Authors
Fabio Galeotti
Astrid Hopfensitz
César Mantilla
Coordination
Cecilia Poggi (AFD)
Linda Zanfini (AFD)

NOVEMBER 2023
No. 297

Research papers

Climate Change Education from the Perspective of Social Norms.

A Systematic Review

ÉDITIONS AFD
AGENCE FRANÇAISE DE DÉVELOPPEMENT

Agence française de développement

Papiers de recherche

Les *Papiers de Recherche de l'AFD* ont pour but de diffuser rapidement les résultats de travaux en cours. Ils s'adressent principalement aux chercheurs, aux étudiants et au monde académique. Ils couvrent l'ensemble des sujets de travail de l'AFD : analyse économique, théorie économique, analyse des politiques publiques, sciences de l'ingénieur, sociologie, géographie et anthropologie. Une publication dans les *Papiers de Recherche de l'AFD* n'en exclut aucune autre.

Les opinions exprimées dans ce papier sont celles de son (ses) auteur(s) et ne reflètent pas nécessairement celles de l'AFD. Ce document est publié sous l'entière responsabilité de son (ses) auteur(s) ou des institutions partenaires.

Research Papers

AFD Research Papers are intended to rapidly disseminate findings of ongoing work and mainly target researchers, students and the wider academic community. They cover the full range of AFD work, including: economic analysis, economic theory, policy analysis, engineering sciences, sociology, geography and anthropology. *AFD Research Papers* and other publications are not mutually exclusive.

The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the position of AFD. It is therefore published under the sole responsibility of its author(s) or its partner institutions.

Climate Change Education from the Perspective of Social Norms

A Systematic Review

AUTEURS

Fabio Galeotti
Groupe d'Analyse
et de Théorie Economique
Lyon St-Etienne (GATE)
France

Astrid Hopfensitz
EMLYON Business School
France

César Mantilla
Universidad Loyola Andalucía
Spain

COORDINATION

Cecilia Poggi (AFD)

Linda Zanfini (AFD)

Abstract
One of the most challenging problems of society today is climate change. Faced with this problem, people must change the way they behave to mitigate the effects of climate change and adapt to them. Such change in behavior can be achieved through a normative shift, i.e., a change in the social norms that regulate people's social interactions and behavior in a society or a group. Climate change education (CCE) can be a powerful tool to achieve this shift. In this paper, we conduct a systematic review (SR) of the literature on CCE, with the aim of offerring a comprehensive overview of the empirical research in this field. We particularly emphasize studies that assess the effects of educational interventions on social norms. Specifically, we focus on studies that either measure actual behaviors or investigate individuals' beliefs regarding the prevalence or acceptability of these behaviors in a society or reference group. We identify 86 studies evaluating CCE interventions. Among these, only 19 look at the effects of CCE on norm-related beliefs or actual behavior. Among the 86 studies, we find a disproportionate focus on interventions conducted in high-income, less climate-vulnerable countries and urban populations, with a general absence of cross-country comparisons. Most studies also employ pre-post evaluations, which are more susceptible to demand effects and social desirability bias. Among the 18 studies that look at norms, only few of them provide a belief-based measure of social norms. The vast majority measures actual behavior, mainly in terms of recycling, trashing and energy saving. Most interventions involve activities aimed at engaging learners. Others focus on nudges (like stickers or posters). A minority is based on lectures, deliberative discussions and interaction with scientists or science in general. The results of this SR reveal important gaps in the literature and potential tensions that can inform future research in this area.

Keywords
Climate change education, systematic review, children and young people, social norms, pro-environmental behavior

Acknowledgements
We would like to thank our research assistants involved in the project: Bengisu Baś, Rito Chakraborty, Anandhu Krishna, Emma Klein, Lisa Salvetti, Silvia Ortíz-Merchán, Sofía Almeida, Ana Gaona, and Jeanne Cunéo. We are also grateful to the people at the French Agency of Development, and especially Cecilia Poggi, Linda Zanfini, Rita Abdel Sater, and Emmanuel Josoa Ramaholimihaso, for helpful comments on an earlier drafts of this paper. This research has benefited from the financial support of the French Agency of Development.

JEL Classification
A20, Q54, Q56.

Original version
English

Accepted
November 2023

Résumé
Le changement climatique est l'un des problèmes les plus pressants pour la société actuelle. Pour y faire face, il est nécessaire de changer les comportements, pour atténuer ses effets et s'y adapter. Ces changements de comportements peuvent être opérés à travers un changement normatif, c'est-à-dire une modification des normes sociales qui régissent les comportements et les interactions dans une société ou un groupe. L'éducation au changement climatique (ECC) peut être un outil puissant pour réaliser ces changements. Dans ce papier de recherche, nous procédons à une revue systématique de la littérature sur l'éducation au changement climatique, en vue d'offrir un panorama exhaustif de la recherche empirique dans ce domaine. Nous mettons particulièrement l'accent sur les études qui évaluent les effets des interventions éducatives sur les normes sociales. Plus précisément, nous nous concentrons sur les études qui mesurent les comportements réels ou qui examinent les croyances des individus concernant la prévalence ou l'acceptabilité de ces comportements dans une société ou un groupe de référence. Nous avons identifié 86 études qui évaluent les interventions en matière d'ECC, parmi lesquelles seulement 19 examinent les effets de l'ECC sur les croyances liées aux normes ou sur les comportements réels. Parmi ces 86 études, nous constatons un accent disproportionné sur les interventions menées dans les pays à revenu élevé - moins vulnérables au changement climatique, et auprès des populations urbaines, et une absence générale de comparaisons entre les pays. La plupart des études utilisent également des évaluations pré-post, plus sensibles aux effets de demande et au biais de désirabilité sociale. Parmi les 18 études portant sur les normes, seules quelques-unes fournissent des mesures basées sur les croyances en matière de normes sociales. La grande majorité mesure le comportement réel, principalement en termes de recyclage, de mise en déchets et d'économie d'énergie. La plupart des interventions impliquent des activités visant à impliquer les apprenants. D'autres se concentrent sur des « nudges » (« coup de pouce »), (comme les autocollants ou les affiches). Une minorité est basée sur des conférences, des discussions délibératives et des interactions avec des scientifiques ou la science en général. Les résultats de cette RS révèlent d'importantes lacunes dans la littérature et des tensions potentielles qui peuvent éclairer les recherches futures dans ce domaine.

Mots-clés
Education au changement climatique, revue systématique, enfants et jeunes, normes sociales, comportement pro-environnemental

Remerciements
Nous tenons à remercier nos assistants de recherche impliqués dans le projet : Bengisu Baś, Rito Chakraborty, Anandhu Krishna, Emma Klein, Lisa Salvetti, Silvia Ortíz-Merchán, Sofía Almeida, Ana Gaona et Jeanne Cunéo. Nous remercions également le personnel de l'Agence française de développement, et en particulier Cecilia Poggi, Linda Zanfini, Rita Abdel Sater et Emmanuel Josoa Ramaholi-mihaso, pour les commentaires utiles qu'ils ont formulés sur les premières versions de ce document. Cette recherche a bénéficié du soutien financier de l'Agence française de développement.

Introduction

The consequences of climate change, including biodiversity loss (Warren et al., 2013), the destruction or alteration of ecosystems (Cooley et al., 2022), and the reduction of global agricultural productivity (Ortiz-Bobea, 2021), are becoming more and more visible (Lee et al., 2015; Trott, 2020). Unfortunately, awareness alone of this global problem seems to be insufficient to solve it due to its 'social dilemma' nature (Milinski et al., 2008; Bisaro and Hinkel, 2016). Even though it would be beneficial on a global scale for humans to reduce their environmental impact and prepare for future climate-related uncertainty, the adoption of more sustainable or climate-responsible actions entails a high individual cost, whereas most benefits are collective and not individual.[1] Thus, individuals are likely to pursue their own private interests by abstaining from true mitigation and adaptation actions while free riding on the eco-sustainable behaviors of others (Barrett and Dannenberg, 2014).

One potential solution to the climate change dilemma is education, especially of children and young people. Following Anderson (2012), we define climate change education (hereafter CCE) as the set of curricular and extracurricular activities aiming at: (i) conveying relevant skills and knowledge on mitigation (e.g., sustainable lifestyles and consumption) and adaptation (i.e., disaster risk reduction and preparedness) issues in the context of an uncertain climate fluctuation; (ii) promoting safe, climate resilient and sustainable learning spaces; (iii) engaging the active participation of the community as agents of change; and (iv) enhancing interactions between education policy-makers and climate researchers. The literature on CCE has grown exponentially in recent years. In this paper, we conduct a systematic review (SR) of this literature with the aim of providing a comprehensive picture of the empirical research on CCE for children, pre-adolescents, and young adults, discussing its main findings and approaches. A particular focus will be devoted to studies that investigate the effects of CCE interventions on social norms or, more generally speaking, studies that measure norms as part of their evaluation instruments.

Social norms are so far often overlooked in this literature, though they might play an important role. Sustainable and lasting transformation in our approach to climate

[1] The collective gains from adopting pro-environmental behaviors are also mainly realized in the long term, while the benefits of pursuing one's own personal interests are short term. This exacerbates even further the social dilemma.

change requires changes in formal as well as informal institutions. While governments focus primarily on formal institutions, such as laws and treaties, informal institutions might play an equally important, if not larger, role. Effective formal institutions are often difficult to put in place because of the resistance of private interests that would be adversely affected by such interventions. This is, for example, the case of laws that strictly regulate the rights of certain companies or sectors to make profitable businesses. Similarly, laws that would control individual behaviors accountable for climate change (e.g., how much meat people eat, what they drive, how much water they use, and where they live) would be met with strong opposition by many citizens (Sparkman et al., 2021).

The effectiveness of formal institutions is thus complemented by informal institutions, such as social norms (Nyborg et al., 2016). Social norms identify what behaviors are socially appropriate and what are not in a society or group. They are defined as the sum of two components: an empirical component (often referred to as "descriptive norm"), which captures what most people typically do, and an injunctive component (often referred to as "injunctive norm"), which captures what most people socially disapprove or approve of (Cialdini et al., 1990; Bicchieri, 2006).[2] Another important distinction is between perceived and actual social norms. Perceived social norms refer to the beliefs that people have about the descriptive and injunctive norm (Farrow et al., 2017). Perceived and actual norms often coincide, but they can also sometimes differ, for example when people overestimate (or underestimate) the prevalence or acceptability of a given behavior in a group or society. In that case, what matters for the establishment of a social norm is what people perceive. According to this conceptualization, social norms can be assessed either by inquiring people about their beliefs regarding the prevalence or acceptability of a given behavior in society (or reference group) or by observing what most people do.[3]

Social norms can be a powerful lever to address large-scale collective action problems such as climate change, for at least two reasons (Ostrom, 2000; Fehr and Fischbacher, 2004; Biel and Thøgersen,

[2] People can also hold personal views about what constitutes appropriate or inappropriate behavior. These are personal norms, which differ from injunctive social norms because they capture what one approves of rather than what society or the reference group approves of. In our literature review, we disregard studies that only look at personal norms.

[3] One could also look at what most people personally disapprove of or approve of as an indicator of the injunctive norm. One problem with that is that there could be a disparity between what people personally think one should do and what they think all other members of their network believe one ought to do. This state is called pluralistic ignorance and could produce wrong inference about the injunctive norm if one only examines the personal beliefs or preferences of people.

2007; Bicchieri, 2016). First, social norms directly tackle the free-rider problem by changing the material and emotional incentives associated with pro-environmental behavior. Incentives change because people feel guilt or shame if they do not behave in line with what is considered as good or appropriate behavior (Bicchieri et al., 2018). People also anticipate approval (or disapproval) by others if they (do not) comply with the norm. Second, social norms increase people's acceptability of formal institutional interventions and control policies.

The problem with using social norms as a solution to climate change is that many of the behaviors responsible for climate change are currently the norm (Sparkman et al., 2021). This requires a deep understanding of how social norms can be changed, and what actions or interventions can facilitate a normative shift to enable climate change mitigation and adaptation. Children and young adults play a key role in this shift, as they can foster child-to-parent learning and transmit their concerns about the future regarding climate change (Lawson et al., 2019; Crandon et al., 2022).

Many policymakers, researchers, and practitioners from education science, environmental science, psychology, and economics argue that one way to induce a normative shift in the global approach to climate change is through education, especially among young people. They believe this "would lead to greater national action and commitment" against climate change (UNESCO and UNFCCC, 2016). The idea that CCE is the key to unlocking a greener future inspired a burst of research on this topic in the last few decades (Monroe et al., 2017).

Why is it important to do this review?

Despite the vast and increasing amount of scientific information on CCE (and more in general on Education for Sustainable Development—ESD hereafter), there is no clear guidance on which educational interventions are most effective in instigating positive norm changes related to the climate and improving environmental engagement. This problem encompasses multiple issues. First, a large body of the literature on CCE is not concerned with measuring the impact of CCE on climate change mitigation and adaptation actions. When it does, the focus is often on assessing the effects on knowledge, attitudes, or psychological constructs like self-efficacy rather than measuring actual behaviors and norms. Second, CCE interventions are extremely heterogeneous in their contents and quality, and in how they could modify behavior and norms (Ategeka et al., 2022). Third, the effectiveness of CCE is usually assessed by the same authors implementing the program, with all the biases that this approach entails. Fourth, in many studies, it is often difficult to track environmentally oriented behaviors and

norm-related beliefs among those receiving CCE (i.e., a measurement problem) or to separate environmental attitudes after receiving CCE from the pre-existing attitudes that initially led individuals to demand CCE (i.e., a selection problem).

Moreover, educational interventions that could work in developed economies may not work in developing or emerging countries. Hence, whereas climate change mitigation and adaptation are global problems, the scalability and adaptability of the interventions that work is unclear. This could stem from different reasons. In developing or emerging countries, people may, for example, be more reluctant to adopt less carbon-intensive lifestyles because they feel less responsible for climate change, thinking that other more advanced economies should carry the burden of climate mitigation. They may also follow, in general, weaker environmental norms or have specific background factors (e.g., poor socio-economic conditions, low levels of literacy and school achievement) that make them less responsive to educational interventions. On the other hand, developing countries might be already more exposed to the consequences of climate change and thus have clearer objectives regarding conservation or mitigation behavior.

A clear account and comparison of the effects of educational interventions in low- and middle- vs. high-income countries is, however, surprisingly missing.

Objectives

Our SR aims at identifying the educational interventions within the definition of CCE whose effect is not limited to information transmission but transcends to changes in actual behavior and norm-related beliefs. We exclude purely conceptual studies and studies that only evaluate or analyze the impact of CCE interventions on knowledge, attitudes, behavioral intentions, psychological states, and self-reported behaviors. This is because we are interested in interventions that have the potential to trigger a normative shift, either by changing what people collectively think one should do (injunctive norms) or what they actually do (descriptive norms). Indeed, knowledge alone about climate change is often not enough to induce a change in behavior, much less a normative shift (see, e.g., Dijkstra and Goedhart 2012; Brownlee et al., 2013). Attitudes, intentions, and self-reports, instead, do not always translate into actual behavior, and are more susceptible to cognitive, social desirability and communicative bias (see, e.g., Schwarz and Oyserman, 2001, Kormos and

Gifford, 2014; Parry et al., 2021; Koller et al., 2023).[4]

More specifically, the objectives of this SR are:

- Taking a snapshot of the existing empirical research on CCE, with a particular focus on studies that measure the impact of CCE interventions on norms and actual behavior, and that use norm elicitation as part of their outcomes;

- Identifying trends, tensions and gaps in the literature that can inform new directions for future research in this area;

- Producing a comparative exercise between CCE interventions conducted in high-income countries and in low- and middle-income countries. The purpose is to identify knowledge, measurement, methodology and impact gaps, as well as potential normative barriers that prevent changes in behavior and normative beliefs in developing or emerging countries.

Preview of findings

Previous systematic reviews on CCE praise the increasing role of more engaging and active teaching methods, while raising questions on the needs of more inter-disciplinary collaborations around CCE (Monroe et al., 2017; Rousell and Cutter-Mackenzie-Knowles, 2020). These call for interdisciplinarity departs from the need to go beyond the natural sciences' information about climate change and see it as a more complex system once one accounts for social interactions.

The majority of studies on CCE that we reviewed are concerned with measuring attitudes, knowledge, intentions or psychological states. Very few studies clearly focus on the effect of CCE on either injunctive (3 studies) or descriptive (16 studies) norms. When they do, they mainly assess actual behavior rather than norm-related beliefs.

Most of the included studies examine interventions conducted in high-income countries and urban settings. Only a few studies compare interventions across multiple countries. In terms of methods, the majority of the studies employ pre-post evaluations, which are more susceptible to demand effects and social desirability bias.

The published and included results further give the impression that any kind of intervention has a positive effect. Almost all studies present positive results in the direction of improving climate change understanding and green skills in children and youth.[5] As a result, it is hard to derive

[4] Focusing on attitudes and intentions is also criticized for being the "low hanging fruits of CCE inquiry" (Reid, 2019).

[5] The few studies that provide more mixed results either show that a program works in the short-run

precise policy recommendations on which interventions work best. It is also difficult to ascertain the extent to which the lack of null results is the result of a publication bias or a genuine capacity to identify interventions that always work. We therefore take an alternative approach. Instead of evaluating the literature based on the reported impact on behavior and norms, we classify the different interventions according to their type and targeted outcome. This mapping allows us to see clearly what type of interventions and outcomes receive the most attention in the literature. We map each one of the 19 studies that focus on injunctive and descriptive norms into an outcome type (energy use, recycling, food waste, provision of infrastructure, interaction with nature, responses on a survey, and a multidimensional category) and an intervention type (engaging learners, lecture-based activities promoting discussion and interactions with science, fostering the communities' involvement, and nudging-type interventions). In terms of outcome types, we find that studies focusing on recycling behavior and clean-up, as well as reduction of energy and electricity use, are the most frequent among those measuring behavior. In terms of intervention types, there is an ample dominance of interventions aiming to engage learners.

In the following, we will detail our search strategy, give an overview of a larger set of studies that still contains studies focusing on knowledge or behavioral intentions to show the distribution of research across countries and fields. We then proceed to a detailed analysis of the final set of papers that study either injunctive or descriptive norms. We will first give an overview of the different behavioral outcomes that these studies focus on. For example, most studies focus on behaviors related to recycling and electricity use, with few studies studying impactful behaviors like transport or food choices. We will then proceed to a discussion of the interventions employed by the different studies. We will show that most of the studies focus on interventions aimed at engaging learners in some way.

but its effects did not persist in the medium/long run (Fröhlich et al. 2013) or that an intervention is effective in one school, where energy renovations were carried out before the intervention, but not in others (Pietrapertosa et al., 2021).

1. Methodology

In this section, we describe in detail the search that was performed to determine the universe of studies that were further refined and scrutinized in the data collection step. We start with a definition of the general inclusion and exclusion criteria. We then provide a description of our search strategy and the data extraction procedure.

1.1. Inclusion and exclusion criteria

The following is a list of characteristics that helped us define what studies entered the main analysis.

Type of interventions

We only included interventions with a clear educational component, encompassing both curricular and extracurricular activities conducted in schools, colleges, or universities. Curricular activities are part of the academic curriculum of any educational institution. Examples are lectures, science laboratories, field trips, and gaming activities. *Extracurricular* activities are carried out outside of the regular academic curriculum. These include awareness-raising campaigns, school-based projects, community-based projects targeting adolescents and young adults, meetings with climate scientists and policymakers, internships, and visits to facilities whose objectives are aligned with nature's preservation (e.g., natural parks or zoos).

