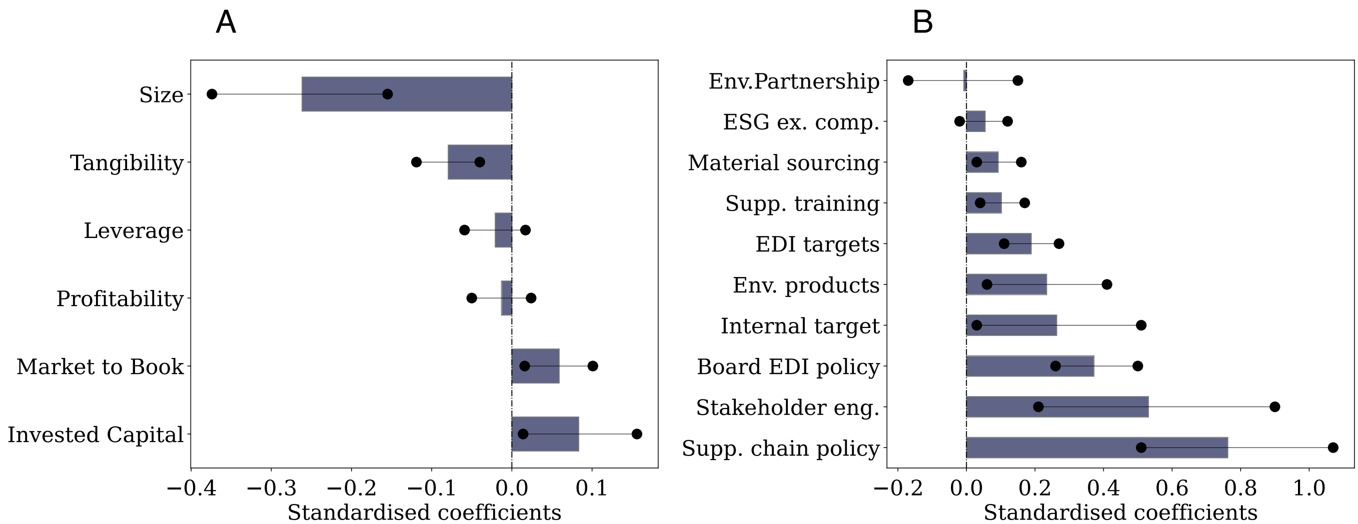

Fig. 3. Organizational system thinking, asset characteristics, and CSR policies. Panel (*A*) shows the association of O-ST with a series of asset characteristics. Specifically, the *y*-axis shows the regression coefficients of the covariates included in a linear model that explains O-ST. Panel (*B*) shows the regression coefficients of O-ST in a series of Probit models, each explaining the presence or absence of CSR policies (*y*-axis). See *Empirical Specifications* for further details on the models. The error bars in both panels show the bootstrapped 95% CIs.

can be seen as the means by which O-ST acts on emissions. Fig. 3*B* shows the estimated regression coefficients of O-ST in each of the Probit models and their bootstrapped 95% CIs. We have found that companies with high levels of system thinking tend to employ several CSR policies, and these policies span both the environmental and social domains. Indeed, in addition to being more likely to adopt sustainable management practices (e.g., developing environmental products and sustainable supply chain policies), companies with a high level of system thinking are also more likely to adopt EDI policies and engage with a broad spectrum of stakeholders. Notably, the probability of setting internal emissions target is positively associated with the characteristic levels of O-ST.

Finally, we estimate the association between O-ST, future GHG emissions, and alignment with climate targets. Fig. 4*A* shows the results of our main estimations. The top bars in the figure show the regression coefficient of O-ST in a linear model that explains future cumulative GHG emissions. Similarly to the previous specifications, in this and each of the subsequent models, we control for asset characteristics, fixed effects, self-selectivity, and text complexity measures. The coefficient is negative and statistically significant, i.e., the higher the level of O-ST, the lower the cumulative future emissions. The coefficient remains negative and statistically significant regardless of whether or not we control for the presence of CSR policies (dotted bars). That is, individual CSR policies do not mediate the effect of O-ST on short-term emissions, suggesting that the mediating factor must be a more complex combination of strategic choices.

The middle bars in the figure show the regression coefficients of O-ST in a Probit model that explains the probability of observing emission pathways aligned with the target of the Paris Agreement of limiting global warming below 2 °C. The coefficient is positive and statistically significant. Similarly to the previous estimation, the coefficient does not change after accounting for the presence of CSR policies (dotted bars). The bottom two bars show the regression coefficient of O-ST in a Probit model that explains the probability of observing alignment with the more stringent target of limiting global warming *well* below 2 °C. The coefficient is again positive and statistically significant with and without CSR policies in the model. Overall, O-ST is associated with lower future emissions and with a greater likelihood of observing alignment with climate targets. The results shown in the panel are robust to different estimation strategies of our O-ST measure (see *Robustness Tests* and *SI Appendix*, Fig. S4), to different cutoffs in the data requirements to estimate O-ST (*SI Appendix*, Table S10), and to the inclusion of alternative and simpler measures of O-ST in the control set (*SI Appendix*, Table S11).