We also included informal activities as long as they specifically target children and young adults. According to the Council of Europe, the definition of *informal education* is that it arises from learner's involvement in activities that are not undertaken with a learning purpose in mind. We argue that, in the case of children or young adults, most of these activities are implicitly captured as part of our definition of extracurricular activities. By contrast, we did not consider informal interventions that were directed at the general population. Hence, we exclude activities that fall within UNESCO's definition of *non-formal education* (i.e., education "caters for people of all ages")[6] unless there is a specific focus on children or young adults.

Regarding the interventions that typically use simple mechanisms of information transmission (e.g., posters, stickers, email or letter reminders) and rely on the psychological methods of nudges, we only considered those interventions in which nudging components were combined with some salient feature of an educational intervention (e.g., stickers reminding to turn off the lights that resemble a paper-board game played in class with the same purpose, as in Mattsson and Laike, 2022) or modifications to the educative environment as part of a larger project (e.g., stickers in a college as part of a "community-based" social marketing program to decrease carbon emissions, as in Frantz et al., 2016).

Target of interventions

We focused on educational interventions targeted at **children, pre-adolescents, and young adults**. Participants ranged from children in pre-primary school to students in master's programs. The age range of interest went from 3 to 24 years old.[7]

[6] https://uis.unesco.org/en/glossary-term/non-formal-education (accessed on the 16th of October, 2023).

[7] We have expanded the age range of interest with respect to the companion document describing the protocol of our SR. The reason is that the final set of studies directly measuring norms or observed behavior, our main focus in this SR, was small. Moreover, in five studies the age groups 5-10 and 11-13 were simultaneously targeted, so this expansion allowed us to have a better track of the studied outcomes from very young children to young adults.

Countries of interest

We did not impose restrictions on the countries of interest. We coded the country's information with the purpose of making a thorough analysis of variation by country's income levels regarding the types of interventions and measured outcomes. We also coded whether the target population had some specific characteristics relative to the country of implementation (e.g., income-based targeting, or population in rural or urban areas).

Outcomes of interest

The focus of our systematic review is to investigate social norms related to environmental behaviors. Therefore, we excluded studies that were unrelated to climate change mitigation, adaptation, conservation, sustainability, or any other forms of environmental awareness that could lead to changes in social norms. During the initial search, we did not impose restrictions on the type of outcomes, which means we also included studies that focused on knowledge, attitudes, behavioral intentions, psychological states, and self-reported behaviors. In the discussion of the results, we will quantify the prominence of these studies compared to studies that explicitly measure injunctive and descriptive norms. However, our primary focus in the discussion will be on the latter.

Language

We only included studies in English, French, Spanish, German, and Italian.

Formats

We limited the analysis to articles published in peer-reviewed journals or as working papers. We excluded conference proceedings, dissertations, theses, book reviews, and evaluation reports because they are more difficult to search efficiently and evaluate. Their contents might also be published elsewhere.

Content criteria for the initial selection of the studies

We define three criteria regarding the environmental intervention that the included studies must initially meet. Specifically, we only consider studies that are clearly related to the environment, to CCE and that present a clear identifiable intervention. Below are the instructions we provide to our coders, regarding these criteria:

- **Must be related to the environment:** Please code as "1" if the entry is discussing an environmental issue. Environmental issues include anything related to climate change, but also about conservation of the environment or particular species, and strategies engaging sustainable behavior (recycling, energy saving, water saving etc.). If the entry does not cover any topic related to environmental issues, please code it as "0".
- **Must be related to education:** Please code as "1" if the entry refers to an educational activity conducted in schools or universities or training activities that are targeting communities or relevant stakeholders (e.g., farmers, fishermen, truck-drivers etc.). If the entry does not fit the criteria described above, please code it as "0".
- **Must include an intervention:** Please code as "1" if the entry describes a program or intervention being analyzed. For instance, the introduction of some specific topics in the curriculum, a program with passive or active learning activities at the school, or training programs. Entries coded as "0" are usually those in which there was only a measurement of attitudes (e.g., applying a survey) without a clear intervention.

When it comes to the latter criterion, the coders were already familiar with the concept of an impact evaluation, which is a term used for various quantitative methods that help gather information about the effects of an intervention. However, in this initial selection of the studies, we instructed the coders to include a wide range of interventions in their sample, even those that were not very specific, as well as more focused analyses, like impact assessments.

All three criteria above must be met to be included in the SR.

1.2. Search strategy for finding eligible studies

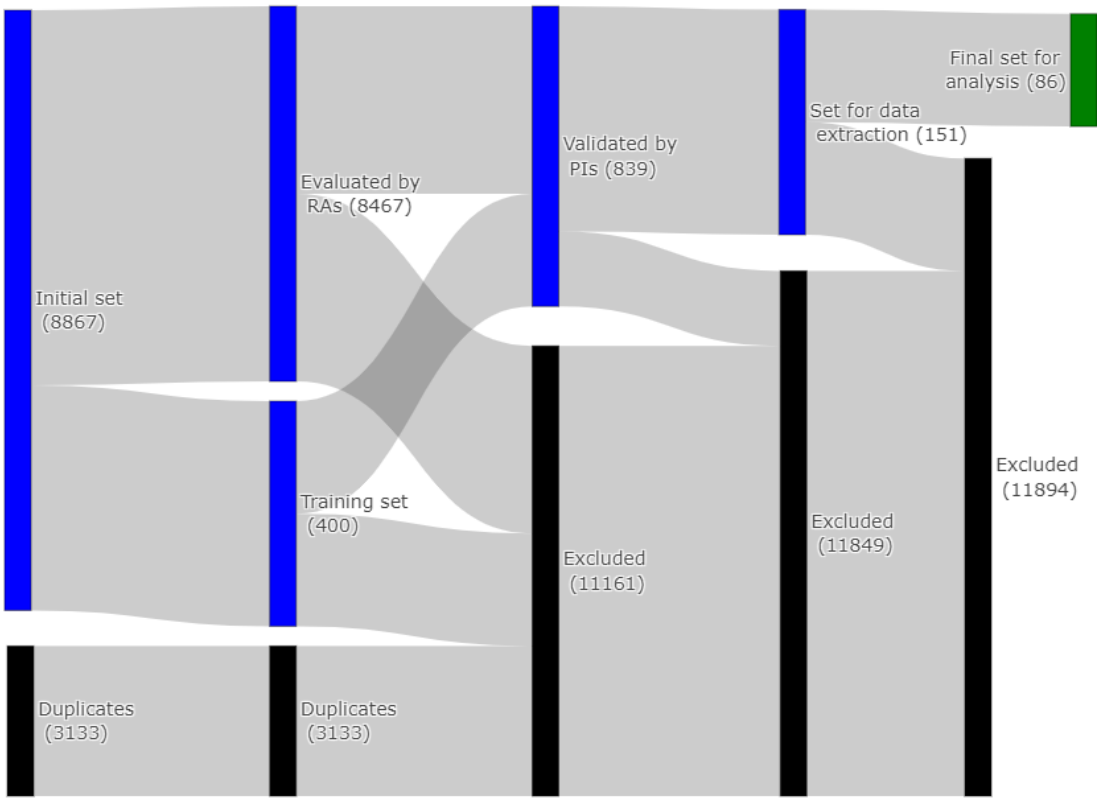
The process for finding eligible studies is depicted in Figure 1. Below, we describe the steps made into this selection strategy.

Initial mapping of terms related to climate change education

We predefine a list of terms retrieved from previously published reviews related to CCE and relevant thesaurus. See this list in Table A.1. in the appendix. The initial set of forty keywords (see Box A.1) was too redundant and some terms too specific, so we refined this list and reduced it to twenty-two keywords, which are listed in the first column in Table A.2.

The next three columns of Table A.2 added terms related to behavioral outcomes (column 2), education (column 3), and interventions (column 4). In columns 1, 2, and 3, the bottom terms that appear shaded in gray were excluded from the search key, as they were redundant with the included terms (e.g., we did not look for "informal education" since the relevant hits would appear when looking at "education").

Figure 1. Summary of the process for selecting eligible studies for the review



Source: Authors' depiction. Original.

Definition of the search key

The terms in columns 1 to 4 in Table A.2 were combined to produce the following search key:

Box 1. Search key

AB "term related to climate change or environment in the first column" AND AB ("behavior" OR "behaviour" OR "beliefs" OR "social norm*" OR "injunctive norm*" OR "descriptive norm*") AND AB ("education" OR "school" OR "teaching") AND AB ("experiment" OR "intervention" OR "program" OR "training")

We conducted the search using EBSCOhost, an online research platform that can simultaneously access multiple databases using a single interface. The number of databases that can be accessed depends on the subscription. We accessed EBSCOhost through the Universidad del Rosario in April 2023, which could, at that time, access 21 databases, including: Academic Search (Complete and Inxed), EBSCO eBooks, EconLIT, ERIC, MEDLINE, Psychology and Behavioral Sciences Collection, among others.

After detecting too much overlap or keys with very few entries, we eliminated the terms located in gray cells in Table A.2.

The search key, after excluding the gray cells, produced over 12,000 hits. The first exclusion of identical duplicates left 8,867 entries to be manually classified by the research team.

Generation of a training set

The three principal investigators (PIs) developed a training set including 400 entries. Each PI coded 200 entries: 100 unique entries, and 100 common to all three PIs. For each entry, the PIs submitted a recommendation between "include", "exclude", or "uncertain"; and coded the three inclusion criteria described in Section 1.1: whether the entry was related to the environment and education, and whether it included an assessed intervention.

Disagreements on the 100 common entries were solved in a live session between the PIs.

All eight research assistants who coded entries first had to process the training set, then we released the PIs' coding so the research assistants could compare their responses. Finally, we held a live session with at least one PI to discuss the disagreements.

Allocation of entries to research assistants and coding procedures

We created eleven sets of 800 entries that were coded by one of the eight research assistants. The research assistants followed the same coding procedure from the PIs: they first coded the recommendation (i.e., either "include", "exclude", or "uncertain") and then they coded the three inclusion criteria from Section 1.1. Research assistants also completed a log file for each batch of coded studies, allowing the PIs to keep track of the time that took them, on average, to code each study. There were some repeated entries across the sets that allowed us to validate the responses across coders.

Processing of research assistants' responses and definition of the final set for data extraction

The PIs gathered responses from all the research assistants, processed the "I don't know" entries, eliminated duplicates, and narrowed the remaining set to 839 entries. This set still included entries in which at most one of the inclusion criteria (i.e., the environment, education, and intervention variables) could have been marked as zero. The purpose was to keep studies that could relate to climate change education (or environmental education, broadly speaking) but did not include an intervention, and

education interventions in which the team needed a further discussion to determine whether the outcome was truly about the environment.

Entries were reassigned to research assistants, who finally selected studies with clear and measured environmental education interventions (i.e., by reassessing the three content criteria), reporting results, and published as articles or working papers. A final set of 151 studies was defined for the data extraction procedure.

Final exclusion during the data extraction and data validation procedures

During the data extraction process, the research assistants reported 20 studies that, after a thorough read, did not meet the criteria. With the remaining 131 studies, the PIs performed a validation check to center the analysis on social norms and observed behavior among the target population of CCE (i.e., students). As a result of this validation, we eliminated another 45 studies that did not measure any outcome related to norms or observed behavior, and studies in which the target population did not include students from any age (i.e., between preschool and master's education). Most of the analysis shown below pertains to the final sample of 86 studies.

1.3. Data extraction

We employed a predefined data extraction tool programmed in the online service Kobo. It included 41 items to be extracted from each study. The list of items is presented in Appendix A.2. This tool is divided into the following sections:

- A section to identify the entry. It includes basic information on the publication's title, first author, and publication type. In addition, it specifies information on whether some funding agency, as well as if there is any conflict of interest.
- A section on the context of the intervention, including the country and sector (in case it involved a sector additional to education).
- A section with the intervention information, with details on the data collection strategy, the target of the intervention, as well as its length and personnel implementing the program.
- A section on the reported outcome, including whether it was a normative or a behavioral outcome, and some characteristics about its measurement.
- A section on the reported results.

In addition to the data extraction tool, we also included a module on the critical appraisal of the selected studies (see Appendix A.3). Our critical appraisal tool borrows and adapts elements from Ategeka et al.'s protocol (2022) and the National Institute for Health and Care Excellence (Great Britain)'s protocol (2012). This module is also helpful to shed light on methodological differences, since we refer to "impact evaluations" in the protocol, though we explained to the coders that other qualitative analyses may have been included.

The first section assesses the quality of the protocol: whether the protocol and measures were clearly defined, and whether selection was random, well-described, and representative. The second section refers to the quality of the intervention's implementation. It assesses whether participants were comparable before the intervention, and if there were contamination issues or differences in how participants were treated within a given condition. The third section pertains to the quality of the statistical analysis. It assesses the appropriateness of analytical methods, and the quality in the report of the effects of the intervention, sample size, and presence of a power calculation.

The PIs then downloaded the information entered in the Kobo tool as a spreadsheet. They double-checked the main variables of interest and proceeded with the analysis, which we report in the next section.

2. Results

In this section, we report the main results of our literature review. We will first cover basic information, such as when the studies were published, where the research was conducted, which population they target, and which empirical methods they use to collect the data. For this initial analysis, we will not yet apply the final inclusion criteria regarding norm measures. This means that we will also consider papers that measure knowledge or psychological constructs, as well as those with self-reports of behavior or intentions to act. We take this approach because we believe it is valuable as a first step to provide a general overview of the broad empirical literature on CCE before finally focusing on the final set of studies that focus on norms. We will then restrict our analysis to studies that clearly measure injunctive or descriptive norms. For these studies, we will examine how social norms are measured and which educational approaches they use. The aim is to identify possible strengths, limitations and gaps that could guide future research on climate change education. Throughout the section, we will also compare studies conducted in developing countries with studies conducted in developed countries to identify possible differences in the implementation and efficacy of CCE between these countries.

Most of the retained studies focus on mitigation and only a minority on adaptation or conservation. The distinction between mitigation and adaptation is not always clear-cut, and often the two terms are interrelated. Moreover, context matters for the definition of mitigation and adaptation, and this is not always specified in the studies analyzed. For example, reducing water or energy consumption are ways to mitigate the impact of a household on the environment. However, if they are done in the context of a drought or an energy shortage, they become adaptation. Considering these limitations, which ultimately reduced the number of studies attributable to an adaptation category, we will largely set aside this broad classification for the rest of the analysis, and, instead, focus on describing the specific behavior that a given intervention aims to change or maintain. Regarding conservation, we reviewed more than ten studies describing interventions that involved experiences with nature, either by learning about other species (mostly visits to zoos with an educational component) or outdoor trips. Nevertheless, only one study met our final criteria of having an outcome of norm elicitation or observed behavior.

2.1. Quality check: exclusion of misclassified entries

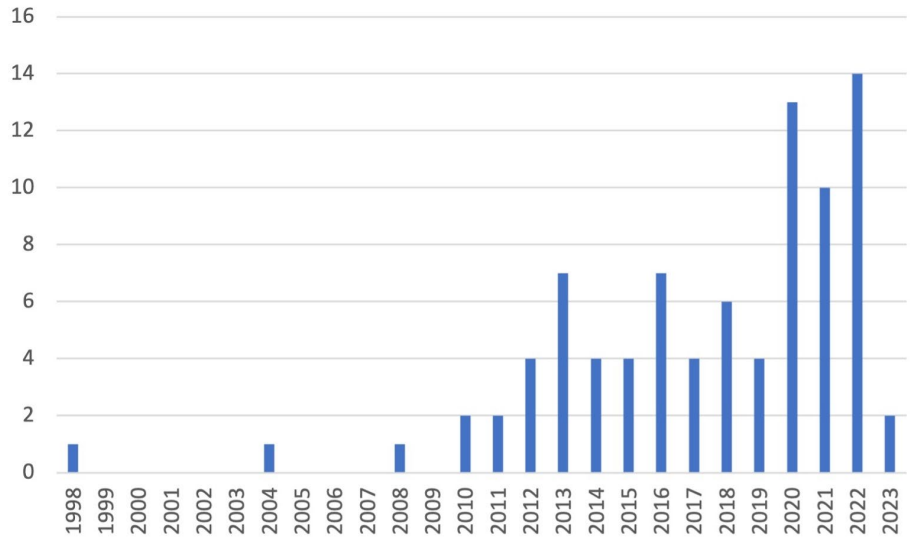
All 131 articles on which data were extracted by our research team were double-checked by one of the authors involved in this study. We did this to provide a quality check of the data extraction and to identify potential mistakes or misclassified studies. Since the original extraction rules were intentionally formulated such as to err rather on inclusion than exclusion, we identified few studies that were wrongly classified. This concerned specifically studies on relevant topics that did not present a clear intervention or studies that were primarily directed at an adult population. In the following, we will continue with an analysis of the remaining 86 studies that survived the quality check, the earliest of which was published in 1998. Of these 86 studies, 85 are written in English, and one in Italian. Figure 2 illustrates the rise of scientific interest in empirical CCE research over the last two decades.

2.2. Geographic distribution

In this section, we give an overview of where the CCE research included in our literature review was conducted. We present this information at the country level, as it is available for all the studies included

in our literature review.[8] There are studies that evaluate multiple interventions or the same intervention in different countries. For these studies, we consider all the countries involved in the research, meaning that each of these studies provides more than one geographical observation (see below), making the total number of observations equal to 103 instead of 86. If a paper concerns different interventions within the same country, this counts as one data point for that country. Figure 3 displays the distribution of CCE research across the world. It shows that CCE interventions are geographically widespread across many countries. Most interventions are concentrated in the United States (20) and in the People's Republic of China (12). If we look at the distribution by continent, Asia has the highest density of CCE research (35), followed by Europe (28), North America (26), Africa (6), Oceania (5) and South America (3).

Figure 2. Year of publication of the 86 retained studies. Studies in 2023 included January-April.



Source: Authors' computation. Original.

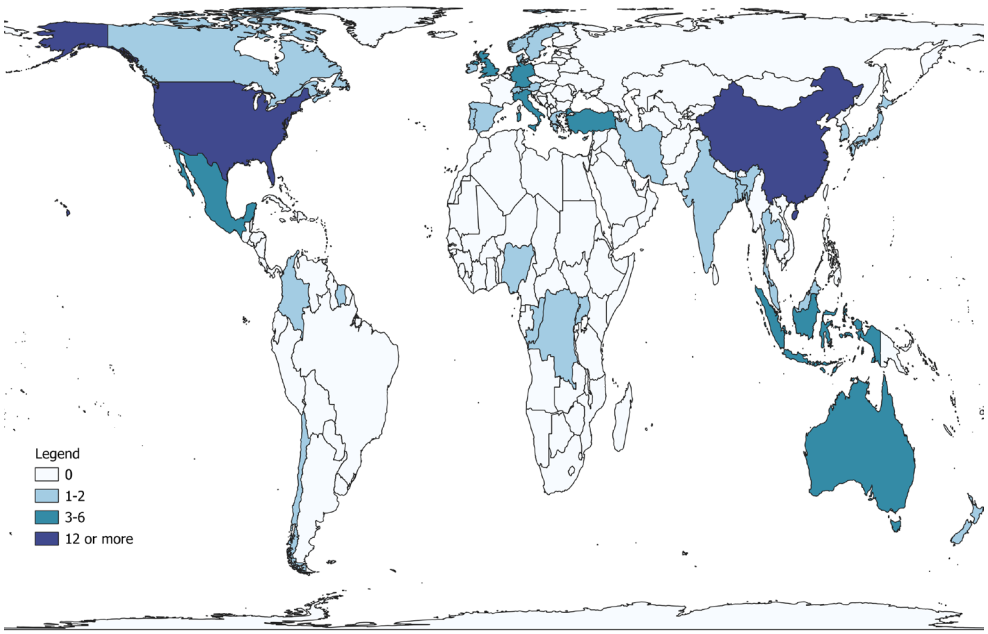
We can also investigate how CCE research is distributed between developing and developed countries. We rely on the World Bank's income classification to determine a country's level of development.[9] This classification is based on the Gross National Income (GNI) per capita of each country and provides four income groups: low, lower-middle, upper-middle, and high. Following the World-Bank's designation of developing countries, we further classify as developing countries those with an upper-middle, lower-middle, or low income. Countries with a high income are designated as developed countries. Table 1 shows the results of this classification. Most of the research is conducted

[8] One study (Baird et al., 2022) only indicates the geographical macroregion (Scandinavia) of one of the places where the educational program took place. According to the Encyclopædia Britannica, Scandinavia is typically said to consist of Denmark, Norway and Sweden (https://www.britannica.com/place/Scandinavia, accessed on the 13th of September 2023). We thus use these countries as the geographical indicator for this study. Throughout the paper, by 'country' we mean "any territory for which authorities report separate social or economic statistics"
(World Bank; https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries, accessed on the 13th of September 2023).

[9] For the analysis, we use the most recent data available from the World Bank's website. We downloaded these data on the 13th of September 2023.

in high income countries (64), and only a minority in lower-middle and low income countries (9), revealing a disproportionate focus of CCE research on developed countries, which is a significant limitation of the current CCE research.[10] In the rest of the analysis, we will compare each time studies conducted in developed countries (high income) with those conducted in developing countries (low and middle income) to identify potential knowledge gaps.[11]

Figure 3. Distribution of CCE research across the world



Source: Authors' computation. Original.

Table 1. Distribution of CCE research per income group and level of development

	WORLD BANK INCOME GROUP	CCE RESEARCH ENTRIES
DEVELOPING	Low income	3
	Lower-middle income	6
	Upper-middle income	30
DEVELOPED	High income	64

Source: Authors' computation. Original.

[10] There might be other papers — written in a language different from English, French, Spanish, German and Italian — that focus on developing countries. However, these studies, if existing, are not included in our literature review because they did not meet our inclusion criteria in terms of language.

[11] For studies with multiple countries, we will, from now on, consider the level of development that is most representative in the data. This simplification only affects five articles and does not change our conclusions. One article only considers developing countries (Kuhar et al., 2012) while a second article only considers developed countries (Deisenrieder et al., 2020). In the other three articles, most of the data (86%, 69% and 94%, respectively) come from developed countries (Arya and Maul, 2016, Sidiropoulos, 2018; Baird et al., 2022).

Along the lines of the geographical distribution of the interventions and the corresponding countries' income, we also shed light on the distribution of the reported funding type. Table 2 reveals that the most commonly reported funding source are government's agencies (either from the national or local levels) with 33.7%, followed by academic institutions with 26.7%, and finally those funded by a charitable, non-governmental, or private organization with 10.5%. Note also that in 29.1% of the studies the funding source was not specified. In comparative terms, funding from academic institutions is more common in high income countries than in countries classified as upper-middle income or below, whereas funding from charitable, non-governmental and private organizations is more frequent in the latter than in the former. Nevertheless, this difference is not statistically significant (p-value from a Chi-squared test is 0.432).

Table 2. Distribution of CCE research's type of funding institutions by country's income levels

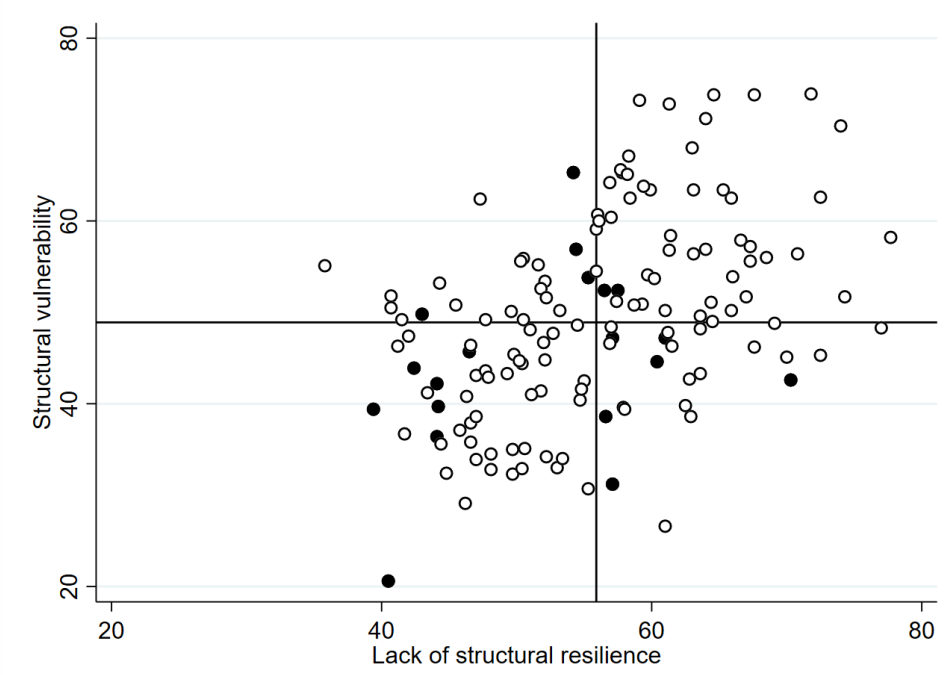
	Government Agency	Academic Institution	Charitable-Governmental, Private Organization	Not Specified
High income	17 (31.5%)	17 (31.5%)	4 (7.4%)	16 (29.6%)
Upper-middle income or below	12 (37.5%)	6 (18.8%)	5 (15.6%)	9 (28.1%)
Total	29 (33.7%)	23 (26.7%)	9 (10.5%)	25 (29.1%)

Source: Authors' computation. Original.

We can also compare the countries targeted by CCE research according to their vulnerability to climate change. Vulnerability is negatively associated to the degree of development of a country, with less developed countries more at risk of being affected by the negative consequences of climate change. For this analysis, we rely on indicators provided by the United Nations and the Notre Dame Environmental Change Initiative (ND-GAIN) of the University of Notre Dame. The United Nations provides a multidimensional vulnerability Index that captures both economic and ecological vulnerability. It is based on two domains of vulnerability: (i) structural vulnerability, which reflects the extent to which a country is exposed to external shocks and stressors, and (ii) structural resilience, which expresses the capacity of a country to resist and adapt to these shocks and stressors (United Nations, 2023).

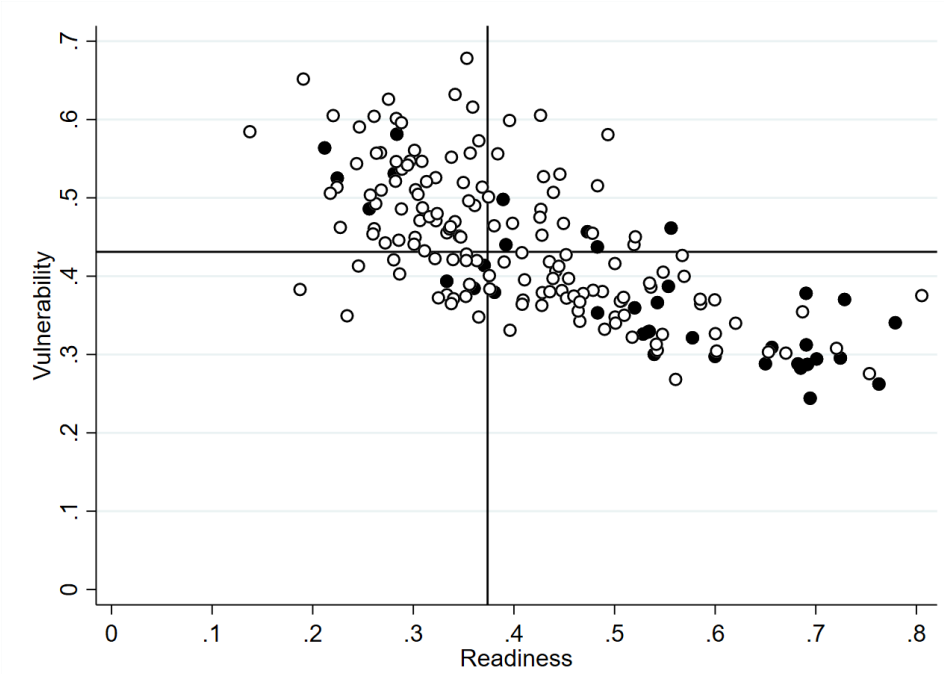
At the time of this study, data on structural vulnerability and (lack of) structural resilience were available only for a set of countries (mostly developing countries). A precise description of how structural vulnerability and structural resilience are measured can be found in the United Nations's report (see United Nations, 2023). In Figure 4, we use a scatter plot to illustrate how structural vulnerability and (lack of) structural resilience vary across countries for which data are available. Black dots identify the countries that were targeted by CCE research. The horizontal and vertical lines represent median values. The figure reveals that the majority of CCE interventions are concentrated in the lower left quadrant, that is, in countries with low levels of vulnerability and high levels of resilience. CCE interventions in countries with high levels of vulnerability and low levels of resilience (upper left quadrant) are almost nonexistent. We replicate this analysis using the data from the ND-GAIN. In this case, we rely on two indices: a measure of vulnerability to climate change and a measure of readiness in leveraging investments for adaptation. For details on how these indices are constructed, see the ND-GAIN (2023)'s technical report. An advantage of using the data from the ND-GAIN is that the information about vulnerability and readiness is available for 182 UN countries, including high-income countries, from 1995 until 2021. We can therefore have a more comprehensive view of how the countries targeted by CCE research differ in terms of vulnerability and resilience. We use the data from 2021. Figure 5 shows how vulnerability and readiness vary across countries. Countries targeted by CCE research are marked with black dots. The horizontal (vertical) line represents the median score of vulnerability (readiness) across all countries and years. Most CCE interventions are conducted in countries with few climate challenges and well positioned to withstand climate change (lower right quadrant). In the other quadrants, there are much fewer observations. In particular, there is a general lack of studies in countries that are not (yet) well positioned to adapt to climate change and that are probably most in need of education to develop adaptive capacity. This is an important limitation of the literature on CCE.

Figure 4. Distribution of CCE research by vulnerability and lack of resilience



Source: United Nations. Data retrieved on October 17, 2023

Figure 5. Distribution of CCE research by vulnerability and readiness



Source: Notre Dame Environmental Change Initiative. Data retrieved on October 17, 2023

Before moving to the next section, we can also check how many studies are conducted in multiple countries and if they target at the same time developing and developed countries. Only 5 out of 86 studies are conducted in multiple countries, which indicates a lack of direct cross-country comparison in CCE research. Kuhar et al. (2012) examine conservation education programs conducted in four developing countries in Africa (Nigeria, Republic of Congo, Democratic Republic of Congo and Uganda). The paper, however, does not compare the effectiveness of these programs across the four locations. It also provides only limited information about the participants in the study. Arya and Maul (2016) assess an educational program designed to encourage secondary students to read scientific studies and discuss them with peers and scientists. The program was conducted in several countries (United States, China, New Zealand, Norway). However, the authors collected only a limited number of observations per country (between 12 and 30 students per country) and did not make any cross-country comparison. No cross-country comparison is also present in Baird et al. (2022). The authors examine the impact of an outdoor nature-experience educational program conducted in different areas of the United States, Canada, Mexico, Chile, Scandinavia, New Zealand, and India. Most participants (276 out of 295; 94%) however came from the US, which severely restricted the possibility to perform any meaningful comparisons across countries. Deisenrieder et al. (2020) examine the impact of a participatory CCE program and the effects of participating in climate protests on 14-year-old students' climate change awareness. The study was conducted in Southern Germany and Austria. The data from these two countries were pooled in the analysis. The only study that provides some comparisons across different geographic contexts is Sidiropoulos (2018). The author examines the influence of sustainability education on tertiary students' worldviews, attitudes, and self-reported behavior towards sustainability. The students were located in two high income countries (Australia and Italy) and one upper-middle income country (Malaysia). Students also differ in several other attributes (e.g., discipline of study), which limited the comparability across the three countries. Moreover, most of the observations (86%) came from only one country (Australia).

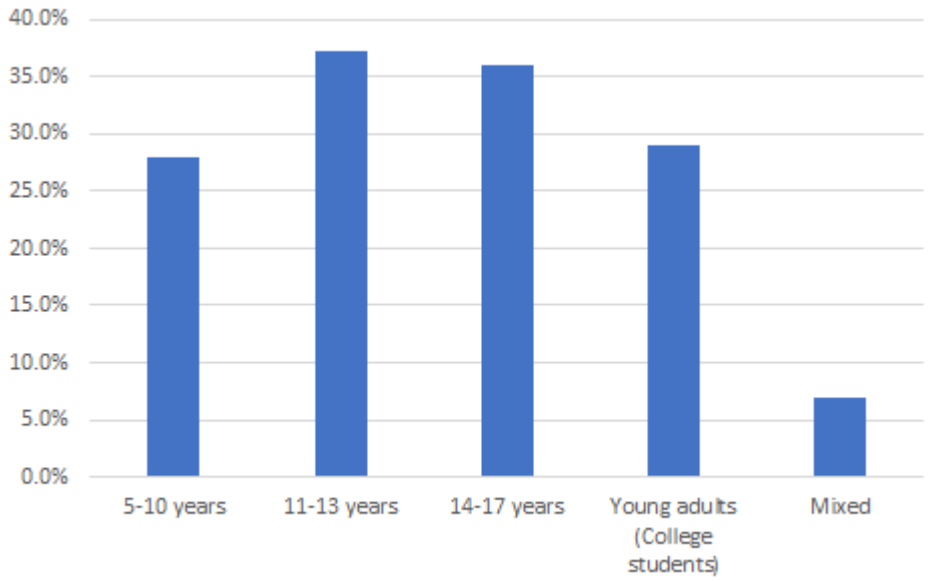
2.3. Demographics: age ranges and targeting of urban and rural populations

This section covers the targeted ages of CCE research. We have divided ages into four large categories, aiming to map students at primary school, middle school, high school, and university. Note that, by our definition of CCE, we do not cover those interventions only targeting adults beyond university. We have a fifth category labeled as "mixed," and it groups studies where both students and the institution's personnel (usually teachers) are targeted in the intervention.

Figure 6 depicts the distribution of studies by age range according to the categories described above. We find that 32 (37.2%) of studies target children between 11 and 13 years old, followed by studies targeting children in the age range 14-17, with 31 entries (36%). Studies with college students (age range 18-25) account for 29.1% of the observations (25 entries in total), while the younger age group for which we have any registry, aged 3-10, was targeted in 24 studies (27.9%). Finally, the category mixed, involving at least one of the former groups and adults beyond college age, has 6 studies (7%). We thus argue that there is a relative balance among the age groups targeted in the reviewed CCE programs involving children and young adults.

Note that these percentages sum above 100%. The reason is that a large amount of the interventions involved more than one age group. This criterion applies to the four most frequent categories, while "Mixed" is mutually exclusive with these four and corresponds to the particular case where the intervention also involved adults. One example is Kuhar et al. (2012), a study describing a joint initiative between the Pan African Sanctuary Alliance (PASA) and Disney's Animal Kingdom. The study describes how personnel from 18 sanctuaries attended an education workshop. These sanctuaries could then obtain funding from Disney's Animal Kingdom to develop their own education programs. In exchange, they had to include a pre-post evaluation in their program. Only five Sanctuaries conducted formal pre-post evaluations. Participants were primary school students, secondary school students and community adults.

Figure 6. Age ranges covered in the 86 retained studies



Source : Authors' computation. Original.

As additional information regarding the target population, we also coded whether the CCE program aimed at urban or rural population, or both. From the 86 studies retained in the analysis, this information could be retrieved for 50 of them (58%). Within this subsample of 50 studies, 66% of them targeted urban students, 22% a mix of urban and rural students, while 12% only covered rural students.

The few studies exclusively targeting the rural students included, for instance, an environment-based school program conducted in the Darajat protected forest area in Indonesia (Rosmaladewi and Poetri, 2020). This program included the establishment of an "environmental ambassador" in each class – i.e. a student who was then trained in waste management, hygiene, and sanitation. Another example is Kendall et al. (2021), who reported a program involving teachers and students in the construction of fuel-efficient stoves in communities nearby the Kibale National Park in Uganda. The objective of the intervention was to reduce encroachment in a chimpanzees' critical habitat.

Another example of environmental education in the surroundings of a natural park is Feilen et al. (2018). The study evaluates a program implemented in six communities in the surroundings of Los Colorados National Park and Los Titíes Biological Reserve, habitats of the cotton-top tamarins. These interventions suggest that, when CCE programs are exclusively targeted at rural areas, they have the direct objective of improving the interactions of the community with the surrounding species, either directly, as in the two examples of primate conservation, or indirectly, by improving hygiene conditions and waste management to protect the forests.

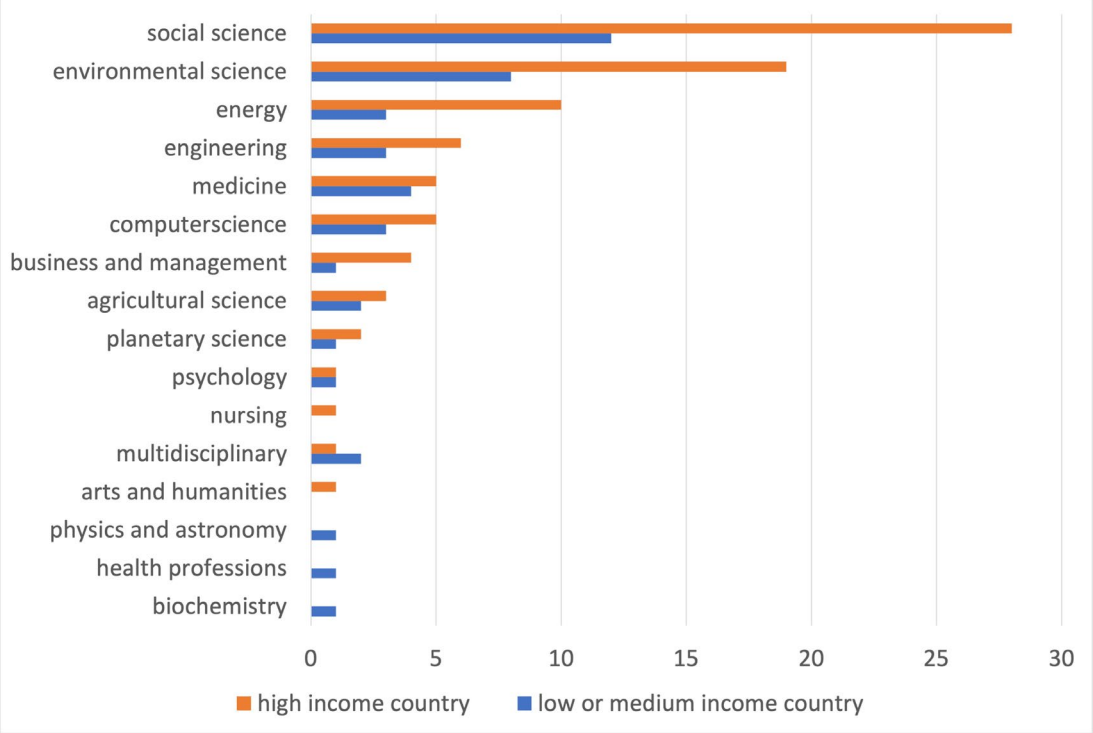
2.4. Methods

The studies included in our literature review come from researchers across a wide range of disciplines. Figure 7 provides an overview of all the listed disciplinary fields attributed to the 69 papers (out of 86) that were published in an indexed scientific journal. Most papers are attributed to more than one field. Not surprisingly, the majority of papers are published in journals related to social sciences (and more specifically education), and in journals related to environmental sciences, energy and engineering. We observe little differences in the distribution of papers across disciplines between high income countries or low/medium income countries, as classified above.