Fig. 4*B* shows the regression coefficients of Panel *A* estimated on a rolling window. The *x*-axis in the panel shows the latest year of estimation. The panel shows that the association between O-ST and GHG emissions is consistent across our sample period (*SI Appendix*, Table S5 for the numerical values associated with the figure). The decline in the uncertainty around the point estimates is due to sample size effects since the estimation is performed on a rolling window. The temporal evolution of the association between O-ST and the probability of observing a company with emission pathways aligned with the Paris Agreement target of lowering global warming below and well below 2 °C is noisier. Alignment with the below 2 °C target is consistently significant while alignment with the well below 2 °C target only became significant in 2020, when sample sizes and characteristic levels of O-ST are larger.[¶]

Finally, we would like to note that the results presented in this section are derived from bootstrapped estimations from several random subsamples of our dataset. Their uncertainty is estimated using percentile bootstrap CIs. This estimation strategy allows us to ensure that our findings are not due to some idiosyncrasy of the sample we end up with after merging data from several datasets. Therefore, it guarantees a level of generalizability to our findings. However, while providing an implicit robustness test to our results, this estimation strategy also generates greater uncertainty around the estimations. Therefore, for completeness, in *SI Appendix*, Tables S6 and S7, we provide the full sample results for the analyses in Figs. 3 and 4. The uncertainty is greater for the analysis of the association with the climate targets on the rolling window (Fig. 4*B*) because sample sizes are significantly smaller. Indeed, *SI Appendix*, Table S8 shows that the results for

[¶]The large error bars around the estimates for the target in 2018 are due to the very small sample of companies with available estimated emission pathways. Indeed, in 2018 we have approximately 25% of the observations that were instead available in 2020.

 https://doi.org/10.1073/pnas.2309510120

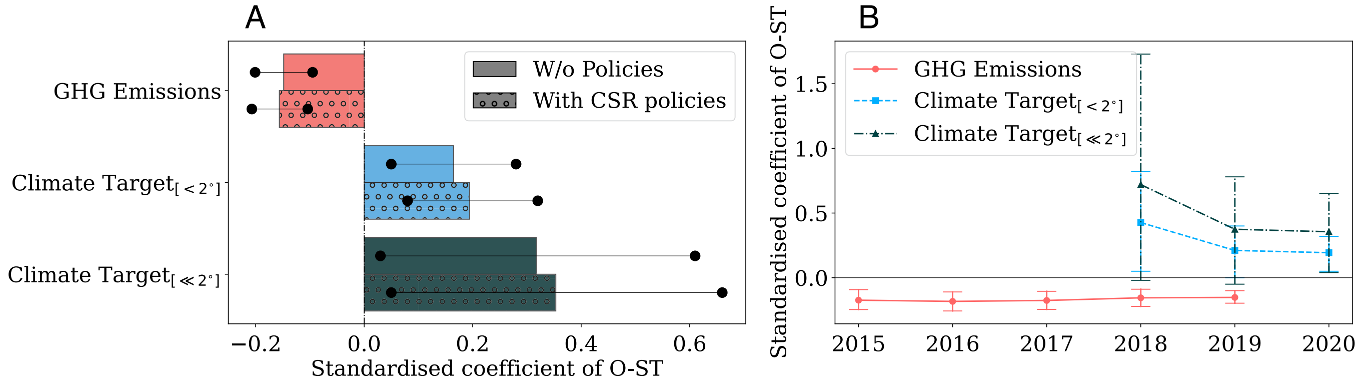

Fig. 4. Organizational system thinking, GHG emissions and alignment with climate targets. Panel (*A*) shows the regression coefficients of O-ST in the models that explain future emissions (red bars) and the probability of alignment with the climate targets (blue and green bars). The coefficients are estimated with (full bars) and without (dotted bars) CSR policies in the control set. Panel (*B*) shows the same regression coefficients estimated on a rolling basis (with CSR policies in the control set). The *x*-axis shows the last year of observations of the independent variables. The error bars in both panels show the bootstrapped 95% CIs. Overall, the results shown in the figure support the hypothesis that organizational system thinking is associated with lower future emissions (red) and with a higher likelihood of alignment with climate targets (blue and green).

the temporal evolution of the association between O-ST and the probability of alignment with climate targets are stronger in the full sample.

Discussion

The ability of individuals to understand the threat posed by climate change and support environmental policies to mitigate its effects is strongly linked to their cognitive capacities and in particular with their system thinking skills (7–11). In this manuscript, we argued that organizations, like individuals, can exhibit system thinking capacities and that organizational system thinking (O-ST) has the potential to provide an essential cognitive framework to address climate challenges. We have provided a theoretical conceptualization of O-ST and an empirical framework to estimate it from observational data in the context of business organizations and, in particular, publicly traded companies.

Main Findings and Limitations. Our main hypothesis was that O-ST is associated with a greater capacity of companies to lower the environmental impact of their business operations. To test this hypothesis, we have analyzed data over the past 10 y from a sample of companies which includes the largest publicly traded polluters in the hard-to-abate and energy-intensive sectors. During the observation period, these companies have taken steps to implement changes in their sustainable management practices as demonstrated by the widespread adoption of environmental and social policies (Fig. 1*A*). Yet they still fall short in meeting global climate targets (Fig. 1*B*). In other words, our analysis suggests that company-level policies are not sufficient to explain both short-term and long-term projections of GHG emissions. This result is in line with previous findings in the oil & gas sector that have found a misalignment between companies' stated goals and actions and their outcomes (19). On the other hand, our O-ST measure is positively and statistically significantly associated with these two key environmental outcomes both cross-sectionally (Fig. 4*C*) and across time (Fig. 4*D*).

The results are robust after controlling for several factors that could provide alternative explanations for these associations (*Materials and Methods*). These factors include 1) self-selectivity, companies deliberately decide to disclose carbon management practices and GHG emissions; 2) complexity measures of the analyzed text, our system thinking measure could have just been a proxy for other cognitive skills that emerge in complex answers to the CDP questionnaire; and 3) company-level CSR policies, system thinking is a property greater than the sum of siloed policies. Moreover, the results are also robust to different empirical estimation strategies (*Robustness Tests* and *SI Appendix*, Fig. S4 and Table S10). Overall, our results suggest that to lower companies' emissions to a level compatible with those expected from science-based targets (such as the targets set by the Paris Agreement), companies need to recognize how their operations affect and are affected by multiple societal actors and environmental factors and that their business needs coevolve dynamically with the environment within which they operate.

There are, however, a number of limitations to our approach. First, we have inferred system thinking capacities from one single source of disclosure, i.e., CDP survey responses. CDP questionnaires are directly related to the environmental outcomes we have analyzed (GHG emissions). Therefore, there could be a potential bias induced by the choice of the underlying data. In other words, what we might be inferring is the level of system thinking observed within a company in its sustainable carbon management practices, not the overall level of system thinking of the company across all business operations. To address this limitation, we have estimated the relationship between O-ST and another environmental impact measure unrelated to emissions (water consumption). We have found that O-ST is associated with lower rates of water consumption. The effect is statistically significant but smaller in magnitude (*SI Appendix*, Table S12). This result suggests that our system thinking measure can capture broader organizational capacities. However, future work could combine several disclosure sources, for example, CSR reports, earning calls, and financial reports to estimate a more comprehensive measure of O-ST.

Second, annotating textual data for inference tasks is notoriously challenging (52). Here, we have employed a classical approach in NLP studies, which is to let multiple individuals annotate the data independently and measure their agreement to provide a robustness measure (*Materials and Methods*). We have also tested the robustness of our results against models trained on data from every single annotator independently and their combinations (*SI Appendix*, Fig. S4). However, future studies could develop better approaches to increase the quality of the annotation task.

Finally, as with any observational study, there could be endogeneity issues with our specifications. To address this concern, we have run multiple tests and used several strategies, such as controlling for self-selectivity, alternative covariates, different subsamples, different time windows, and different measures of our key variable (see *Empirical Specifications* and *Robustness Tests*). However, because our study is purely observational, we still cannot claim that the O-ST capacities cause behavioral choices that are able to lower companies' emissions and align them with climate targets. We can only claim that the two factors are positively correlated. Future research could design experimental studies to further investigate the causal nature of our estimated association.

Implications of Our Work and Future Avenue of Research. Our work has relevant implications for policymakers and business practices. Meeting global climate targets, such as those set by the Paris Agreement, is one of the greatest challenges of our times. Climate targets are set at the country level through the formulation of nationally determined contributions. Delivering on these contributions depends on the actions and behaviors of large organizations, such as publicly listed companies (18). Therefore, understanding which factors drive the necessary internal changes in management practices that can help companies reduce their emissions is crucial to design better incentive schemes, such as targeted environmental policies and market-based solutions that can help countries meet their nationally determined contributions. Our results suggest that policies aimed at fostering system thinking within companies can nudge effective changes in their sustainability behavior.

Importantly, policymakers already appreciate the importance of individual-level system thinking, and several countries, such as the United States and the United Kingdom, are setting standards (15) and guidelines[#] to foster system thinking in educational curricula and government body. Our study suggests that these policies and guidelines should also take into account the way in which system thinking changes as a result of interactions between individuals in an organization, between a group of organizations, and between organizations and decision-making bodies.

The main implication of our results is that system thinking can provide an essential organizational-level cognitive framework to address companies' environmental challenges. There are, however, two additional important contributions of our framework and results beyond their business implications, which we believe can open interesting avenues of research. First, while the existence of emergent capacities from the interactions of individuals within groups is a well-known phenomenon that has been measured in several studies (21, 22, 53); to the best of our knowledge, there is limited understanding about the emergence of organization-level system thinking capacities. Our work provides empirical evidence supporting the hypothesis that system thinking can be an organizational capability. In our analysis, we build on the notion that group interactions in the production of a cognitive output (i.e., CDP survey responses) allow the observation of emergent system thinking capacities. However, our study does not shed light on the processes underlying the emergence of such a collective cognitive framework. Investigating those processes is crucial to investigate further the characteristics of our theoretical construct and its relationship with organizations' outcomes.

Another interesting theoretical result that emerges from our analysis is the connection between system thinking capacities and adverse selection (information asymmetry). Here, we have found that companies with greater system thinking capacities tend to be smaller, with high value of invested capital, intangible assets, and market-to-book ratios (Fig. 3*A*). These characteristics are often associated with greater levels of information asymmetry between insiders and outsiders due to greater uncertainty around the value of intangibles, growth opportunities (market-to-book), and idiosyncratic risk.[‖] One speculative hypothesis that emerges from this result is that lowering the impact of adverse selection forces motivates companies to take on a greater stakeholder-oriented approach, which results in a greater estimated measure of system thinking. Indeed, our operational definition of O-ST is closely related and inspired by stakeholder theories (55). Further research is needed to investigate this relationship in greater depth within the context of novel frameworks in behavioral corporate finance (56).

Finally, we believe that our framework can be further expanded methodologically in two major directions. First, we focused our analysis on uncovering the relationship between O-ST and GHG emission reduction and, in particular, on how company-level reductions are, or are not, in line with climate targets, which are a well-defined but potentially narrow domain. Therefore, a crucial step in further research is to broaden the scope of systemic challenges to consider in evaluating the effects of O-ST. Indeed, business operations have direct consequences on several global environmental challenges like biodiversity loss or soil pollution and social challenges like North–South global inequality or health justice. Second, here, we developed a general definition for organizational system thinking, which can be applied to a broad set of forms of organizations (e.g., businesses, NGOs). However, our empirical approach only focused on publicly traded companies due to their crucial role in tackling climate change. Further research could develop alternative empirical strategies that rely on different forms of disclosure of organizational processes to extend our framework to other forms of organizations.

Overall, in this study, we proposed a theoretical construct and derived an empirical estimation strategy for a cognitive framework that can help organizations address environmental challenges. More broadly, we believe that our work can be the foundation for several future studies investigating the behavioral drivers and implications of organizational system thinking.

Materials and Methods

Data. In this study, we focus on publicly traded companies in the Energy, Utilities, Material, and Industrial sectors as defined by the GICS. Companies in these sectors share similar environmental challenges because production relies significantly on tangible assets, it is energy intensive, and costs strongly depend on commodity prices. Moreover, because carbon management is a material issue for all companies in our sample, the answers to CDP questionnaires are particularly relevant for communicating core business processes and progress to stakeholders in a standardized fashion. Below we provide specific details of the variables we used in our study for reproducibility purposes. In *SI Appendix*, Table S4, we provide the exact name of the datasets and a link to the data providers.