To evaluate the quality of the published papers, we also categorized the quartile ranking according to SCImago. We observe that most papers are published in journals that can be considered in the first (44 out of 69) or second quartile (18 out of 69) of their first listed disciplinary field.

As anticipated, the observed methods regarding measures of the impact of CCE vary greatly. This is partly due to different standards across disciplines and the methodological challenges of obtaining reliable measures of lasting behavioral change. Since the behaviors addressed through CCE are often behaviors that might be influenced by predispositions, previous knowledge, or attitudes, it is crucial for reliable data to have a clear baseline or comparison. We observe that many studies (69 out of 86) employ a pre-post approach. This implies a measurement administered before the behavioral intervention to provide a baseline measure. This baseline can then be compared to a post-treatment evaluation.

Figure 7. Distribution across academic fields (following SCImago journal classification) and country of concern of the 69 indexed papers



Source: Authors' computation. Original.

Post-treatment evaluations bear the risk of bias if demand effects are present. For example, a questionnaire administered right after a training course on recycling, might lead to higher self-reported intentions to do so due to social pressure or a desire to please the teachers or the researchers. Some studies try to address this problem by repeating post-treatment measures at different time intervals (e.g., Feilen et al. 2018). These repetitions might indeed reduce demand effects to some degree. However, in the case of very specific programs the reminder might again trigger similar effects as initially experienced. Some studies that include repeated measures observe indeed weaker effects during re-testing. However, it is unclear to which degree repetition can actually control for demand effects.

An alternative approach to mitigate demand effects is to allocate the participants to a treatment and a control group. We observe that 31 out of 86 studies report results from either a non-treated control group or compare multiple treatment interventions. It should be noted however that the pure

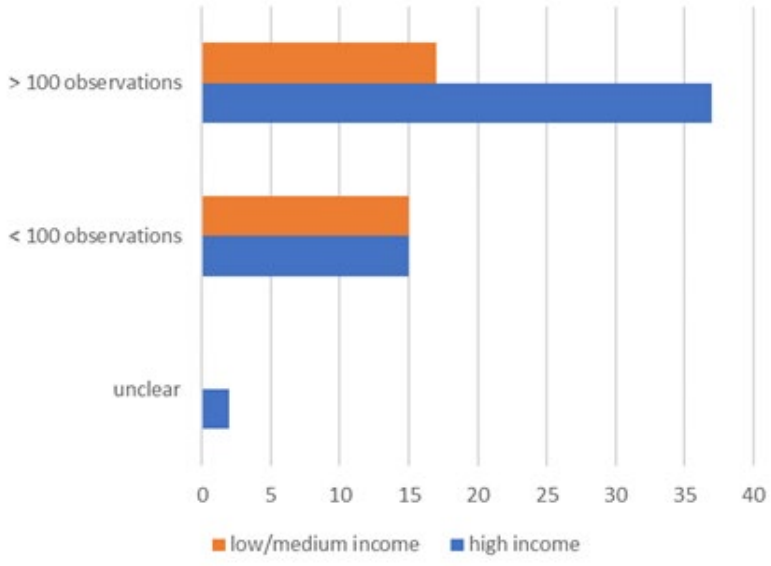
existence of a control treatment is not necessarily a guarantee for higher data reliability. Notably, the allocation should be done randomly to avoid selection effects. However, in many cases the allocation to treatment and control is non-random (e.g., because treatment participants volunteered to participate in an environmental program or because researchers assign the treatment and control groups to very different classes). Another important factor to consider is that treatment and control groups should not interact to prevent spill-over effects from one group to the other. In many cases however teachers 'treat' randomly some of their classes in the same school (e.g., Tarng et al., 2015). In such cases, we cannot exclude the possibility of interactions occurring between students, which could potentially influence the observed effects.

Although we know that 69 studies employed a pre-post approach and 31 had a treatment and control group, their cross tabulation is also helpful. There are 12 studies that do not include any of these features in their design. Hence, they are purely observational or descriptive in their content. On the other hand, there are 26 studies combining both methodologies, meaning that they can perform a better comparison that takes into account differences in trends between those receiving the intervention and those who do not.

Data quality is not solely dependent on valid measurement but also on the availability of sufficient observations to draw valid conclusions. We classified the number of reported observations in the studied papers. We observe that while 54 studies report more than 100 observations, 30 report less than 100 observations (with 2 studies not providing enough information to judge the number of observations). The quality analysis of the studies highlights the importance of a careful evaluation of the different statistical analyses. Specifically, educational interventions are often administered on a school or group level. However, data, especially when recorded through individual questionnaires and self-reports, is often analyzed on an individual level. Larger numbers of observations reported in the papers might thus sometimes be misleading in promising greater data reliability.

We also investigated the reported numbers of observations separately for the countries categorized as either low/medium income or high income (see section 2.2). Figure 8 presents the distribution according to the categories discussed above and we observe no differences across the two types of country categories.

Figure 8. Reported observations across studies and country of concern



Source: Authors' computation. Original.

2.5. Coordination efforts: are the reviewed studies part of larger education initiatives?

We find that 46 out of the 86 selected studies were performed without any type of coordination beyond a single educational institution (either school or university): they were not part of any local or national initiative, and neither explicitly mentioned that their CCE intervention derived from a known and more general curriculum transformation guideline.

We have another 18 studies in which the coordination efforts come from the researchers conducting the study without any steering national policy. They were implementing the CCE intervention or collecting data from students, often designed by the same team, in different institutions. These coordination efforts involve at least two schools, though we also found interventions with more complex coordination efforts that involve up to 13 and 16 schools (Seybold et al., 2014; Mikami et al., 2022).

We divided the remaining 22 studies, which are the main focus of this section, into three groups. First, we have eight studies that explicitly mention that the implementation was coordinated with at least one governmental organization, either at the local or national level, plus one case in which the organization conducting the CCE intervention is directly funded by a national institution (i.e., the National Science Foundation provided the grant for MADE CLEAR, the Maryland and Delaware Climate Change Education Assessment and Research, hosting the Climate Academy teacher program detailed in Shea et al., 2016). In this group, we highlight a study that is perhaps the largest coordination effort in the reviewed studies in terms of national reach: Somwaru (2016) reports how 86% of the schools in Suriname adopted a "Green School" model encouraging more practical activities in education for sustainability. Although the study describes the implementation in 286 schools, it acknowledges that the monitoring of cleaning and waste-related outcomes was done for 76 schools (which remains the largest number in our review). The other initiative at the national level is named k.i.d.Z.21 and corresponds to the Austrian Climate Change education initiative, monitored in 8 schools of Austria and Bavaria (Deisenrieder et al., 2020). There are other four initiatives coordinated with local governments in Portugal (Rocher et al., 2020; Rolim and Baptista, 2021) and Italy (De Dominicis et al., 2017; Pietrapertosa et al., 2021). The remaining one is an afterschool program coordinated with the city of Sacramento (California) to teach about water management (Bird and Subramaniam, 2022).

The second group concerns coordination efforts whose focal agency is a private institution rather than a government. From the eleven studies in this group, five are directly related to wildlife conservation programs, either as collaborations between foreign or domestic zoos and specific communities (Kuhar et al., 2012; Feilen et al., 2018; Kendall et al., 2021), or as wildlife clubs or trusts (Damerell et al., 2013; LaCombe and Danoff-Burg, 2013). The remaining six are quite different in how they connect with the educational system, though most of them are non-profit organizations. For instance, Sarrasin et al. (2022) explore the connection between pro-environmental behaviors and climate anxiety registered in a set of activities called "Youth Climathon." The study was made in partnership with the non-profit company organizing these Climathon events. The Gould League, a non-profit Australian organization, evaluated in a couple of schools the effects of a program named Waste Wise (Armstrong and Grant, 2004). Another couple of studies coordinated through non-profit groups target students in the United States. First, the Alliance for Climate Education program evaluated the engagement and concern with climate change across 49 schools and more than 1,200 students (Flora et al., 2014). Second, a study using stickers to promote laundry with cold water and turning classrooms lights off was based on a project from the Urban Sustainability Directors Network, a network founded by and for sustainability practitioners working on sustainability at the community level (Frantz et al., 2016).

We have three remaining studies in the last group. They correspond to implementations that followed a national or local program or guideline on education for sustainability but, based on the study's description of the implementation, seem to be an uncoordinated effort with the program provider. The three initiatives come from Asia. In Indonesia, one intervention follows the Adiwiyata School Program, a key program of the Ministry of Environment aimed at promoting knowledge and awareness about environmental conservation efforts among students (Sulistyarini et al., 2022). The other intervention in

Indonesia is documented in Lestari et al. (2022), who implemented a learning model called RADEC (read-answer-discuss-explain-create), based on the ESD guidelines. In the remaining example, the authors declare to follow the "National energy talent cultivation/integration-establishment of experiential curriculum" in their design of hands-on teaching materials (Hsu and Shih, 2014).

On the aggregate, our categorization exercise indicates that only in 22% of the studies (19 in total) there is a report of a coordination effort with a national or local government agency, or with a non-profit institution promoting or implementing CCE or ESD interventions. In this group, collaborations with government agencies, especially locally, tend to occur more often in Europe. Collaborations led through private efforts are often connected with conservation programs, including those from zoos and their domestic and foreign activities. There are also a couple of examples where non-profit organizations conduct this type of interventions in the United States. In light of the few reported coordination efforts, a crucial aspect to point out is that they often come from schools but not from universities. Funding incentives may need to be revised to encourage such collaborations at this level.

2.6. Studies that measure injunctive and descriptive norms

In line with the objective of this systematic review, we will now shift our focus to studies that measure either injunctive or descriptive norms. Studies on injunctive norms investigate what individuals think others (dis)approve of. Instead, studies on descriptive norms investigate what people do or what people think others do. We thus do not consider studies that solely measure knowledge, intentions, attitudes, or psychological constructs (like self-efficacy). We also exclude papers that include questions that could be interpreted as related to injunctive or descriptive norms, but that were included in some larger scales or constructs and not separately analyzed. Finally, we exclude studies only measuring personal norms or views with no link to social norms.

A first observation is that only a limited number of studies are concerned with measuring the effects of CCE interventions on real behavior and norm-related beliefs, despite one of the primary objectives of CCE is to change people's normative opinions and motivate them to take concrete actions against climate change. In particular, only 18 out of 86 studies measure the effects of a CCE intervention on norms. 3 of these studies report clear injunctive norms, while 16 studies report clear descriptive norms by observing actual behaviors or by measuring what people think others do (one study analyzes both descriptive and injunctive norms). Our inclusion of observed behaviors into the analysis of descriptive norms gives us some comprehension on studies that crossed the intentions-actions gap. Nonetheless, their interpretation comes with a caveat: behavioral change is aligned with norm shifting, but we cannot tell whether such changes are taking effect in a sufficiently large population to currently alter the norm. In other words, learning about CCE interventions that steer behaviors is a first step for norm shifting.

We will now present a description of these 18 studies. Each time we will first mention a study, we will add, in brackets and italics, a brief description of the intervention. We first consider studies measuring injunctive norms. In the social sciences, injunctive norms are often measured through questionnaires that are not incentivized. There are also incentivized methods for identifying injunctive norms (see, e.g., Bicchieri and Xiao, 2009; Krupka and Weber, 2013).[12] None of the studies that we identified in our literature review collect incentivized measures of injunctive norms. The use of non-incentivized measures is in itself not a limitation, but it could be useful to complement these measures with incentivized elicitation methods to encourage more precise and neutral answers.

[12] Bicchieri and Xiao (2009) and Krupka and Weber (2013) propose two different methods to elicit norms by providing incentives on belief accuracy. In the former, individuals are paid based on the accuracy of their beliefs regarding what others believe is appropriate or inappropriate behavior. In the latter, individuals play a pure coordination game where they are paid if they match the responses of others regarding what is considered as socially appropriate or inappropriate.

Two of the three studies that measure injunctive norms rely on survey answers to questions of the type: 'My friends and family supported me in adopting environmental protection behaviors' (Tsai and Tan, 2022) [*Lectures followed by hands-on sorting of recyclables*] or 'The pupils from my school find it important that I eat little meat' (Jans et al., 2023) [*Vegan cooking workshops*]. Participants could answer these questions on multiple-point Likert scales. Both studies find a positive effect of their respective CCE interventions on pro-environmental injunctive norms. These two articles also reveal that there may be substantial variation in the way the questions about injunctive norms are framed. For example, Tsai and Tan (2022) examine injunctive norms from the perspective of friends and family members, while Jans et al. (2023) consider the perspective of other pupils who were exposed to the same educational intervention as the respondent. This difference is crucial because the choice of the reference group in eliciting injunctive norms can profoundly affect how people respond to injunctive questions. Specifically, the closer the reference group is to the respondent's own situation (e.g., sharing similar circumstances or being exposed to the same intervention), the stronger the response regarding injunctive norms may be (see, e.g., Goldstein et al., 2008). The same principle applies when considering 'broad' reference groups with a prominent social identity (e.g., environmentally conscious individuals).

The third study involving injunctive norms is Zhang et al. (2019) [*Voluntary participation in a nature club*]. It explores the change in personal norms—defined as internalized social norms (the authors call them "subjective norms")[13]—regarding hunting birds and eating insects. The authors do not explain how they precisely measure these norms, which represents a limitation of this study. However, they acknowledge a moderate shift in norms following the intervention, despite the prevailing strong local cultural norms that discourage the protection of birds and insects. This is a noteworthy aspect, as it recognizes the role of cultural and social barriers in impeding the broader-scale shift in norms.

Turning to descriptive norms, these are either measured by observing actual behaviors or through questionnaires that elicit the beliefs of people about what others do. Beliefs about the behavior of others can be incentivized or not. We found only one paper that elicited descriptive beliefs via a non-incentivized survey question (Jans et al., 2023). In this study, children aged 10 to 17 across nine German schools (three urban, six rural) took part in a vegan cooking workshop session. There were 17 such sessions, where pupils cooked in small groups (52) ranging from 3 to 6 members. A questionnaire applied before and after the workshop included four items on pro-veg*n injunctive norms (where veg*n stands for vegetarian and vegan), plus another set of items on pro-veg*n descriptive norms. The questions on norms were elicited for two reference groups: the workshop session and the school. To elicit injunctive norms, students were asked to what extent they perceived other pupils from their workshop session/school to find it important: (i) to eat vegetarian, (ii) to eat vegan, (iii) to reduce their meat consumption, and (iv) that others find it important for the respondent to eat little meat. Descriptive norms were measured by asking pupils how often other pupils from their cooking group and school had/will have a vegetarian meal within the previous/upcoming 7 days. The descriptive questions were framed in terms of intensive margin (e.g., how often others engage in a given behavior) rather than extensive margin (e.g., what percentage of others engage in a given behavior). An open question is whether the authors opted for the intensive margin due to an expectation of lower numeracy, particularly among younger individuals, or if it reflected the researchers' preferences for measuring norms. If the latter is the case, it suggests a potential avenue for further research in this direction. The authors find that the cooking workshop strengthened pupils' injunctive and descriptive pro-veg*n norms. This result is consistent with the findings of a systematic review by Dudley et al. (2015) on factors that promote healthy eating in schools. Dudley et al. (2015) highlights how experiential learning strategies are the most effective for improving eating habits.

We detected more articles measuring real behavior. Observing behaviors involves some methodological challenges, because behaviors must be quantified, attributed to the agent causing it, and observation should not be evident to avoid the above discussed problem of social desirability biases.

[13] Even though the authors discuss mention "subjective norms" and not social norms, we kept this study in our SR for being a borderline case that nourishes the few examples of direct norm measurement.

We observed a range of different approaches across the 16 studies that measure descriptive norms using real behavior. We group them into different categories that we describe below and list in Table 3.

Recycling behavior & clean-up

The majority of the studies involve some sort of recycling behavior. Recycling and cleaning up trash from the environment can be seen as an element of climate change mitigation. Recycling saves energy and prevents extraction of raw materials. It also helps reduce waste, and therefore pollution. Recycling thus reduces the environmental impact, contributing to a "green" mentality.

Due to regulation or pedagogic purposes, schools often need to have a formal waste management plan. The visible infrastructure of these plans (e.g., multiple bins and any larger separate storage facilities) makes it possible to increase adolescents' knowledge of what is and what is not recyclable. This is a first step in using adolescents to promote recycling to adults and reinforce their globally minded and future-oriented attitudes toward sustainability (Prestin and Pearce, 2010). Hence, recycling is an instantaneous behavior that allow students to take action and also convenient for the researcher (i.e., data collection strategies are relatively simple to implement).

Recycling behavior can be easily observed and in studies experimentally manipulated, for example by handing out to participants a specific item that will be thrown away. It is then relatively easy for the researchers to count the number of times the item was correctly recycled or not. This approach is taken by Donmez-Turan and Kiliclar (2021) [*Environmental training with or without incentives*]. In this study, students were presented with a 25-minute environmental training provided by the researchers to change students' knowledge about the environment. After the training, students were handed a tool with a recyclable package. The outcome variable in this study was whether the package was indeed correctly disposed in the recycle bin. Across different treatments (control and reward) the effect of the intervention was thus observed. Arya and Maul (2016) [*Exchange program focused on climate change*] presented participants with a piece of paper that they were asked to dispose of. It was then counted how many participants correctly disposed the paper in the recycling versus the regular bin. LaCombe and Danoff-Burg (2013) [*Lecture on trash disposal part of a larger EE program*] gave participants as part of the reward for their participation a wrapped candy and counted how many participants in the treatment versus control group correctly threw their candy wrapper in the appropriate trash bin.