We used annual COMPUSTAT (57) and Refinitiv (58) data for companies' fundamentals. Specifically, we define Size as the log of sales (SALE, in USD) adjusted for inflation (https://fred.stlouisfed.org/series/CPIAUCSL); Invested Capital is long plus short-term debt (DLTT + DLC), plus book equity (CEQ) plus cash and short-term investments (CHE); Tangibility is property plant and equipment (PPENT, in USD) divided by book assets (AT, in USD), Profitability

[#] See for example https://www.gov.uk/government/publications/systems-thinking-for-civil-servants

[‖] Idiosyncratic risk is inversely proportional to companies' size. Indeed, Size is proportional to the number of projects a company can take; hence, it directly drives the volatility of the assets; i.e., larger companies can be seen as well-diversified portfolios of investment projects (54).

is Earnings Before Interests, Tax, Depreciation, and Amortization (EBITDA in USD) over lagged book asset. Market leverage is long-term plus short-term debt (F.DebtTot) divided by market value of assets: total assets (F.TotAssets) − book equity (F.ShHoldEqCom) + market equity (F.MktCap); market to book is the market value of assets divided by Total Book Asset.

Environmental Social and Governance (ESG) data are from Refinitiv Asset4 (59), which is the leading and most comprehensive data provider of ESG and nonfinancial data (see ref. 60 for an extensive review of this dataset). Specifically, we use data on supply chain policy (TR.SupplierEsgTraining, TR.PolicyEnvSupplyChain), environmental partnerships (TR.EnvPartnerships), internal target setting (TR.TargetsEmissions), executive compensation (TR.PolicyExecCompESGPerformance), environmental material sourcing (TR.EnvMaterialsSourcing), environmental products (TR.EnvProducts), EDI targets (TR.TargetsDiversityOpportunity), board diversity policies (TR.PolicyBoardDiversit), and stakeholder engagement (TR.StakeholderEngagement). These variables take on the value of one if a company has a policy in place and zero otherwise. For example, if a company has set an internal emission target, then $TR.TargetsEmissions = 1$.

Company-level emissions data are from Trucost (49), which is one of the most widely used data providers of emission data in the climate finance literature (51, 61). Trucost reports emission data for companies that disclose this information in sustainability reports, annual reports or to CDP, and it estimates the emissions from all the other companies based on proprietary input–output models and an extensive database on production process data (61). Importantly, Trucost covers emissions data for approximately 15,000 publicly traded corporations across sectors and geographical regions, and it has the greater coverage among comparable emissions data providers (61). In this work, we measure total GHG emissions as Direct plus first-tier indirect emissions which are defined as GHG protocol scope 1 emissions, plus any other emissions derived from a wider range of GHGs relevant to a company's operations, plus GHG protocol scope 2 emissions, plus the company's first-tier upstream supply chain. This is Trucost's default measure of emissions. We focus on direct emissions for two reasons: 1) They can be directly related to management practices (and therefore O-ST) and 2) previous works have found that the quality of these emissions is substantially better than the quality of indirect emissions (61).

Data on alignment with climate targets are also from Trucost (50). Specifically, we use the difference between the projected emission pathway of companies as of 2018, 2019, and 2020 and the required pathway to limit global warming below and well below 2 °C. The base year for the alignment calculation is included as a firm fixed effect in the regression. Trucost estimates the transition pathway using the methodologies highlighted by the Science Based Targets Initiative (SBTI). Specifically, they use the Sectoral Decarbonization Approach (SDA) for high-emitting companies with a homogeneous business activity and The Greenhouse Gas Emissions per Unit of Value Added (GEVA) approach for low-emitting companies with heterogeneous business activities.

Finally, information on sustainable carbon management processes and environmental outcome is from the Carbon Disclosure Project (CDP), which is an international nonprofit organization that help organizations disclose their sustainable management practices and environmental impact (62). To merge the four datasets, we create a mapping from ISIN numbers and COMPUSTAT gvkey, since the latter is a unique entity identifier while the former is a security identifier and there can be multiple ISINs for the same entity. Companies that could not be matched with the ISIN number were matched by company name after removing common suffixes such as "corp," "llc," …. After merging, the joint dataset has a total of 615 companies with average emissions in the top 88th percentile of the total emissions data available from Trucost. *SI Appendix*, Table S2 shows a summary statistics of our sample.

Empirical Estimation of Organizational System Thinking. Our task is to use companies' answers to open-ended questions in the CDP questionnaire to infer if an organization (hereafter company) uses system thinking at the core of its carbon management processes. Since CDP includes both open-ended and multiple-choice questions, we first filter out the latter based on the length of the text. Preliminary exploratory data analysis showed that answers to open-ended questions tend to be larger than 200 characters, which is the cutoff value we used in our study. After isolating open-ended questions, we manually annotate approximately 2,000 examples using a human-in-the-loop natural language processes approach divided into two steps described below.

The first step is motivated by the fact that, within the ∼100,000 answers identified, the relative incidence of system thinking cues is fairly low. This implies that a manual annotation on a random subsample would require increasingly large numbers in order to achieve a critical mass of positive examples. Therefore, we build a "high-likelihood" subset of answers that overrepresents system thinking occurrences compared to the original population. The subsequent manual annotation of random draws from the high-likelihood subset can thus capture both system thinking and non-system thinking characteristics much more efficiently.

We begin to parse together the subset by using a dictionary of bigrams that might be related to system thinking in the CDP context. Bigrams are text expressions made of two words that covey a joint meaning, e.g., "system thinking." We avoid including monograms (i.e., individual words) in our dictionary because finding system thinking cues requires nuanced concepts that do not easily emerge in any individual word. We decline our definition of organizational system thinking in different connotations and build the dictionary with potentially related words. It is important to notice that our objective is merely to identify keywords whose occurrence in an answer is (even mildly) correlated with system thinking. We do not believe that isolating the concept completely with any keyword set is possible. *SI Appendix*, Table S3 illustrates the system thinking areas considered and the associated keywords for the total of 20 that forms the initial dictionary.