Other studies do not focus on specific acts of recycling but rather attempt to measure the general rate of recycling for a school or class. This avoids the "demand effect" that might be experienced by a participant that is asked to recycle after a lecture on the topic but has the disadvantage of not providing exact data on individuals due to measurement issues. Thus, it is more feasible to explore this type of general behavior at the level of the intervention (e.g., the class or school) with repeated measures of the outcome of interest. Nourmoradi et al. (2022) [*Four weeks educational program*] observe behavior in school measured through checklist by the researchers. These lists involved items related to the use of recycling bins, preserving green space, keeping the school clean, etc. The data analysis from this paper however illustrates the problem of independence of observations: while the intervention was at a school level, data are mainly presented from questionnaires and self-reports on the individual level. García-Vásquez et al. (2021) [*Role-playing with Spanish children for awareness of child miners in the DRC—Democratic Republic of the Congo*] take a more methodologically sound approach. The treatment in this case concerned interventions at the class level. The behavioral outcomes concerned intentions regarding buying a new phone but also recycling of old (Post-First-Life) mobiles in three weeks following the intervention. For this, after the interventions, boxes were left in the classes for collecting Post-First-Life mobile phones for recycling. The boxes were opened after three weeks, mobiles were counted and taken to civic amenities for reuse or recycling.

An alternative approach to measure recycling and cleaning behavior is observing participation in specific activities aimed at this. Druen and Zawadzki (2021) [*Social dilemma simulation: "The Climate Trap"*] study, for example, the effect of a simulation exercise on various behaviors, one of which concerned volunteering to a campus clean-up activity and the hours spend on doing the activity.

Reduction of energy and electricity use

One of the most impactful approaches to reduce our impact on the climate is through saving energy and thus electricity. Reducing electricity use is thus a very prominent mitigation approach.[14] We observe multiple studies that focus on electricity use. Energy use can be objectively measured through meters, but it is a very difficult task to match administrative data of households' energy consumption with educational interventions in which children from these households effectively took part.[15] Even if it happens, household participation may lead to self-selection issues (e.g., parents more eager to report their address or consent the use of their data might also be more involved in their children's education outside the school, inflating the treatment effects).

Given these difficulties, in most cases, the researchers rely on observing the behavior that leads to the energy saving (e.g., turning off the lights, selecting cold washing cycles). For example, Frantz et al. (2016) [*Stickers and posters on energy saving in the school's laundry rooms and classrooms*] studied behavior by students in different situations. On the one hand, the types of washing cycles in machines in three large dormitories on campus were recorded and on the other hand it was observed if lights in classrooms were turned off. Similarly, Mattsson and Laike (2022) [*Board and digital games followed by sticker reminders of energy use*] had data loggers record if children turned off the light in rooms.

The data collection strategy of Mikami et al. (2022) [*Home economics class including measurement of energy consumption*] is somewhere in-between directly observing energy-saving behavior and validating the electricity meters. Three hundred students from High School in the Kanto region (Japan) were in charge of directly measuring and reporting their household's energy consumption for electricity and gas, before and after a 6-lessons program on energy-savings behaviors. Although the definition of observed behavior becomes fuzzy in this study depending on whether students correctly reported their household's energy consumption or not, this measurement is accompanied by a 16-item questionnaire asking for specific energy-saving behaviors that suggest some consistency. Moreover, the compliance with observed behavior could be an outcome itself, since being asked to measure and report may change behavior.

Actual energy consumption was also measured by Pietrapertosa et al. (2021) [*Competition between schools in art and energy consumption*]. In this study, a complex intervention between schools involved creative thinking, learning and competition regarding energy use. The outcome variable in this study was the actual natural gas and electricity consumption of the schools involved in the program. Rolim and Baptista (2021) document a similar intervention [*Digital competition between schools*] that involved the registration of children and their families in an app that allows reporting energy saving activities (e.g., trips substituting the use of a car by walks and bicycle trips) and taking quizzes on sustainability issues to earn points for their school, so it receives a prize. Both studies involving competitions are measuring behaviors, but they are also implicitly using descriptive norms: between-school performance comparisons give an idea of what the population of interest is doing, with a stark remark on the reference groups (i.e., my school and other schools).

From a comparative perspective, we want to stress two points. First, it calls our attention that, according to the Energy Vulnerability Index reported in Liu et al. (2023), most countries with a high vulnerability are located in sub-Saharan Africa and South Asia, but the studies targeting energy use in our review were implemented in Europe, United States, and Japan. Given that energy aid is well-targeted to the most energy vulnerable countries (Dong et al., 2024), there is room for designing projects involving educational institutions. Second, note all the described interventions aimed at reducing consumption, though it remains as an open question whether educational institutions can become a vehicle for switching energy sourcing. Whereas in the developing world the installation of

[14] As mentioned ealier, energy saving behavior might constitute adaption if it is done after, for example an electricity shortage.

[15] Gill and Lang (2018) devised an intervention along these lines. However, as a reminder of the difficulties of such methodological implementations, they matched 50 (out of 60) households using their addresses. A small number compared to their control group, including over 1,500 households.

solar panels may increase education enrolment in deprived areas (see Gillani et al. (2022) for the case in Pakistan), in developed countries may foster environmental attitudes in children, parents and teachers (see Izadpanahi et al. (2017) for a study in Australia). Linking these infrastructure adequations to the school's CCE plans may help accelerate the private adoption of such alternative energy sources.

Table 3. Overview of types of interventions and measured outcomes

		Intervention Type			
		Engaging learners	**Lectures, deliberative discussion, and interactions with scientists/science**	**Nudges**	**Involving teachers and communities**
Outcome Type	Recycling behavior and clean-up	**García-Vásquez et al. (2021), Spain:** Role play activity of children in mines recycling of phones. **Druen and Zawadzki (2021), USA:** Games /simulation signing up for clean-up activity. **Tsai and Tan (2022), China:** Lectures / Hands-on in sorting recyclables / Documenting environmentally friendly behavior family and friends' norms. **Donmez-Turan & Kiliclar (2020), Turkey**: Environmental training recycling of packages.	**Nourmoradi et al. (2021), Iran:** Lectures & group discussions recycling in school. **LaCombe and Danoff-Burg (2013), Bangladesh:** Learning on proper disposal of trash, clean-up candy wrapper disposal. **Arya and Maul (2016), USA:** Discussion of scientific studies on climate change issues correct recycling of paper.		
	Reduction of energy and electricity use	**Pietrapertosa et al. (2021), Italy:** Competition between schools gas consumption at school.		**Frantz et al. (2016), USA:** Posters and stickers washing with cold water / turning lights off.	**Rolim and Baptista (2021), Portugal:** Competition between schools (other community members could sign up)

				Mattsson & Laike (2022), Sweden: Boardgames turning off lights. **Mikami et al. (2022), Japan:** Stickers and reminders (self-reported) energy use at home.	sustainable mobility (bicycle and walking trips)
	Reducing food waste			**Yazdankhah et al. (2020), Iran:** Posters, stickers, smaller portions food waste.	
	Respecting nature and species	**Collins & O'Riordan (2022), Ireland:** Lecture on pinguins and lemurs /Zoo visit feeding, chasing of animals.			
	Creation of new infrastructures				**Kendall et al. (2021), Uganda:** Teacher training building of fuel-efficient stoves.
	Multidimensional	**Somwaru (2016), Suriname:** Environmental "do-it yourself" activities water saving, cleanliness, environmental brigades. **Sulistyarini et al. (2022), Indonesia:** Waste handling / Environmental extracurriculars / Planting trees school cleanliness and compost making, among others.			

	Survey	**Jans et al. (2023), Germany:** Vegan cooking workshop believed norms of others. **Zhang et al. (2019), China:** Nature club beliefs about others' behavior regarding insects.			

Source: Authors' creation.

Reducing food waste

School canteens are a promising scenario for pedagogic practices regarding sustainability: they have the potential to revert the large quantities of generated waste (which can be as high as 45% of the food served, see Byker et al., 2014), while interventions using active learning can be easily intertwined with behavior during meals. Nevertheless, the awareness regarding food waste generated at schools appears to be typically low. For instance, Derqui et al. (2020) conducted a survey responded by 420 school headteachers and find that even if Spanish schools are largely engaged in sustainability issues, this does not often include canteen food waste reduction.

Food and food waste is another very specific resource that can be studied as a resource protection behavior aimed at climate change mitigation. Yazdankhah et al. (2020) [*Pamphlets on food wastage when buying lunch*] observe the amount of food wasted in two university restaurants after introducing a set of different measures at the restaurant level. Again, it should be noted that outcomes from this intervention have to be taken at the restaurant level and cannot be used at the individual level.

Respecting nature and species

Children's understanding of conservation and ecological thinking can be fostered through direct (e.g., zoo visits) and indirect experience (e.g., documentaries or books) with animals. Based on a study with 171 children interviews as part of a zoo visit, Myers Jr. et al. (2004) reveal that conservation thinking gets behind ecological thinking during children's age development. Hence, children's curiosity about animals may be redirected toward more conservation thinking as part of CCE and ESD interventions.

Although in our list of 86 studies we had 10 directly related to conservation, often describing interventions with zoo visits or outdoor explorations, most of them were measuring increases in knowledge or changes in attitudes, and only one looked at behavior. Collins and O'Riordan (2022) [*Lecture prior to zoo visit*] describe an educational intervention performed with several classes at an Irish zoo (in the city of Cork), in which children saw a slideshow presentation with pictures of two of the zoo's species, lemurs and penguins, to increase knowledge and create an emotional connection with both species. The observed outcome are incidences of negative behaviors such as chasing, feeding, and touching animals, or flash photography and banging on glass.

Creation of new infrastructures

Energy efficient behavior sometimes relies on the existence of appropriate infrastructures. A way to implement long lasting change can thus be by focusing on behavior that will improve the existing infrastructure. This approach is taken by Kendall et al. (2021) [*Construction of fuel efficient stoves*] that observe the construction of new more fuel-efficient stoves that use less wood and water. Teachers

receive training and, together with their students, they build these stoves on sites jointly selected by the teachers, the students, and the community. Sites include schools, student and teacher homes, and homes of local leaders or other community members.

While the paper reports the number of new stoves actually built for the treatment group, unfortunately no data is provided for the control treatment.

Studies where CCE leads to the actual provision of infrastructure are scarce. We list below a couple of references that, despite not being included in our study, illustrate the importance of these creations: they reveal how channeling students' engagement into action can bring important cues of a sustainability mindset. The first example comes from Lestari et al. (2022), who implemented the RADEC learning model. Indonesian students from elementary school that participated in the RADEC program proposed, in their last stage devoted to creation, the construction of recyclable and portable trash cans. Although there is no report that the bins were indeed constructed (and therefore the study is excluded from our final dataset), it hints the role of creativity in the transmission of sustainable behaviors. The second example, not included in the review because it does not directly refers to a CCE, reports how students' engagement transformed into actions (i.e., signed petitions), and ultimately led to infrastructure provision: a bike lane to commute to the school in Denmark (Jensen, 2019).

Qualitative, multidimensional approaches

Finally, some of the studies observe behavior in a qualitative sense. This also often implies grouping multiple types of behaviors together that are considered as somewhat environmentally friendly. This approach can be very subjective but gives some interesting insights into the effects to be expected in general. For example, Somwaru (2016) [*Teachers' training workshop on "Green schools"*] report on the general cleanliness of school yards and classrooms, the presence of school gardens, recycling, energy and water saving, and the installation of environmental brigades. Data are based on observations at the schools, discussions with the teachers and by consulting logbooks. The authors themselves mention that this approach has the limitations of subjectivity.

Similarly, Sulistyarini et al. (2022) [*Voluntary action-based environmental program*] study a complex "Green School Program" called the "Adiwiyata Program". The program is carried out during school hours and outside school hours, divided into curricular and extracurricular activities according to school policy. To evaluate the impact, data were collected through observation and structured interviews. Also, in this case the assessment of the effectiveness of the intervention is purely qualitatively and made by the authors.

2.7. Intervention types

As previously discussed, our systematic review included all studies that reported clear climate change education interventions. Given the multitude of observed approaches we present a categorization of the observed interventions below. We will separately discuss interventions that focused on engagement, cognitive (and scientific) arguments, nudges, and involvement of teachers and the school community.

Engaging learners

Engagement of learners is the most prominent approach we observe and can happen on multiple different levels. Indeed, we observe that in some cases the intervention relied on almost minimal engagement. One example is Nourmoradi et al. (2021), which presented educational interventions to secondary school girls. The intervention primarily consisted of five lectures of 30 minutes each, concerning topics related to recycling, importance of water, energy saving, wastewater and diseases and air pollution. To accomplish engagement, lectures were combined with group discussions and brainstorming.

Other approaches to engagement rely on activations of emotional reactions, since emotional engagement can positively impact learning (Ballantyne et al. 2011). This can be achieved sometimes with rather minimal additional effort as in the intervention presented in Collins & O'Riordan (2022). The educational Intervention in this study consisted of a PowerPoint presentation regarding the biology of lemurs and penguins. Then, to achieve emotional engagement the presentation included appealing pictures of the study animals. The children also participated in a hands-on activity during which they made enrichment devices for the study species. This involved filling clear plastic bottles with shiny pieces of foil for the penguins (Clarke, 2003) and cutting up fruit for a scatter feed for the lemurs. Another way of emotional engagement is studied by García-Vásquez et al. (2021). The paper studies the effect of role-play where participants put themselves in the place of children working in cobalt mines.

Simulations and games related to social dilemmas are another approach to engage the learners and make them experiencing the feelings and decisions related to actual behavior and outcomes. Druen and Zawadzki (2021) studied climate-specific social dilemma simulations. For this they developed a simulation of the social dilemma of climate change, called "Climate Trap: Social Dilemma Simulation", which was designed for people to gain insight into the dilemma tradeoffs, to observe how they themselves behave under pressures, and to help them feel empowered to overcome short-term decisional traps that sacrifice the collective.

Engagement is also often achieved through hands-on activities and social interactions. For example, Jans et al. (2023) delivered vegan cooking workshop for children. This was a one-time activity, and not embedded in a broader curriculum. To engage learners, after a presentation on reasons for a vegan diet, cooking workshops were carried out. The researchers adapted these cooking workshops specifically to stimulate shared social identity formation in the cooking group. A similar approach was studied by Pietrapertosa et al. (2021) that investigated a 'Schools4energy" program involving a multifold "learning by doing" methodology. It combined three sets of activities: a competition among schools to reduce energy consumption (School Race) based on the adoption of good energy practices (based on a common Energy Charter), an artistic competition among students to support creative thinking (Art4energy) and gamification as an alternative method to learn about energy (Play4energy). Other studies focus on the specific behaviors that they want to change. For instance, Tsai and Tan (2022) studied the effect of four weeks of lessons involving environmental and sustainability topics, followed by hands-on sorting of recyclables.

Finally, certain CCE achieve engagement through complex interventions that involve schools or even larger areas on multiple different levels. Somwaru (2016) studies the implementation of the 'Green School' project at school level. This involved information about endangered species and school gardens, recycling and its importance, lessons about the impact of pollution on the environment, lessons about the importance of water and energy, formulation of rules for a 'The Green School', installation of environmental brigades to make sure the entire school abides by the rule and environmental games. Sulistyarini et al. (2022) studies the "Green School Program" called the "Adiwiyata Program" by the Indonesian Ministry of Environment. It provides opportunities for schools, especially students, to be involved in pro-environmental activities. The program is carried out during school hours and outside school hours, divided into curricular and extracurricular activities according to school policy. The curricular activities are embedded in the subjects, i.e., students get environmental materials integrated into school subjects such as chemistry, physics, biology, citizenship, geography, and entrepreneurship education. Extracurricular activities happen after school hours for 3-4 hours per week, with a compulsory character. They include participatory-based environmental activities, like making compost, utilizing used goods that can be recycled, giving botanical names to trees, and planting trees, flowers, and medicinal plants.

Lectures, deliberative discussion and interaction with scientists or science

As a close example, LaCombe and Danoff-Burg (2013) studies a "proper disposal of trash" activity. For this the authors introduced the concept of littering versus proper disposal of trash to participants in the treatment group and then discussed the potential consequences of litter on humans, animals, and

potential income streams like ecotourism. This was followed up by a trash clean-up activity in the surrounding area.

Another example is Arya and Maul (2016) who study an eight-week iteration of an educational program called CELL (Climate Exchange for Language and Learning), during which students read and discussed adapted versions of seminal scientific studies related to climate change issues (e.g., rise in sea level). The idea of the program was to foster evidence-based discussions about climate change concepts and issues. It also provided online opportunities to connect with students in various parts of the world – China, New Zealand, Norway, and the United States – as well as scientists across different disciplines in order to discuss and evaluate the evidence for global climate change.

Nudges

The use of nudges in education settings has substantially increased in the last two decades (Damgaard and Nielsen, 2018). The most popular nudging categories according to Damgaard and Nielsen's review include reminders (to students, teachers, and parents), social comparison (mostly about test outcomes), the transmission of cues on social belonging to activate the students' identity with their institution and their learning role, and nudges providing information. The latter is the wider in scope, and it is divided in Damgaard and Nielsen's study in information transmitted to parents, about behavior and ability, about the returns to schooling, and about financial aid.

In our review, some of the CCE activities use methods and approaches that can be categorized as informational nudges. They include reminders, posters, and stickers that target attention limitation and other informational barriers. Though very simple to implement, they indeed can have positive effects. Examples include Frantz et al. (2016), who tested the effects of introducing specific stickers and posters visible in the campus laundry areas regarding the advantages of washing with cold water, and reminders and signs prompting at turning off the lights in the classrooms. These interventions were carried out at Oberlin College (Ohio, United States), and observations were collected by inspecting washing machines at specific times during the weekends (when are more often used) and classrooms on two moments of the day during the weekdays.

Mikami et al. (2022) specifically refers to the approach of nudging. Their approach was to use illustrations related to energy-saving behavior, that appeared many times in a text and were printed as stickers that could be used as reminders at home. This was combined with showing 8 concrete behaviors that involve changing device settings and 8 behaviors that involve daily actions and to provide a concrete image of the action to be taken. Finally, they explicitly recommend to children the changing of default settings, such as the temperature for heating and cooling devices, so that children could quickly feel the effects.

Nudge-based educational interventions might be particularly effective on young children, who do not yet have the maturity to understand complex concepts. This has been investigated by Mattsson and Laike (2022). The authors used digital and paper-board games and visual prompts to stimulate pre-school children aged between 3 and 6 years old to turn off the lights. They found that only visual stimuli have a significantly positive effect on behavior.

Other nudge approaches are more complex and combine multiple behavioral interventions. For instance, Yazdankhah et al. (2020) studied food wastage in two university restaurants. The wastage was observed before and after intervention, and customers in the restaurants also responded to questionnaires on knowledge, attitudes, and behavior. The intervention was complex: it relies on the distribution of flyers and posters, but also on training staff to serve less and only on demand. In addition, containers were made available to keep leftover food, and bread was served in smaller pieces. Jointly these interventions indeed reduced the amount of food wastage observed.

Involving teachers and communities

Though all CCE interventions rely to some degree on teachers and other members of the school community, some approaches specifically focus on training. For example, Kendall et al. (2021) investigated the impact of exposing students to teachers trained under the UNITE program.[16] UNITE is a professional development program for instructors that is focused on sustainable activities around the Kibale National Park in Uganda. Although UNITE encompasses several educational activities over the year, the researchers explain their focus on fuel efficient stoves (i.e., using less wood to conserve the environment for animals) due to the current environmental degradation in the park's surroundings.

Control schools did not have such teachers. The evaluation is then conducted up to three years after the teachers have received their training, allowing also the UNITE program to keep track of further stoves constructed by teachers and/or students, requesting visual confirmation of the provided infrastructure.