Our initial search led us to 4,565 answers that contained any of the initial keywords as part of the high-likelihood subset. Furthermore, following ref. 63, we add syntactically similar answers to the ones already included. Concretely, we train various machine learning methods (nearest neighbor, logit regression, decision tree, random forest, singular vector machine) on a training set built by assigning a value of 1 to all answers already in the high-likelihood subset, and 0 to a random subsample of the remaining answers. We then use inference analysis to assign labels within a test set, which, in this case, comprises all answers available outside of the training set. A correct model would assign 0 to all answers in the test set since they are virtually equivalent to the 0s used in the training set. However, if any of the algorithms misclassifies answers giving them a label of 1, then those answers must be particularly "close" to those already present in the high-likelihood subset. They are therefore added to it. In its final iteration, 8,606 answers were part of the high-likelihood subset.

From the full body of texts, and their preliminary classification, we sample a random number of questions, and we manually check whether the answers to those questions classified as system thinking in the previous step meet our theoretical definition of system thinking (*SI Appendix*, section S1). If they do, we assigned them a value of one; otherwise, we classified them as zero. We provide examples of system thinking and non-system thinking answers in *SI Appendix*, section S2, where we also showcase our annotation process practically by walking through our reasoning. The two-step procedure is necessary to simplify the annotation task and increase the quality of the training data. Two of the authors annotated the text independently, and to measure interrater reliability, we used Cohen's kappa (∼0.74). To generate the final dataset, we train a pretrained BERT model, using the HuggingFace Transformers library. To increase the reliability of the predictions, we run three independent models and then average their forecasts (i.e., their predicted labels). The three models are derived from the classifier trained on three different datasets, i.e., the two training sets of the individual annotators and an intersection training set where an answer is classified as system thinking if both annotators agreed on this classification. In the training process, we split the annotated data in a training and test set of size 80% and 20%, respectively. Subsequently, we divided the training set into training (80%) and validation (20%). Answers are randomly distributed in the training, validation, and test set in such a way as to keep an approximatively equal incidence of system thinking behavior (∼15%) across the three samples. The accuracy of the models in the hold-out test set are ∼85%, ∼80%, and ∼90%.

Finally, we use the trained models on the full body of texts of the CDP questionnaire. Organizational system thinking for company i in year y is then defined as the yearly average labels across all the company's answers in the given fiscal year. To ensure stability in the averages, we remove yearly observations with less than 10 responses. The total number of companies in the dataset goes down to 622, which is the final sample we use for our analyses. In Robustness tests, we show that our results are robust to different training strategies and to different cutoffs in the number of responses necessary to include an observation in our sample.

Notice that in our theoretical construct, organizational system thinking is an organization-level capacity. For our empirical estimation to be consistent with the theoretical construct, we need to ensure or assume that enough bodies within the company are involved with the answers to the CDP questionnaire. Otherwise, our estimation would capture the individual-level capacities of the questionnaire respondents, not group-level properties. In *SI Appendix*, section S3, we provide evidence in support of our assumption.

Empirical Specifications. To estimate the relationship between organizational system thinking and environmental outcome, we take a three-stage approach. First, we estimate the association between system thinking and several asset characteristics, namely: Market Leverage, Tangibility, Profitability, Size, Market to Book, and Invested Capital. To address endogeneity issues arising from properties of the text that could confound our estimation of organizational system thinking, we also control for several characteristics of the answers. Precisely, we control for the length of the text in the CDP answers. This is necessary because there is a strong positive association between the likelihood of observing system thinking and the length of the text from which the likelihood is estimated. This association is expected because system thinking emerges from detailed descriptions of processes and operations. Therefore, the annotated dataset had a higher incidence of system thinking among the longest answers. However, long texts do not necessarily reflect system thinking, and therefore, this association could bias the BERT model which, in theory, could simply predict one or zero based on this factor solely. To exclude this possibility, we control for the average length of the text per year in each of our specifications. Similarly, we control for standard text complexity measures (see *SI Appendix*, section S4 for further details on these measures). This control is necessary because text complexity is positively correlated with system thinking, but system thinking should not be just a proxy for complexity. To further address endogeneity issues, we control for geography, years, and sector-fixed effects in every specification. We collectively denote the control set of this regression as $\mathcal{X}$. We do not control for firm fixed effects because a) some companies go in and out of the sample; therefore, for some observations, we have a limited number of years [and so subtracting average values would not be a well-defined operation (64)] and b) organizational system thinking is measured with error, and companies' fixed effects can significantly increase the noise to signal ratio in the presence of measurement error (65).

Because disclosure to CDP is voluntary, there is a self-selectivity bias in our sample. To correct for this bias, we run a Heckman's two-stage model (66). Specifically, we first run a Probit model where the dependent variable is an indicator variable that takes the value of one if company *i* discloses information to CDP in year *y* and zero otherwise. The independent variables include Size, Tangibility, and Invested Capital as well as the proportion of companies in the same sector and country that also disclose information to CDP. In the Probit model, we also control for average emissions in the 2 y before the disclosure, since companies with lower emissions might be more likely to report their carbon management processes to their stakeholders. We collectively denote the control set of this regression as $\tilde{\mathcal{X}}$. From the fitted Probit, we estimate the inverse Mills ratio, l^{-1}, which is defined as $l^{-1} = \frac{f(x)}{1-F(x)}$, where $f(x)$, $F(x)$ are the (normal) probability density function and the cumulative distribution, respectively. The inverse Mills ratio, l^{-1}, is then used as an additional covariate in the estimation of the coefficients in the linear model. Specifically,

$$\mathbb{P}[\text{Disclose to CDP}|\tilde{\mathcal{X}}] = \Phi(\tilde{\mathcal{X}}^T\tilde{\vec{\beta}}) + \epsilon$$
$$\text{O-ST} = \mathcal{X}\vec{\beta} + \delta l^{-1} + \epsilon, \qquad [1]$$

where $\tilde{\mathcal{X}}$, O-ST, and $\mathcal{X}$ are the control set of the Probit model in the first step of the Heckman correction, organizational system thinking, and the control set of the linear model, respectively. Because the particular value and statistical significance of a coefficient could depend on the specific sample we end up with after merging the different datasets, we estimated bootstrapped regression coefficients and 95% CIs ($\alpha = 0.05$). Specifically, we generate $k = 2{,}000$ random subset of our dataset (with replacement), each of which is 85% of the total sample, and we sort the vector of coefficients obtained from estimating Eq. **1** in each of the subsamples. Then, we calculate the upper and lower bound of the 100(1-α)% CI as the $(0.5\alpha)k$th and the $(1-0.5\alpha)k$th value of the sorted vector (67).