A recent review, explaining on how public-sphere actions are discussed as part of CCE programs in schools (Kranz et al., 2022), remarks that education strategies involving communities often stem from local problems. This is fundamental to emotionally connect students with topics in more traditional classes, but also leads students to "take action" and help solving their communities' problems. The experience reported by Kendall et al. (2021) aligns with the connection to local problems (i.e., encroachment in the National Park for wood collection) to trigger interactions between students and their communities. It is nonetheless particular that the outcome was the production of infrastructure. Kranz et al. (2022) cite a related example where, in a context as different as Denmark, students' action (by sending petitions to private companies and local village boards) led to the construction of a bike lane heading to their school (Jensen, 2019).

2.8. A comparative analysis of interventions

Let us close this section by presenting a comparative analysis of the studies implemented in high-income countries with respect to those grouped in the category "developing countries" according to Table 1, including from upper middle- to low-income countries.

Our first finding is that, in this list of 19 studies measuring norms or behavior, nine were implemented in high-income countries. The single multi-country study in this group involved three high-income countries and China, in the upper middle-income category (hence, for simplicity when aggregating this information for comparative purposes, we will treat it as a study mainly conducted in a high-income setting). Compared to the initial sample of 86 studies, the proportion of developing countries has increased. The most important implication of this sample reconfiguration is that measuring norms or behavior, which we consider more adequate from a methodological standpoint (see Section 3 for a thorough discussion), is not a more prominent feature from high-income countries.

One more signal of quality in our selected sample of 19 studies, that do not differ by the country's income category, is that these interventions with measured norms or behavior are over representative of the coordination efforts reported in Section 2.5: with respect to the sample of 8, the proportion of studies without any coordination effort falls from 53 to 26 percent. By contrast, those with coordination from the researchers increased from 21 to 36%, and those with coordination with governments or private institutions rose from 22 to 32%. This positive correlation between measuring outcomes beyond

[16] UNITE was founded in 2002 by the North Carolina Zoo and started as a teachers' exchange program. Kendall et al. (2021) presents two reasons for focusing on teachers: they are "multipliers" of the effects as they will have contact with multiple cohorts of students, and they are highly respected by their community and thus have more potential to influence other members of the community.

knowledge and attitudes with larger coordination efforts constitutes a positive signal in the design of more ambitious interventions: the additional coordination efforts can be compensated by more robust tests of behavioral change.

Let us delve into the main difference we found between high-income and developing countries: all the interventions measuring changes in energy consumption were implemented in high income countries, whereas recycling and food waste interventions are more frequent in developing countries (5, which are 55% of all interventions in this group) compared to three interventions (30%) in the high-income countries' subset. Within this comparison group of all interventions in energy and waste (13 studies in total), we can observe differences in the funding bodies[17]: government and international aid agencies are the majority in high-income countries, whereas the funding from academic institutions is the majority 'n developing countries. This might be one case in which aid agencies may look for targeting energy consumption (or sourcing) as the outcome in CCE interventions. In terms of methodology, the standard of a randomized allocation of the treatment and control groups is only met in four studies, two in high-income countries, Spain and Sweden, whereas the other two come from Uganda and China. Whereas the low number of studies employing such experimental design gives some room for improvement, it is noteworthy to validate that we do not see differences between the country's income categories.

On the other hand, the comparative analysis by intervention type is less conclusive because more than half of the studies (10/19) were classified as being of the type "engaging learners." There are nonetheless two elements to highlight. First, nudges are more frequent in high-income countries (3 out of 4), whereas the intervention type "Lectures, deliberative discussion, and interactions with scientists/science" were slightly more frequent in developing countries (2 out of 3). It is probably safe to say that nudges are more scalable than lectures and other interactions involving experts. In this line, there is also an opportunity for funding a different nature of interventions in developing countries, such that they rely less on lecture-based methods and more in school's environment through the nudges' choice architectures.

There is one final categorization we developed for the 19 interventions studied in depth that corresponds to the degree of social interaction required in the intervention. We coded interventions as having a *required* "high" degree of social interaction (HSI) the competitions between schools, the use of interactive games (i.e., boardgames, role-playing, and simulation), group-based workshops such as the one on cooking and the nature club, and also the construction of stoves, and the lecture with cross-country interactions (since it was the main characteristic of this virtual exchange program). On the other hand, we coded interventions as having a *required* "low" degree of social interaction (LSI) the lectures and training programs, the lectures followed by activities that can be performed individually (even if they were done by groups) such as measurement of energy consumption and recycling, and the nudges operating as reminders. We emphasize on the required nature, because even if the activities were done in groups, they could be implemented on an individual basis if one neglects the decreasing costs of scale.

We find that nine interventions required HSI, and the other ten were categorized as requiring LSI. More importantly, we find a stark difference by the country's income levels: the proportion of HSI interventions is 78% in high-income countries, and only 30% in the complementary group pooling upper-middle income and lower levels. This is a remarkable difference that may indicate an income gap in terms of the technical and budgetary requirements to conduct interventions that are collective

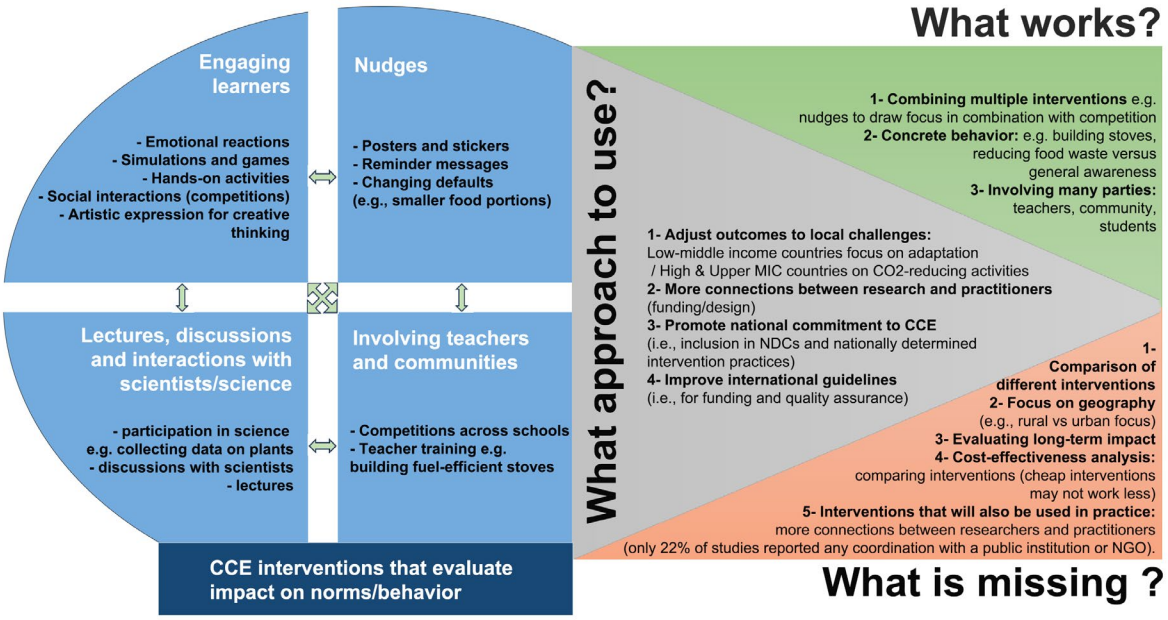
[17] Whereas we found these differences for the subset of 13 out of 19 studies, we did not find systematic differences between funding agencies for the subset of 19 studies.

in their nature, an aspect particularly important if such interactions have a multiplying effect resulting from the additional social interplay.

3. Policy implications

Figure 9 summarizes the interventions related to norms and behavior that we covered in depth in our SR. The oval diagram groups these interventions according to their types, while the internal arrows illustrate their synergic potential. These interventions can be jointly explored to address the complexities of both climate science and social interactions. For instance, abandoning traditional lectures may be difficult for schools, but they can be redesigned in combination with engagement activities for the students and their communities to foster behavioral changes. The combination of multiple interventions is highlighted as a top priority among the list of effective measures that we have identified (see the upper-right triangle), as documented, for example, in Collins & O'Riordan (2022) where lectures on animal behavior were followed by a zoo visit.

Figure 9. Exploring the new paradigm of research capturing revealed behavior: What works, what approach to use (in future interventions), and what is missing?



The bottom-right corner lists aspects that we found to be missing in the reviewed interventions. For instance, we did not find comparisons of different interventions assessed within the same study. This is something that could alleviate the general tendency for everything to be effective in this literature, making it hard to distinguish what works from what does not. Second, interventions conducted in coordination with schools could be designed to evaluate long-term impacts as well. The presence of students inside the educational system for a considerable number of years has the advantage that long-term effects could be more easily assessed with respect to other settings where follow-up measurements would be harder or very costly to implement due to attrition problems. The third missing aspect is the need for stronger connections between researchers and practitioners. These connections should be extended to educational institutions to facilitate the implementation of interventions. Our analysis of the reviewed studies reveals a decentralization of efforts. There is no focal point, in the form of general curriculum guidelines or a set of encouraged practices, that provide a "cookbook" of educational interventions. Article 11 of the Paris Agreement delegates capacity building and climate empowerment through education to the national governments, as they must respond to national and local contexts. Nevertheless, not having a set of recipes that work (and can be adapted

to the local taste) may be delaying our understanding of how to steer CCE interventions. One first step would be to explicitly include CCE expansion efforts into the country's Nationally Determined Contributions (NDCs). Two examples are Italy, who declared in their NDCs that environmental education will be mandatory in all schools from early childhood to high school, and Chile, that committed to the creation of technical schools to create a workforce for the transition to green energies (solar and wind).[18] In what follows, we provide a discussion on three key elements for improving the design and evaluation of CCE interventions using the UNESCO's framework for the implementation of Education for Sustainable Development as a focal point in terms of objectives.

The 2020-2030 UNESCO's framework for the ESD acknowledges climate change as a "real and rapidly-evolving threat for humanity" (UNESCO, 2019). This framework calls for more engagement in ESD because education is a key element in the response to climate change. This engagement is channeled through five priority action areas (PAAs) targeting different stakeholders in education:

- PAA#1: Integration of ESD into the regional and national policies of education
- PAA#2: Innovative learning environments
- PAA#3: Development opportunities for teachers and educators
- PAA#4: Engagement opportunities for the youth
- PAA#5: Empowerment of communities as platforms for activities beyond the classroom

Below, we dig deeper into the implications for intervention design when it comes to state-of-the-art intervention evaluation, comparing our findings from the SR with the prism of analysis of the UNESCO's PAAs.

3.1. Understanding the intention-action gap by changing the measured outcomes

Farjam et al. (2019) reveals that the gap between environmental attitudes and behavior widens when the environmentally desired action becomes more costly. This result, revealed with a lab experiment in which participants could sacrifice personal gains to send money to a carbon offsetting project, shows that people are willing to pay to make their actions consistent with their intentions, but only when it is cheap to do so.

This result shows why studies that measure intentions have a major drawback: they cannot tell us much about the change in behavior when the switching costs are high. Moreover, the educational environments can foster the idea that there is only one correct intention to report, increasing the social desirability bias.

Since PAA #4 and #5 call for the engagement and empowerment of students and their communities, this "call to action" beyond the classroom should be accompanied by the measurement of revealed rather than reported or intended behavior. We acknowledge that measuring behavior can be more costly, but technology is making its part to lower these costs. Consider, for example, the "Sharing Lisboa" program and how children and their parents reported various environmentally desired behaviors, such as using bikes or walking instead of using the car, as part of a competition between schools (Rolim and Baptista, 2021). Recall from our description of this intervention that the behaviors were validated through a mobile app. While there may have been an initial cost associated with programming and deployment of this app, the majority of the expenses for behavior measurement were shifted to the participants, who were responsible for uploading proof of their actions in order to earn points.

When measuring behavior is unfeasible, assessing social norms via norm-related beliefs offers a potential solution for two key reasons. First, norm-related beliefs have a more specific content and definition, and they are strongly related to real behavior, which is not the case for other constructs such as intentions and attitudes (Bicchieri, 2016). Second, even if a social desirability bias may still be

[18] https://www.careaboutclimate.org/blog/climate-education-as-a-foundation-for-achieving-the-goals-of-the-paris-agreement (accessed on the 20th October 2023).

present when eliciting norm-related beliefs, small incentives may help correct such bias and motivate more accurate responses (something that is not possible when one elicits attitudes and intentions). In particular, methodological advances in behavioral economics offer techniques to accurately measure norm-related beliefs using real incentives (Bicchieri and Xiao, 2009; Krupka and Weber, 2013). A simple reward system that can be implemented in educational settings to elicit accurate norm-related beliefs is to use tokens that can be exchanged for small presents like stickers, candies, or small toys at the end of the study (Sutter et al., 2019).

We found only three studies that directly measure social norms. Further research should be pursued in this direction, so that scholars, policymakers and decision-makers in the education sector can assess the potential of social norms as an outcome in itself or as a first step (i.e., a diagnostics) that may lead to an intervention to foster behavioral change through changes in norms. One further advantage, in line with PAA #4 and #5, is that the reference group for expected behavior could be either students or other groups identified by a close social identity (e.g., neighbors in the assessment of energy consumption or transportation of their household).

3.2. Increasing coordination efforts

Our analysis of whether CCE interventions were following some national or local guidelines revealed that only eight studies reported any coordination with a public institution, and another eleven reported that a non-government organization was behind the coordination efforts required for the intervention to be implemented. Together, they constitute 22% of the 86 studies in our review, a figure suggesting that there are coordination opportunities for increasing the scale and comparability of such interventions, in line with the PAA #1 and its call for integration of ESD into the regional and national policies of education.

The suggested increase in coordination efforts emerges from the need to conduct larger CCE interventions, potentially involving more schools with a clearer delimitation of specific activities as part of the intervention. This could help to achieve the goal of understanding what works, and therefore should be expanded in several directions: (i) increasing sample sizes, both in the number of students and number of classes, would improve comparability between the treated and baseline groups; (ii) activities provided by external institutions (e.g., zoo visits) may become more attractive to such providers if a larger number of institutions participate; (iii) national and regional education offices, may find it simpler to oversee the implementation of ESD practices when specific activities serve as focal points for coordination efforts; (iv) coordination between more schools would lead to more precise interventions, ensuring replicability and enabling a better understanding of which specific activities or intervention components work.[19]

The limited number of studies revealing coordination efforts suggest the need to review the incentives that may encourage teachers, communities, and scholars to participate in more extensive and ambitious projects. An open question that arises is how to reward the costs of coordination. The documented experiences from non-profit organizations can offer a clue, aligned with PAA #3, on how to address this question: certifiable training and capacity-building programs that can engage teachers and other adult community members in programs recognizing these influential figures as the first step in transmitting ESD and CCE concepts. The UNITE program and the MADE CLEAR initiative, as documented in Kendall et al. (2021) and Shea et al. (2016), serve as two good examples of the positive effects of teachers' engagement through hands-on training activities.

[19] The latter point may not only be beneficial in terms of policy outcomes but also in terms of scientific dissemination: in a non-negligible number of the reviewed studies, the description of the intervention was lengthy and unclear, as it involved several steps (e.g., curriculum discussions, curriculum modifications, development and implementation of new materials, with multiple focus groups and changes in the unit of analysis throughout the process).

Another category of rewards revolves around recognition. One effective way to combine recognition with improved coordination efforts is by funding and promoting a platform where teachers and students can access educational materials on climate change. This platform should allow students and teachers to record their teaching experiences with such materials using a predefined template. There are already some websites that fullfil the first requirement, that is, the promotion of CCE materials. However, in absence of coordination efforts it is hard to learn what works best.[20] Predefined templates can also facilitate other dissemination processes, such as making popular materials available in multiple languages.

We want to emphasize that scalability and replicability are related but distinct concepts. Scalability guarantees that a given intervention can be expanded, for example, from one school where the pilot study took place to all the schools in an educational district. Replicability, on the other hand, guarantees that the procedures are sufficiently documented, allowing the same intervention to be successfully implemented in a different environment (i.e., another school). Consider, for example, the multi-country studies we have reported. Those that only required a survey for their implementation can be more easily scaled up and replicated, as distributing surveys mostly involve translation costs and coordination between researchers and/or practitioners. However, a multi-country study involving an exchange of experiences among young adults from different countries (see Arya and Maul, 2016) could be easily replicated (by the same scholars or another group of scholars) but it would pose significant challenges when it comes to scalability. This is because interactions between participants from a more diverse set of countries would become more complex, involving issues such as language barriers and coordination between different time zones.

Lastly, we acknowledge that more coordination may also increase the costs in terms of data collection. In particular, the scalability of the CCE interventions may represent considerable costs for registering outcomes based on observed rather than self-reported behavior. This is another opportunity to adopt a framework that exploits the potential of social norms' elicitation.

Since there is a will to implement CCE activities given the current climate situation, we believe it is a matter of improving coordination and agreeing on measurement targets and instruments to rapidly learn about what works best and how to adapt it to local contexts.

3.3. Reallocating data collection efforts

One surprising finding from our SR is that 67 out of 86 studies employ a pre-post approach. We argue that this approach has several inherent issues. Data collection efforts should be reallocated. Instead of two "waves" of data collection at different points in time, it would be preferable to have two groups: one receiving the intervention and the other serving as a reference. It is essential to pay close attention to the comparability of these groups and measure the outcome only once. While having both treatment and reference groups with multiple measurements would be ideal, it may require extra efforts as data collection would occur at two different moments in time (pre and post) for two different groups (those receiving the intervention and those who do not).

The primary issue with the pre-post approach is that it increases the demand for an "appropriate" response after the intervention. This demand can be particularly pronounced in educational settings, where students are supposed to learn something new after an activity. For example, when assessing attitudes or intentions, regardless of their baseline response, students may feel external or internal pressure to report a change in their views and attitudes toward sustainability after they have been made thinking about the environment.

[20] Two examples are the webpage from the United States Environmental Protection Agency (https://www.epa.gov/climate-change/climate-change-resources-educators-and-students) and the platform Subject to Climate (https://subjecttoclimate.org/). The latter is an example of a very rich set of teaching resources (the search tool lets you navigate over 2,400 results), where the vast availability of materials would make hard the instructor's task of selecting which one to adopt.

A similar issue arises when the outcome is knowledge. Students take a test before receiving relevant information and then take another test after receiving the information as part of a lecture. Obtaining positive results on knowledge after a CCE intervention is unsurprising. This can lead to difficulties in understanding which interventions are genuinely effective. When everything appears to work, it creates a paradox in which the few interventions that do not work become more informative. We do not advocate for stopping to collect information on knowledge outcomes, but rather for going beyond and gathering information about behaviors and social norms. Knowledge outcomes can serve as a validation check when examining other outcomes such as norms and observed behaviors.

We ackowledge that one potential objection to having a treated and a baseline group is rooted in fairness concerns: why do some kids receive the intervention while others do not? We believe that the simplest way to avoid this conundrum is to delay the CCE interventions in the baseline group, rather than suppressing them altogether. Moreover, the idea of experimentation, understanding what works and what does not, is aligned with the pursuit of innovative learning environments, as outlined in PAA #2.

4. Conclusions

In this systematic review (SR), we studied educational interventions within the definition of CCE that focus on changes in actual behavior and normative beliefs related to climate change. The SR explores the new paradigm of research capturing revealed behavior, providing some initial evidence for practitioners and researchers from education, climate and behavioral science on what has been rigorously tested and what should be better documented. We exclude purely conceptual studies and studies that only evaluate or analyze the impact of CCE interventions on knowledge, attitudes, behavioral intentions, psychological states, and self-reported behaviors. The main contribution of this study is to critically appraise the characteristics of CCE interventions that have the potential to trigger a normative shift, either by changing what people collectively think one should do (injunctive norms) or what they actually do (descriptive norms).