In the second stage of our three-stage estimation approach, we retain the statistically significant factors (asset characteristics) estimated from Eq. **1**, and we used them in the control set of a series of Probit specifications, each explaining the presence or absence of a CSR policy. In the control set, we also include our organizational system thinking variable, text complexity measures, and fixed effects. We again repeat the estimation of the Probit models for different subsamples, calculate the CIs of the coefficients with bootstrapping, and extract the policies that are consistently statistically significantly associated with organizational system thinking. To compare the Probit results with those of the linear model, we standardized the coefficients following the empirical approach described in ref. 68. Specifically, we compute the empirical estimator of the standard deviation of the latent variable y^* (i.e., the class assignment probability), and then, we standardize the coefficients as in a traditional linear model using this estimator.

Finally, in the third stage of our estimation approach, we use the statistically significant factors estimated from the two previous steps to estimate the association between organizational system thinking and a) cumulative GHG emissions intensity 2 y ahead and b) alignment with the climate targets, as explained in the main text, *Results*. We calculate emission intensity as GHG emissions over total invested capital in USD. Emission intensity measures the level of emissions generated by the unit of USD invested. Also in this last step, we estimate the coefficients and their statistical significance with bootstrapping. Similarly to the previous specifications, we also control for self-selectivity, text characteristics, geography, sector, and year-fixed effects. The results of the regressions are shown in Figs. 3 and 4 in the main text. For completeness, *SI Appendix*, Tables S6 and S7 shows the estimations of the regressions from the full sample without bootstrapping.

Robustness Tests. We already run explicit robustness tests to sample size and choice of covariates in the main estimation through bootstrapping. However, to further validate our results, we run three additional robustness tests. Here, we describe the tests, and in *SI Appendix*, we show the results.

In the first test, we include two additional variables in our estimation: 1) the total number of projects implemented to lower GHG emissions and disclosed by a company in the CDP questionnaire; 2) the number of objectives covered by these projects. The objectives include, for example, "process emissions reduction," "fugitive emissions reduction," "Behavioral change," "Commuting-Employee incentive," "Raw material reduction," and "Green Financing," to cite a few. Information on the objectives is disclosed in the description of the projects in the CDP questionnaire. We include these factors to investigate whether the effect of O-ST on the environmental outcomes changes when we account for alternative, and simpler, measures of system thinking capacities. Indeed, the total effort invested in carbon management processes (quantified by the total number of projects a company is involved with) and the diversification of this effort (quantified by the number of objectives) can be seen as the capacity of a company to spread its attention toward multiple goals, which is a crucial characteristic of our operationalization of the O-ST measure. We have found that none of these factors (and their joint presence) substantially influence the association of O-ST with the environmental outcomes (*SI Appendix*, Table S11). Interestingly, however, even after accounting for the number of projects across the objectives, the number of objectives itself is statistically significantly associated with lower emissions and a higher likelihood of alignment. The regression coefficient, however, is substantially smaller than the coefficient on O-ST.

The label predictions used to measure the level of organizational system thinking in the main text were based on an ensemble of models. Each model in the ensemble was trained on a specific dataset, i.e., the two manually annotated dataset and their intersection. The final prediction was the average label predicted by the models. Here, we generate four additional label predictions to test the robustness of our results to different empirical strategies to estimate O-ST. In the first two tests, we repeat the estimation using label predictions obtained from a model trained on the training set generated by the two annotators independently. In the third test, we use label predictions obtained from training the model on an intersection dataset in which a text is classified as system thinking if and only if both annotators agree on the classification (ex-ante agreement). Finally, in the last model, we use label predictions obtained from the agreement of the predictions of the model trained on the single annotators training sets (ex-post agreement). Results are shown in *SI Appendix*, Fig. S4. Overall, we have found that the results are robust to different training strategies.

Finally, we run a robustness test to show that the results of our study are left unchanged when we change the cutoff level in the number of yearly answers to the CDP questionnaire that are necessary to include an observation in the sample. Specifically, we estimate the association between O-ST and GHG emissions and alignment with climate targets for different cutoff levels. Results are shown in *SI Appendix*, Table S10.

Data, Materials, and Software Availability. In our study, we use several datasets that we have purchased from third parties and that we therefore cannot share. These include COMPUSTAT, Trucost, Refinitiv and CDP. However, these datasets can be accessed directly from the data providers for a fee, see refs. 49, 50, 57–59, and 62, *Materials and Methods*, and *SI Appendix*, Table S4 for further information. The O-ST dataset derived in this work is publicly available together with the code to reproduce the main analyses. Data and Code have been deposited in Harvard Dataverse at https://doi.org/10.7910/DVN/5TZ16W (69).

ACKNOWLEDGMENTS. We would like to thank Jeremy Woods and Nikolaos Voulvoulis for valuable comments on earlier drafts of this manuscript.

1. A. Bjørn, S. Lloyd, D. Matthews, From the Paris agreement to corporate climate commitments: Evaluation of seven methods for setting 'science-based' emission targets. *Environ. Res. Lett.* **16**, 054019 (2021).
2. M. Bergquist, M. Thiel, M. H. Goldberg, S. van der Linden, Field interventions for climate change mitigation behaviors: A second-order meta-analysis. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2214851120 (2023).
3. R. Gifford, The dragons of inaction: Psychological barriers that limit climate change mitigation and adaptation. *Am. Psychol.* **66**, 290–302 (2011).
4. R. Salahodjaev, Is there a link between cognitive abilities and environmental awareness? Cross-national evidence. *Environ. Res.* **166**, 86–90 (2018).
5. B. Xie, M. J. Hurlstone, I. Walker, Correct me if I'm wrong: Groups outperform individuals in the climate stabilization task. *Front. Psychol.* **9**, 2274 (2018).
6. Y. Luo, J. Zhao, Motivated attention in climate change perception and action. *Front. Psychol.* **10**, 1541 (2019).
7. O. J. H. Bosch, C. A. King, J. L. Herbohn, I. W. Russell, C. S. Smith, Getting the big picture in natural resource management–systems thinking as 'method' for scientists, policy makers and other stakeholders. *Syst. Res. Behav. Sci.* **24**, 217–232 (2007).
8. T. Porter, J. Córdoba, Three views of systems theories and their implications for sustainability education. *J. Manage. Ed.* **33**, 323–347 (2008).
9. A. C. Davis, M. L. Stroink, The relationship between systems thinking and the new ecological paradigm. *Syst. Res. Behav. Sci.* **33**, 575–586 (2016).
10. S. B. Lezak, P. H. Thibodeau, Systems thinking and environmental concern. *J. Environ. Psychol.* **46**, 143–153 (2016).
11. M. T. Ballew, M. H. Goldberg, S. A. Rosenthal, A. Gustafson, A. Leiserowitz, Systems thinking as a pathway to global warming beliefs and attitudes through an ecological worldview. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8214–8219 (2019).
12. N. Voulvoulis *et al*., Systems thinking as a paradigm shift for sustainability transformation. *Glob. Environ. Change* **75**, 102544 (2022).
13. J. M. Randle, M. L. Stroink, The development and initial validation of the paradigm of systems thinking. *Syst. Res. Behav. Sci.* **35**, 645–657 (2018).
14. National Research Council, *Science for Environmental Protection: The Road Ahead* (The National Academies Press, Washington, DC, 2012).
15. Next Generation Science Standards, The three dimensions of science learning (2018). https://www.nextgenscience.org/. Accessed 5 May 2023.
16. R. E. Dunlap, K. D. Van Liere, A. G. Mertig, R. E. Jones, New trends in measuring environmental attitudes: Measuring endorsement of the new ecological paradigm: A revised nep scale. *J. Soc. Issues* **56**, 425–442 (2000).
17. IAE, Cumulative emissions reduction by mitigation measure in the net zero scenario, 2021–2050 (2023). https://www.iea.org/data-and-statistics/charts/cumulative-emissions-reduction-by-mitigation-measure-in-the-net-zero-scenario-2021-2050. Accessed 5 May 2023.
18. S. Dietz *et al*., An assessment of climate action by high-carbon global corporations. *Nat. Clim. Change* **8**, 1072–1075 (2018).
19. S. Dietz, D. Gardiner, V. Jahn, J. Noels, How ambitious are oil and gas companies' climate goals? *Science* **374**, 405–408 (2021).
20. J. E. McGrath, *Groups: Interaction and Performance* (Prentice-Hall, Englewood Cliffs, NJ, 1984).
21. A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, T. W. Malone, Evidence for a collective intelligence factor in the performance of human groups. *Science* **330**, 686–688 (2010).
22. A. Almaatouq, M. Alsobay, M. Yin, D. J. Watts, Task complexity moderates group synergy. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101062118 (2021).
23. X. Wu *et al*., A survey of human-in-the-loop for machine learning. *Futur. Gener. Comput. Syst.* **135**, 364–381 (2022).
24. P. H. Thibodeau, C. M. Frantz, M. L. Stroink, Situating a measure of systems thinking in a landscape of psychological constructs. *Syst. Res. Behav. Sci.* **33**, 753–769 (2016).
25. R. D. Arnold, J. P. Wade, A definition of systems thinking: A systems approach. *Proc. Comput. Sci.* **44**, 669–678 (2015).
26. K. E. Maani, V. Maharaj, Links between systems thinking and complex decision making. *Syst. Dyn. Rev.* **20**, 21–48 (2004).
27. A. C. Davis, W. Leppanen, K. P. Mularczyk, T. Bedard, M. L. Stroink, Systems thinkers express an elevated capacity for the allocentric components of cognitive and affective empathy. *Syst. Res. Behav. Sci.* **35**, 216–229 (2018).
28. J. D. Sterman, Does formal system dynamics training improve people's understanding of accumulation? *Syst. Dyn. Rev.* **26**, 316–334 (2010).
29. E. Moxnes, Misperceptions of basic dynamics: The case of renewable resource management. *Syst. Dyn. Rev.: J. Syst. Dyn. Soc.* **20**, 139–162 (2004).
30. S. Grewatsch, S. Kennedy, P. Tima Bansal, Tackling wicked problems in strategic management with systems thinking. *Strat. Organ.* **21**, 721–773 (2021).
31. J. D. Sterman, Risk communication on climate: Mental models and mass balance. *Science* **322**, 532–533 (2008).
32. G. P. Hodgkinson, M. P. Healey, Cognition in organizations. *Annu. Rev. Psychol.* **59**, 387–417 (2008).
33. J. P. Walsh, Managerial and organizational cognition: Notes from a trip down memory lane. *Organ. Sci.* **6**, 280–321 (1995).
34. D. Nicolini, Comparing methods for mapping organizational cognition. *Organ. Stud.* **20**, 833–860 (1999).
35. M. G. Bougon, Congregate cognitive maps: A unified dynamic theory of organization and strategy. *J. Manag. Stud.* **29**, 369–387 (1992).
36. S. Nadkarni, V. Narayanan, Validity of the structural properties of text-based causal maps: An empirical assessment. *Organ. Res. Methods* **8**, 9–40 (2005).
37. S. W. Kozlowski, D. R. Ilgen, Enhancing the effectiveness of work groups and teams. *Psychol. Sci. Public Interest* **7**, 77–124 (2006).
38. S. Mohammed, L. Ferzandi, K. Hamilton, Metaphor no more: A 15-year review of the team mental model construct. *J. Manage.* **36**, 876–910 (2010).
39. L. A. DeChurch, J. R. Mesmer-Magnus, The cognitive underpinnings of effective teamwork: A meta-analysis. *J. Appl. Psychol.* **95**, 32 (2010).
40. M. P. Healey, T. Vuori, G. P. Hodgkinson, When teams agree while disagreeing: Reflexion and reflection in shared cognition. *Acad. Manag. Rev.* **40**, 399–422 (2015).
41. J. A. Cannon-Bowers, E. Salas, Reflections on shared cognition. *J. Organ. Behav.* **22**, 195–202 (2001).
42. P. A. Corning, The re-emergence of emergence, and the causal role of synergy in emergent evolution. *Synthese* **185**, 295–317 (2012).
43. M. Tripsas, G. Gavetti, Capabilities, cognition, and inertia: Evidence from digital imaging. *Strat. Manage. J.* **21**, 1147–1161 (2000).
44. G. Gavetti, Cognition and hierarchy: Rethinking the microfoundations of capabilities' development. *Organ. Sci.* **16**, 599–617 (2005).
45. G. P. Hodgkinson, M. P. Healey, Psychological foundations of dynamic capabilities: Reflexion and reflection in strategic management. *Strat. Manage. J.* **32**, 1500–1516 (2011).
46. L. Luo, Q. Tang, Does voluntary carbon disclosure reflect underlying carbon performance? *J. Contemp. Accoun. Econ.* **10**, 191–205 (2014).
47. B. Doda, C. Gennaioli, A. Gouldson, D. Grover, R. Sullivan, Are corporate carbon management practices reducing corporate carbon emissions? *Corp. Soc. Respon. Environ. Manage.* **23**, 257–270 (2016).
48. UNFCCC, The Paris Agreement (2016). https://unfccc.int/process-andmeetings/the-paris-agreement. Accessed 5 May 2023.
49. Trucost, Trucost environmental (2023). https://www.marketplace.spglobal.com/en/datasets/trucost-environmental-(46). Accessed 24 February 2023.
50. Trucost, Trucost Paris alignment (2023). https://www.marketplace.spglobal.com/en/datasets/trucost-paris-alignment-(186). Accessed 24 February 2023.
51. P. Bolton, M. Kacperczyk, Do investors care about carbon risk? *J. Financ. Econ.* **142**, 517–549 (2021).
52. A. Clark, C. Fox, S. Lappin, *The Handbook of Computational Linguistics and Natural Language Processing, Blackwell Handbooks in Linguistics* (Wiley, 2012).
53. L. L. Thompson, E. R. Wilson, "Creativity in teams" in *Emerging Trends in the Social and Behavioral Sciences*, R. Scott, S. Kosslyn, Eds. (John Wiley & Sons, Inc., Hoboken, NJ, 2015), pp. 1–14.
54. S. Cenci, S. Kealhofer, A causal approach to test empirical capital structure regularities. *J. Finance Data Sci.* **8**, 214–232 (2022).
55. R. E. Freeman, J. S. Harrison, S. Zyglidopoulos, *Stakeholder Theory: Concepts and Strategies, Elements in Organization Theory* (Cambridge University Press, Cambridge, 2018).
56. U. Malmendier, "Chapter 4—Behavioral corporate finance" in *Handbook of Behavioral Economics–Foundations and Applications. Handbook of Behavioral Economics: Applications and Foundations* B. D. Bernheim, S. DellaVigna, D. Laibson, Eds. (North-Holland, 2018), vol. 1, pp. 277–379.
57. Compustat, Compustat financials (2023). https://www.marketplace.spglobal.com/en/datasets/compustat-financials-(8). Accessed 24 February 2023.
58. Refinitiv, Fundamentals (2023). https://www.refinitiv.com/en/financial-data/company-data/fundamentals-data. Accessed 24 February 2023.
59. Refinitiv, Environmental, social and corporate governance–ESG (2023). https://www.refinitiv.com/en/financial-data/company-data/esg-data. Accessed 24 February 2023.
60. C. de Villiers, J. Jia, Z. Li, Corporate social responsibility: A review of empirical research using Thomson Reuters asset4 data. *Account. Finance* **62**, 4523–4568 (2022).