Starting from a large set of initial studies (12,000 hits during original search) we systematically identified studies that are related to the environment and use clear educational interventions on children or young adults. Our final set of studies is surprisingly small given the large interest in CCE. Within the 86 retained studies, we observe a disproportionate focus of CCE research on developed countries. An analysis of the countries involved, based on measures of vulnerability to climate change and readiness to leverage investments for adaptation, shows that there is a lack of studies in countries that are not (yet) well positioned to adapt to climate change. These are the nations that likely require education to enhance their adaptive capacity the most. Very few studies are conducted in multiple countries, which indicates a lack of direct cross-country comparisons. We also observe that most studies target urban populations (75%). The lack of focus on rural populations is surprising, given that these populations might be involved in activities directly linked to the environment. Finally, we observe that, in high-income countries, the majority of studies are funded by government and international agencies, while in developing countries, academic institutions provide the majority of funding for CCE research.

Regarding norms, we identify only very few studies that clearly focus on the effect of CCE on either injunctive (3 studies) or descriptive (16 studies) norms. The identified studies mostly focus on behaviors related to recycling and waste. The second most covered topic relates to behaviors associated with energy reduction or electricity consumption. If we consider the impact of these behaviors in terms of CO2 emissions, acts such as recycling and turning off lights are on the lower end regarding overall effectiveness. This raises questions as to why there is such a strong focus of CCE research on these behaviors, rather than on more impactful behaviors, such as those related to transportation (with no studies identified from the selection criteria of this SR) or food (with only two studies available).

The employed interventions aim in most cases at engaging learners. This ranges from group discussions, participating in role playing, artistic competitions to hands-on activities. Other studies focus on nudges (like stickers or posters) to change norms. Few studies focus on the effects of more traditional lectures or exposing students to scientific results and scientists. We briefly discussed the scalability of outcomes, highlighting the potential usefulness of eliciting social norms when interventions target a larger number of pupils. When it comes to engagement strategies, we must also consider scalability. Nudges with small reproducibility costs (such as the aforementioned stickers and posters) are good candidates for scalability. Inter- and intra-school competitions are another promising option for scalability, and with the right incentives, they may benefit from network externalities (e.g., having additional competitors can lead to non-linear improvements in environmentally desired outcomes and information transmission). One relatively unexplored category with scalability potential for adolescents' engagement is social media content. This potential must nevertheless be explored with caution, given the limited control over social media beyond the scope of the intervention. On the other hand, specific experiences may be more difficult to scale. Two good examples are visits to zoos and interactions with scientists. In these cases, it is important to further explore whether substitutes (e.g., outdoor exploration for zoos, and virtual encounters with scientists) can retain part of the positive effects.

The results of this SR highlight the importance of conducting further studies on CCE that consider actual behavior and norms. These studies should not exclusively concentrate on knowledge transfer or intended actions, as we have seen that the intentions-behavior gap depends on the costs to take action (Farjam et al., 2019). The results also highlight the lack of studies conducted across multiple countries and the need of further coordination efforts at different scales (regional, national, global) and with both public and private actors, and the lack of studies eliciting descriptive norms.

Regarding the studied interventions, it is not clear which type is the most effective. There are two main issues to reach more specific conclusions in terms of effectiveness. First, our SR methodology was very ample in terms of fields of publication of CCE results (e.g., social and environmental sciences were more frequent, but also education, engineering and medicine were included), making comparability harder. We pointed out the methodological issues from the before-after comparisons, which were quite popular among the reviewed studies, but their overall positive results may be overestimated due to a social desirability bias. Second, studies testing knowledge and attitudes as main outcomes may have a "lower bar" in showing positive results, and therefore remains unclear which of these interventions may have worked in closing the intention-action gap.

With this in mind, it is necessary that future studies target more specific interventions, which may also ease the systematic study of heterogeneities (e.g., intervention's effectiveness by gender) in whether some programs work better than others. Moreover, funding institutions might help raise the bar in terms of measured outcomes by allocating resources to projects measuring outcomes beyond knowledge and attitudes. As we have stressed throughout our analysis, eliciting social norms has potential for those scenarios in which directly measuring behaviors is too costly or unfeasible to scale. Another lesson in terms of targeting for funding institutions comes from the difficulties in finding specific studies testing adaptation interventions. Since some CCE interventions can be considered to aim mitigation or adaptation purposes, depending on the context, there is a need to state clearer measurable outcomes directly reflecting adaptation.

We also observe a focus on urban populations without evidence of these populations having the most impact. For instance, CCE interventions targeting changes in energy sourcing and those aimed at climate change adaptation (which were scarce in our SR) in rural communities may have larger impacts because, as Kendall et al. (2021) argued, rural teachers and schools are an important source for successful information transmission, especially in the developing world. We finally observe a focus on behaviors related to recycling and energy saving that might be driven by the relative ease of observing these behaviors. However, many other important domains regarding environmentally relevant behavior, for example transportation or food choices, are so far little studied. Educational interventions aimed at biodiversity conservation are frequent, but most of them still need to take the step from measuring knowledge and attitudes toward observed behaviors.

Among the reviewed interventions, there is one intervention type that we consider having the potential to articulate most of the desirable elements discussed: competitions between schools in which some environmentally desirable behaviors are encouraged, as they add points to a scoring system allocating prizes to schools. Their potential dwells on several elements: (i) they directly target verifiable behaviors, since they are part of the competition's scoring rules; (ii) comparison between schools are implicitly using descriptive norms (i.e., how is my group performing with respect to others); (iii) the perimeter of competition can be expanded so families and other community members may also get involved in behavioral change, combining in-school and out-of-school approaches; (iv) they imply the articulation of coordination efforts between schools, but also with some public offices or private organizations that serve as the competition organizers and referees; and (v) they strongly benefit from the involvement of funding agencies, which may send a credible signal that prizes would be granted and their implementation oversaw (e.g., prizes in terms of infrastructure). This comes also with methodological challenges, such as the evaluation of a program that by definition may have strong spillovers, as well as key features in the contest design such that it does not trigger discouragement feelings among the children whose schools are getting behind in the competition.

Bibliography

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9–10), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Anderson, A. (2012). Climate Change Education for Mitigation and Adaptation. *Journal of Education for Sustainable Development*, 6(2), 191–206. https://doi.org/10.1177/0973408212475199

Armstrong, P., & Grant, J. (2004). How research helped us to move from awareness to action and then to systems development. *Australian Journal of Environmental Education*, 20(1), 13-24.

Arya, D., & Maul, A. (2016). The building of knowledge, language, and decision-making about climate change science: a cross-national program for secondary students. *International Journal of Science Education*, 38(6), 885-904.

Ategeka, J., Booth, S., Cavatassi, R., Curtis, B., Filippi, V., Kim, D. S., Kim, Y., Langer, L., Mapitsa, C. B., Mokgano, E., Prowse, M., Puri, J., & Robertsen, J. (2022). Behavioural Science Interventions Within The Development And Environmental Fields In Developing Countries: A Systematic Review. https://www.ifad.org

Baird, J., Dale, G., Holzer, J. M., Hutson, G., Ives, C. D., & Plummer, R. (2022). The role of a nature-based program in fostering multiple connections to nature. *Sustainability Science*, 17(5), 1899-1910.

Ballantyne, R., Packer, J., & Sutherland, L. A. (2011). Visitors' memories of wildlife tourism: Implications for the design of powerful interpretive experiences. *Tourism Management*, 32(4), 770-779.

Barrett, S., & Dannenberg, A. (2014). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, 4(1), 36–39. https://doi.org/10.1038/nclimate2059

Bicchieri, C. (2006). The grammar of society: The nature and dynamics of social norms. *Cambridge University Press*.

Bicchieri, C. (2016). Norms in the wild: How to diagnose, measure, and change social norms. *Oxford University Press*.

Bicchieri, C., Muldoon, R., and Sontuoso, A. (2018). Social Norms, *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2018/entries/social-norms/.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2), 191-208.

Biel, A., & Thøgersen, J. (2007). Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *Journal of Economic Psychology*, 28(1), 93–112. https://doi.org/10.1016/j.joep.2006.03.003

Bird, M., & Subramaniam, A. (2022). 4-H Water Wizards: Lessons learned for effective afterschool science programming. *California Agriculture*, 76(4), 112-120.

Bisaro, A., & Hinkel, J. (2016). Governance of social dilemmas in climate change adaptation. *Nature Climate Change* (Vol. 6, Issue 4, pp. 354–359). Nature Publishing Group. https://doi.org/10.1038/nclimate2936

Brownlee, M. T., Powell, R. B., & Hallo, J. C. (2013). A review of the foundational processes that influence beliefs in climate change: Opportunities for environmental education research. *Environmental Education Research*, 19(1), 1-20.

Byker, C. J., Farris, A. R., Marcenelle, M., Davis, G. C., & Serrano, E. L. (2014). Food waste in a school nutrition program after implementation of new lunch program guidelines. *Journal of nutrition education and behavior*, 46(5), 406-411.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6), 1015.

Collins, C., & O'Riordan, R. (2022). Data triangulation confirms learning in the zoo environment. *Environmental Education Research*, 28(2), 295-317.

Cooley, S., Schoeman, D., Bopp, L., Boyd, P., Donner, S., Ito, S. I., ... & Yool, A. (2022). Oceans and coastal ecosystems and their services. *In IPCC AR6 WGII*. Cambridge University Press.

Crandon, T. J., Scott, J. G., Charlson, F. J., & Thomas, H. J. (2022). A social–ecological perspective on climate anxiety in children and adolescents. *Nature Climate Change*, 12(2), 123-131.

Damerell, P., Howe, C., & Milner-Gulland, E. J. (2013). Child-orientated environmental education influences adult knowledge and household behaviour. *Environmental research letters*, 8(1), 015016.

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, 64, 313-342.

De Dominicis, S., Bonaiuto, M., Carrus, G., Passafaro, P., Perucchini, P., & Bonnes, M. (2017). Evaluating the role of protected natural areas for environmental education in Italy. *Applied Environmental Education & Communication*, 16(3), 171-185.

Deisenrieder, V., Kubisch, S., Keller, L., & Stötter, J. (2020). Bridging the action gap by democratizing climate change education—the case of kidZ 21 in the context of Fridays for Future. *Sustainability*, 12(5), 1748.

Derqui, B., Grimaldi, D., & Fernandez, V. (2020). Building and managing sustainable schools: The case of food waste. *Journal of Cleaner Production*, 243, 118533.

Dijkstra, E. M., & Goedhart, M. J. (2012). Development and validation of the ACSI: Measuring students' science attitudes, pro-environmental behaviour, climate change attitudes and knowledge. *Environmental Education Research*, 18(6), 733-749.

Dong, K., Jiang, Q., Liu, Y., Shen, Z., & Vardanyan, M. (2024). Is energy aid allocated fairly? A global energy vulnerability perspective. *World Development*, 173, 106409.

Donmez-Turan, A., & Kiliclar, I. E. (2021). The analysis of pro-environmental behaviour based on ecological worldviews, environmental training/knowledge and goal frames. *Journal of Cleaner Production*, 279, 123518.

Druen, P. B., & Zawadzki, S. J. (2021). Escaping the Climate Trap: Participation in a Climate-Specific Social Dilemma Simulation Boosts Climate-Protective Motivation and Actions. *Sustainability*, 13(16), 9438.

Dudley, D. A., Cotton, W. G., & Peralta, L. R. (2015). Teaching approaches and strategies that promote healthy eating in primary school children: a systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 12, 1-26.

Eom, K., Kim, H. S., & Sherman, D. K. (2018). Social class, control, and action: Socioeconomic status differences in antecedents of support for pro-environmental action. *Journal of Experimental Social Psychology*, 77, 60–75. https://doi.org/10.1016/j.jesp.2018.03.009

Farrow, K., Grolleau, G., & Ibanez, L. (2017). Social norms and pro-environmental behavior: A review of the evidence. *Ecological Economics*, 140, 1-13.

Farjam, M., Nikolaychuk, O., & Bravo, G. (2019). Experimental evidence of an environmental attitude-behavior gap in high-cost situations. *Ecological Economics*, 166, 106434.

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *In Trends in Cognitive Sciences* (Vol. 8, Issue 4, pp. 185–190). https://doi.org/10.1016/j.tics.2004.02.007

Feilen, K. L., Guillen, R. R., Vega, J., & Savage, A. (2018). Developing successful conservation education programs as a means to engage local communities in protecting cotton-top tamarins (*Saguinus oedipus*) in Colombia. *Journal for Nature Conservation*, 41, 44-50.

Flora, J. A., Saphir, M., Lappé, M., Roser-Renouf, C., Maibach, E. W., & Leiserowitz, A. A. (2014). Evaluation of a national high school entertainment education program: The Alliance for Climate Education. *Climatic Change*, 127, 419-434.

Frantz, C. M., Flynn, B., Atwood, S., Mostow, D., Xu, C., & Kahl, S. (2016). Changing energy behavior through community based social marketing. *The Contribution of Social Sciences to Sustainable Development at Universities*, 259-272.

Garcia-Vazquez, E., Cartón, P., Domínguez, M., Rodríguez, N., Bustillos, A., & Dopico, E. (2021). For a good selfie. Enhancing mobile phone recycling through simulated exposure to cobalt mining. *Sustainable Production and Consumption*, 26, 44-53.

Gill, C., & Lang, C. (2018). Learn to conserve: The effects of in-school energy education on at-home electricity consumption. *Energy Policy*, 118, 88-96.

Gillani, A. A., Khan, S., Nasir, S., & Niaz, S. (2022). The effectiveness of installing solar panels at schools in Pakistan to increase enrolment. *Journal of Environmental Studies and Sciences*, 12(3), 505-514.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research*, 35(3), 472-482.

Hsu, H. G., & Shih, H. C. (2014). Energy Operation Network Promotion for Models of Informal Education. *International Journal of Research-GRANTHAALAYAH*, 2(1), 26-46.

Intergovernmental Panel on Climate Change (IPCC). (2023). Oceans and Coastal Ecosystems and Their Services. In Climate Change 2022 – Impacts, Adaptation and Vulnerability (pp. 379–550). *Cambridge University Press*. https://doi.org/10.1017/9781009325844.005

Izadpanahi, P., Elkadi, H., & Tucker, R. (2017). Greenhouse affect: the relationship between the sustainable design of schools and children's environmental attitudes. *Environmental Education Research*, 23(7), 901-918.

Jans, L., Koudenburg, N., & Grosse, L. (2023). Cooking a pro-veg* n social identity: the influence of vegan cooking workshops on children's pro-veg* n social identities, attitudes, and dietary *intentions*. *Environmental Education Research*, 1-16.

Jensen, B. B. (2019). Environmental and health education viewed from an action-oriented perspective: a case from Denmark. *Curriculum and Environmental Education* (pp. 277-297). Routledge.

Kendall, C. J., Leeds, A., Tinka, J., Lukas, K. E., & Folta, E. (2021). Teacher training as a means to sustained and multiplicative behavior change: An example using fuel-efficient stoves. *American Journal of Primatology*, 83(4), e23193.

Koller, K., Pankowska, P. K., & Brick, C. (2023). Identifying bias in self-reported pro-environmental behavior. *Current Research in Ecological and Social Psychology*, 4, 100087.

Kormos, C., & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: A meta-analytic review. *Journal of Environmental Psychology*, 40, 359-371.

Kranz, J., Schwichow, M., Breitenmoser, P., & Niebert, K. (2022). The (un) political perspective on climate change in education—A systematic review. *Sustainability*, 14(7), 4194.

Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3), 495–524. https://doi.org/10.1111/jeea.12006

Kuhar, C. W., Bettinger, T. L., Lehnhardt, K., Cartwright, B., & Cress, D. (2012). Education program evaluation at multiple primate sanctuaries in equatorial Africa. *International journal of primatology*, 33(1), 208-217.

LaCombe, C. M., & Danoff-Burg°, J.A. (2013). Measuring Behavioural Change in Bangladesh. *IZE Journal*, (49).

Lawson, D. F., Stevenson, K. T., Peterson, M. N., Carrier, S. J., L. Strnad, R., & Seekamp, E. (2019). Children can foster climate change concern among their parents. *Nature Climate Change*, 9(6), 458-462.

Lee, T. M., Markowitz, E. M., Howe, P. D., Ko, C. Y., & Leiserowitz, A. A. (2015). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change*, 5(11), 1014–1020. https://doi.org/10.1038/nclimate2728

Lestari, H., Ali, M., Sopandi, W., Wulan, A. R., & Rahmawati, I. (2022). The Impact of the RADEC Learning Model Oriented ESD on Students' Sustainability Consciousness in Elementary School. *Pegem Journal of Education and Instruction*, 12(2), 113-122.

Liu, Y., Dong, K., & Jiang, Q. (2023). Assessing energy vulnerability and its impact on carbon emissions: A global case. *Energy Economics*, 119, 106557.

Mattsson, P., & Laike, T. (2022). Young children's learning about lighting and turn-off behaviour in preschool environments. *Energy and Buildings*, 268, 112193.

Mikami, A., Akaishi, N., Hirayama, S., Yata, M., & Nagao, K. (2022). Development of a climate change education program using knowledge of behavioral science such as nudges and verification of its effects in decreasing CO2 at home. *International Journal of Home Economics*, 15(1), 157-170.

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J. R., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*. www.pnas.orgcgidoi10.1073pnas.0709546105

Monroe, M. C., Plate, R. R., Oxarart, A., Bowers, A., & Chaves, W. A. (2017). Identifying effective climate change education strategies: a systematic review of the research. *Environmental Education Research*, 25(6), 791–812. https://doi.org/10.1080/13504622.2017.1360842

Myers Jr*, O. E., Saunders, C. D., & Garrett, E. (2004). What do children think animals need? Developmental trends. *Environmental Education Research*, 10(4), 545-562.

National Institute for Health and Care Excellence (Great Britain). (2012). Methods for the development of NICE public health guidance. *National Institute for Health and Care Excellence*.

Nourmoradi, H., Asadi, M., Naghizadeh, M. M., Mazloomi, S., Asadi, H., & Ghazanfari, Z. (2022). The influence of an education program on students' environmental responsibility in developing countries: evidence from Iran. *Journal of Human Behavior in the Social Environment*, 32(7), 970-980.

Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W. N., Arrow, K. J., Barrett, S., Carpenter, S., Chapin, F. S., Crépin, A. S., Daily, G., Ehrlich, P., Folke, C., Jager, W., Kautsky, N., Levin, S. A., Madsen, O. J., ... De Zeeuw, A. (2016). Social norms as solutions. In Science (Vol. 354, Issue 6308, pp. 42–43). *Science*. https://doi.org/10.1126/science.aaf8317

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4), 306–312. https://doi.org/10.1038/s41558-021-01000-1

Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of economic perspectives*, 14(3), 137-158.

Parry, D. A., Davidson, B. I., Sewall, C. J., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 5(11), 1535-1547.

Pietrapertosa, F., Tancredi, M., Salvia, M., Proto, M., Pepe, A., Giordano, M., Afflitto, N., Sarricchio, G., Di Leo, S., & Cosmi, C. (2021). An educational awareness program to reduce energy consumption in schools. *Journal of cleaner production*, 278, 123949.

Prestin, A., & Pearce, K. E. (2010). We care a lot: Formative research for a social marketing campaign to promote school-based recycling. *Resources, Conservation and Recycling*, 54(11), 1017-1026.

Reid, A. (2019). Climate change education and research: possibilities and potentials versus problems and perils?. *Environmental Education Research*, 25(6), 767-790.

Rocher, M., Silva, B., Cruz, G., Bentes, R., Lloret, J., & Inglés, E. (2020). Benefits of outdoor sports in blue spaces. the case of School Nautical Activities in Viana do Castelo. *International journal of environmental research and public health*, 17(22), 8470.

Rolim, C. C., & Baptista, P. (2021). Sharing lisboa: A digital social market to promote sustainable and energy efficient behaviours. *Climate*, 9(2), 34.

Rosmaladewi, O., & Poetri, D. T. (2020). Study of Environmental-based School Model in Village Around Protected Village Areas (Case Study: Barusari Village, Indonesia). *JournalNX*, 6(06), 197-204.

Rousell, D., & Cutter-Mackenzie-Knowles, A. (2020). A systematic review of climate change education: Giving children and young people a 'voice'and a 'hand'in redressing climate change. *Children's Geographies*, 18(2), 191-208.

Sarrasin, O., Henry, J. L., Masserey, C., & Graff, F. (2022). The Relationships between Adolescents' Climate Anxiety, Efficacy Beliefs, Group Dynamics, and Pro-Environmental Behavioral Intentions after a Group-Based Environmental Education Intervention. *Youth*, 2(3), 422-440.

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *The American Journal of Evaluation*, 22(2), 127-160.

Seybold, B., Braunbeck, T., & Randler, C. (2014). Primate conservation—An evaluation of two different educational programs in Germany. *International Journal of Science and Mathematics Education*, 12, 285-305.

Shea, N. A., Mouza, C., & Drewes, A. (2016). Climate change professional development: Design, implementation, and initial outcomes on teacher learning, practice, and student beliefs. *Journal of Science Teacher Education*, 27, 235-258.

Sidiropoulos, E. (2018). The personal context of student learning for sustainability: Results of a multi-university research study. *Journal of Cleaner Production*, 181, 537-554.

Somwaru, L. (2016). The Green School: a sustainable approach towards environmental education: Case study. *Brazilian Journal of Science and Technology*, 3(1), 1-15.

Sparkman, G., Howe, L., & Walton, G. (2021). How social norms are often a barrier to addressing climate change but can be part of the solution. *Behavioural Public Policy*, 5(4), 528–555. https://doi.org/10.1017/bpp.2020.42

Sulistyarini, A., Komalasari, A., Dewantara, J. A., Prasetiyo, W. H., & Rahmanie, E. F. (2022). Futures for pro-environment and social responsibility activities in Indonesian schools: An Adiwiyata case study. *Issues in Educational Research*, 32(2), 746-764.

Sutter, M., Zoller, C., & Glätzle-Rützler, D. (2019). Economic behavior of children and adolescents–A first survey of experimental economics results. *European Economic Review*, 111, 98-121.

Tarng, W., Ou, K. L., Yu, C. S., Liou, F. L., & Liou, H. H. (2015). Development of a virtual butterfly ecological system based on augmented reality and mobile learning technologies. *Virtual Reality*, 19, 253-266.

Trott, C. D. (2020). Children's constructive climate change engagement: Empowering awareness, agency, and action. *Environmental Education Research*, 26(4), 532–554. https://doi.org/10.1080/13504622.2019.1675594

Tsai, A. Y. J., & Tan, A. Y. K. (2022). Analysis of Undergraduates' Environmentally Friendly Behavior: Case Study of Tzu Chi University Environmental Education Program. *Energies*, 15(13), 4853.

UNESCO and UNFCCC. (2016). Action for Climate Empowerment. http://www.unesco.org/open-access/terms-use-

United Nations (2023). High level panel on the development of a multidimensional vulnerability index. https://www.un.org/ohrlls/sites/www.un.org.ohrlls/files/final_mvi_report.pdf

University of Notre Dame Global Adaptation Initiative (2023). Country Index Technical Report. https://gain.nd.edu/assets/522870/nd_gain_countryindextechreport_2023_01.pdf

Warren, R., Vanderwal, J., Price, J., Welbergen, J. A., Atkinson, I., Ramirez-Villegas, J., Osborn, T. J., Jarvis, A., Shoo, L. P., Williams, S. E., & Lowe, J. (2013). Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. *Nature Climate Change*, 3(7), 678–682. https://doi.org/10.1038/nclimate1887

Yazdankhah, Z., Mehrabi, Y., Rakhshanderou, S., Safari-Moradabadi, A., & Ghaffari, M. (2020). Behavioral approach to food consumption and waste production: A quasi-experimental study. *Journal of Education and Health Promotion*, 9.

Zhang, W., Zhao, J., & Chen, J. (2019). Nature club programs promote adolescents' conservation behavior: A case study in China's biodiversity hotspot. *The Journal of Environmental Education*, 50(3), 192-207.

List of acronyms and abbreviations

AFD	Agence française de développement
CCE	Climate Change Education
ESD	Education for Sustainable Development
EVI	Energy Vulnerability Index
OECD	Organization for Economic Cooperation and Development
PAA	Priority Action Area
RADEC	Read-Answer-Discuss-Explain-Create
SR	Systematic Review
UNFCCC	United Nations Framework Convention on Climate Change
UNITE	Unite for the Environment
UNESCO	United Nations Educational, Scientific and Cultural Organization

Appendix

A.1. Elements from the search strategy

Table A.1 includes the list of existing reviews and thesaurus employed to make an initial list of keywords for our search strategy. This list of keywords is reported in Box A.1.

Table A.2 summarizes the set of keywords finally included in the search key.

Table A.1. Sources for initial set of keywords

Articles
Anderson, A. (2012). Climate change education for mitigation and adaptation. *Journal of Education for Sustainable Development*, 6(2), 191-206.
Ardoin, N. M., Bowers, A. W., & Gaillard, E. (2020). Environmental education outcomes for conservation: A systematic review. *Biological conservation*, 241, 108224.
Ardoin, N. M., Bowers, A. W., & Gaillard, E. (2023). A systematic mixed studies review of civic engagement outcomes in environmental education. *Environmental Education Research*, 29(1), 1-26.
Bhattacharya, D., Carroll Steward, K., & Forbes, C. T. (2021). Empirical research on K-16 climate education: A systematic review of the literature. *Journal of Geoscience Education*, 69(3), 223-247.
Crandon, T. J., Scott, J. G., Charlson, F. J., & Thomas, H. J. (2022). A social–ecological perspective on climate anxiety in children and adolescents. *Nature Climate Change*, 12(2), 123-131.
Jorgenson, S. N., Stephens, J. C., & White, B. (2019). Environmental education in transition: A critical review of recent research on climate change and energy education. *The Journal of Environmental Education*, 50(3), 160-171.
Kranz, J., Schwichow, M., Breitenmoser, P., & Niebert, K. (2022). The (Un) political perspective on climate change in education—A systematic review. *Sustainability*, 14(7), 4194.
McKenzie, M. (2021). Climate change education and communication in global review: Tracking progress through national submissions to the UNFCCC Secretariat. *Environmental Education Research*, 27(5), 631-651.
Monroe, M. C., Plate, R. R., Oxarart, A., Bowers, A., & Chaves, W. A. (2017). Identifying effective climate change education strategies: A systematic review of the research. *Environmental Education Research,* 25(6), 791-812.
Thesaurus
Thesaurus of Environmental Education Terms
OneLook Theasurus
GreenFILE Thesaurus database in EBSCO
CRC Thesaurus
UNESCO Thesaurus

Box A.1. Initial set of keywords

Action research, CCEC, Climate action, Climate change literacy, Climate change mitigation, Climate change prevention, Climate change research, Climate education, Climate literacy, Community education, Community education for the environment, Conservation, Conservation behavior, Conservation education, Conservation education, CRC article 24, CRC article 29, Earth science education, Ecology education, Education and development, Education for mitigation, Education for Sustainability, Education for sustainable development, Energy education, Environmental attitude, Environmental awareness, Environmental behavior, Environmental education, Environmental information, Environmental research, Environmental sciences education, Informal education, Natural resources education, Non-formal education, Population education, Public education, Science and society, Social biology, Sustainability education, Sustainable development

Table A.2. Set of keywords employed in the EBSCO systematic search

Column 1 (Environment)	Column 2 (Behavioral outcome)	Column 3 (Education)	Column 4 (Intervention)
Climate change	behavior	education	experiment
Conservation	behaviour	school	intervention
Earth science	beliefs	teaching	program
Ecology	social norm*	informal education	training
Energy	injunctive norm*	non-formal education	
Environmental awareness	descriptive norm*	public education	
Adaptation	attitudes	community education	
Mitigation			
Natural resources			
Science and society			
Sustainability			
Climate action			
Climate change literacy			
Climate change adaptation			
Climate change mitigation			
Climate change prevention			
Climate change research			
Environmental			
Environmental information			
Environmental research			
Environmental sciences			
Sustainable development			

A.2. Data extraction tool

DESCRIPTION	QUESTION
Log information	
Date when form was completed	Date when form was completed
ID of person extracting data	ID of person extracting data
Report identification	
Publication title	Title of publication
Publication ID	Same ID from the original dataset (Excel file)
Author details	Surname of first author
Publication date	Year
Publication type	What is the impact evaluation publication type? ☐ Academic journal article ☐ Research report ☐ Government report ☐ Dissertation / thesis ☐ Online book chapter
Funding agency name	Who is funding the evaluation/study? Please add name of the agency funding the evaluation.
Funding agency type	Type of agency funding the evaluation/study: ☐ Academic institution ☐ Charitable or private foundation ☐ For-profit firm ☐ Government agency ☐ International aid agency ☐ International financial institution ☐ Non-profit organization ☐ Not specified ☐ Don't know
Conflict of interest	Is there a potential conflict of interest associated with the study which could influence the collected/reported results? (e.g. Is there a declaration of conflict "of interest? Is any of the authors related in any way to the funding or implementing institution? Is the evaluation funded or undertaken by funders and/or implementers?) ☐ Yes ☐ No ☐ Not clear
Comments on conflict of interest	Please add reason for your answer to whether there is a conflict of interest.
Language of publication	Language of publication of the study (e.g. Spanish, English etc) ☐ English ☐ Spanish ☐ French ☐ German ☐ Italian

DESCRIPTION	QUESTION
Context	
Country	List countries the study was conducted in.
Detailed location	If provided, give detailed information on where the study took place within a country (e.g. regions/districts covered; schools/university).
World Bank Region	Select region(s) the study was conducted according to the World Bank. For more info on region classification see: https://datatopics.worldbank.org/sdgatlas/archive/2017/the-world-by-region.html ☐ East Asia & Pacific ☐ Europe & Central Asia ☐ Latin America & the Caribbean ☐ Middle East & North Africa ☐ North America ☐ South Asia ☐ Sub-Saharan Africa
World Bank Income category	Select the World Bank income classification of the country at the time of the study. Check here: https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html ☐ High Income ☐ Upper Middle Income ☐ Lower Middle Income ☐ Low Income ☐ Not classified
Sector	Choose sector options below: ☐ Agriculture ☐ Education ☐ Energy and extractives ☐ Forestry ☐ Financial ☐ Industry and Trade/Services ☐ Information and Communication ☐ Public Administration ☐ Transportation ☐ Water sanitation and hygiene (WASH) ☐ Environmental and disaster management ☐ Other. Specify:
Intervention information	
Data collection strategy	Select the type of data collection strategy: ☐ Laboratory experiment or lab-in-the-field ☐ Field experiment or randomized controlled trials ☐ Natural experiment ☐ Survey - One stage ☐ Survey - Before and after ☐ Observational data ☐ Other

DESCRIPTION	QUESTION
Justification study design	Brief description of the data collection strategy.
Description of Intervention(s)	Write a short paragraph to describe the intervention type and characteristics. The description should be as detailed as possible.
Objectives of intervention	State any objectives stated in study or other document.
Unit of intervention	At which level was the intervention implemented? ☐ Individual ☐ Household ☐ Cluster of individuals (e.g., classroom, team, group of visitors) ☐ School / University ☐ Community ☐ Other ☐ Unclear
Target population age	Indicate the population either : ☐ 5-10 (Grades K-5) ☐ 11-13 (Grades 6-8) ☐ 14-17 (Grades 9-12) ☐ Young adults in college (18-25) ☐ Adults (25-65) ☐ Elderly (65+) ☐ Mixed ☐ Not specified
Target population income	Indicate the target population income (relatively to the country where the study is carried out) ☐ Low ☐ Middle ☐ High ☐ Diverse ☐ Not specified
Target population living environment	State the target population living environment between ☐ Rural ☐ Urban ☐ Both
Date of intervention	Date (month-year) of intervention.
Intervention length/exposure to intervention (in months)	Intervention length ☐ One day ☐ More than one day and less than a week ☐ Between one and four weeks ☐ One or more months
Incentives	Were incentives provided to intervention participants? ☐ Yes ☐ No ☐ Not clear

DESCRIPTION	QUESTION
Personnel implementing the program	Who was in charge of implementing the program? ☐ PI/researchers (study authors) ☐ Implementing agency staff ☐ External agency (e.g. survey firm) ☐ Personnel from educational institution (different from study authors) ☐ Others (please specify) ☐ Not clear
Other personnel implementing	Specify the "Other" personnel implementing the program
Outcome information	
Outcome type	Select the outcome type:
	Normative outcomes ☐ Personal beliefs (personal opinions or beliefs about what is considered appropriate or inappropriate behavior) ☐ Normative expectations (opinions or beliefs about what the society considers appropriate or inappropriate behavior) ☐ Empirical expectations (beliefs about what others do or not do) ☐ Unclear which category ☐ No measure of beliefs/expectations
	Behavioural outcomes ☐ Start behavior ☐ Increase behavior ☐ Decrease behavior ☐ End behavior ☐ No change in behavior ☐ No measure of behavior
Impact	***Dimension of impact of the intervention*** ☐ Mitigation ☐ Adaptation ☐ Conservation
Outcome indicator description	Extract the exact name of the indicator being used as the dependent variable in the analysis. Use this open answer field to enter, in the author's own words, a description of the outcome, in a sentence or so. Be selective and concise with the excerpts being transcribed here as to ensure accurate and precise descriptions of the outcome. Include page numbers with every excerpt extracted.
Outcome timing (from end of the interventionand the point at which an outcome measure is measured)	☐ Within the day ☐ Within the week ☐ Within the month ☐ Within one year ☐ More than one year
Timing of outcome measurement	☐ Only after ☐ Before and after ☐ Not clear

DESCRIPTION	QUESTION
Effect size calculations	
Nature of the reported results	***Nature of the reported results*** ☐ Descriptive qualitative ☐ Descriptive quantitative ☐ Non-parametric analysis ☐ Regression analysis
Direction of the effect	☐ Positive effect on mitigation/adaptation/conservation ☐ Negative effect on mitigation/adaptation/conservation ☐ Zero effect ☐ Unclear
Effect is statistically significant?	☐ Yes. Please specify. ☐ No ☐ Unclear
Independent observations	Provide the number and explain (specify also how many independent observations per condition if there are multiple conditions)
Subgroup	Is there an analysis of a subgroup (e.g., gender, status)? ☐ Yes ☐ No
If yes to subgroup, describe the subgroup if applicable	Free text, describe the subgroup if applicable (e.g. boys, girls).

A.3. Critical appraisal

DESCRIPTION	RESPONSE			
	-	+	++	Can't tell
I. Quality of the protocol and writing				
Does the study address a clearly focused research question?				
Is the study protocol clearly defined? (Consider whether interventions and comparisons are described in sufficient detail to enable replication)				
Are the outcome measures (e.g., behaviors, norms) clearly defined and assessed using objective instruments and indicators? (e.g., real behavior ++ vs. self-reported measure of behavior -)				
Do outcome measures reflect what the experiment set out to measure?				
Were the methods to assess outcome measures comparable across groups (control, treatments)?				
Is the source population from which participants are selected well described? (Consider whether the country (e.g., developed or non-developed), setting (primary schools, community centers etc.), location (urban, rural), population demographics etc. are adequately described)				
Are the selected participants representative of the source population? (Consider whether the authors explain how and why the sample was chosen (i.e., identified/selected/recruited), whether some individuals did not agree to participate, whether there are other sources of selection bias, whether inclusion or exclusion criteria are applied and whether they are appropriate)				
Are participants in the control[1] group sampled from the same population as those exposed to the intervention (treatment)? The terms 'control' and 'comparison' group refer to any group with the treatment of interest is compared and is presumed to represent conditions in the absence of that treatment, whether it is true random or not.				
Is the assignment of participants to interventions randomized?				
Is the number of participants per group clearly stated?				
Is the study internally valid (i.e. unbiased)? (Consider whether there are significant flaws or sources of bias in the study design)				
Is the paper written in correct language and style? Recall that language refers to English/Spanish/French/German/Italian				
If one or more of the questions of this section were marked with (-), please provide in this space the justification. You can also put here any other relevant comment regarding the quality of the protocol or the writing.				

II. Quality of the intervention's implementation				
Were participants, outcome assessors and/or implementers (e.g. teachers) blind to the group/intervention allocation?				
Are the characteristics of participants (e.g. age, sex, socio-economic characteristics) in the different groups (e.g. control and treatment) comparable before the intervention?				
Were all participants accounted for at study conclusion? (Consider whether there was attrition, other forms of missing/incomplete data or data exclusion)				
Was contamination acceptably low? (Consider whether the intervention was taken by the controls or vice versa)				
Apart from the intervention, were each study group treated equally?				
If one or more of the questions of this section were marked with (-), please provide in this space the justification. You can also put here any other relevant comment regarding the intervention's implementation.				

III. Quality of the statistical analysis				
Do the authors explain how and why the sample size was chosen?				
Was a power calculation undertaken to determine sample size?				
Were the effects of intervention reported comprehensively? (Consider the results are reported for each outcome in each study group, whether the authors report on all variables they aimed to study, whether statistical tests are reported, whether p-values are reported)				
Are the analytical methods appropriate?				
The reported effect estimate does not look prone to selective reporting.				
Are the findings generalizable to the source population (i.e., externally valid)?				
If one or more of the questions of this section were marked with (-), please provide in this space the justification. You can also put here any other relevant comment regarding the statistical analysis.				

Agence française
de développement
5, rue Roland Barthes
75012 Paris I France
www.afd.fr

What is AFD?
Éditions Agence française de développement publishes analysis and research on sustainable development issues. Conducted with numerous partners in the Global North and South, these publications contribute to a better understanding of the challenges faced by our planet and to the implementation of concerted actions within the framework of the Sustainable Development Goals.
With a catalogue of more than 1,000 titles and an average of 80 new publications published every year, Éditions Agence française de développement promotes the dissemination of knowledge and expertise, both in AFD's own publications and through key partnerships. Discover all our publications in open access at editions.afd.fr.
Towards a world in common.

Publication Director Rémy Rioux
Editor-in-Chief Thomas Melonio

Legal deposit 4th quarter 2023
ISSN 2492 - 2846

Graphic design MeMo, Juliegilles, D. Cazeils
Layout Denise Perrin, AFD
Printed by the AFD reprography service

To browse our publications:
https://www.afd.fr/en/ressources-accueil