61. T. Busch, M. Johnson, T. Pioch, Corporate carbon performance data: Quo vadis? *J. Ind. Ecol.* **26**, 350–363 (2022).
62. CDP, Corporate response dataset (2023). https://www.cdp.net/en/data/corporate-data. Accessed 24 February 2023.
63. G. King, P. Lam, M. E. Roberts, Computer-assisted keyword and document set discovery from unstructured text. *Am. J. Polit. Sci.* **61**, 971–988 (2017).
64. R. Rajan, P. Ramella, L. Zingales, *What Purpose do Corporations Purport? Evidence from Letters to Shareholders* (National Bureau of Economic Research, 2023).
65. Z. Griliches, J. A. Hausman, Errors in variables in panel data. *J. Econ.* **31**, 93–118 (1986).
66. J. J. Heckman, Sample selection bias as a specification error. *Econometrica* **47**, 153–161 (1979).
67. K. J. Preacher, A. F. Hayes, Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* **40**, 879–891 (2008).
68. J. B. Grace, D. J. Johnson, J. S. Lefcheck, J. E. K. Byrnes, Quantifying relative importance: Computing standardized effects in models with binary outcomes. *Ecosphere* **9**, e02283 (2018).
69. M. Burato, S. Tang, V. Vastola, S. Cenci, Organizational system thinking as a cognitive framework to meet climate targets. Harvard Dataverse. https://doi.org/10.7910/DVN/5TZ16W. Deposited 15 September 2023.