



A Protocol for the Review of Examples of Transformational Change in the Energy and Public Health Sectors to Inform Climate Mitigation and Adaptation Interventions

Zharas Aitmambet¹ · Joseph Dickman² · Markus Frölich¹ · Atika Pasha¹ · Martin Prowse^{3,4} · Jyotsna Puri^{5,6} · Neha Sharma² · Arne R. Weiss¹

Accepted: 19 December 2023
© The Author(s) 2024

Abstract

This article describes the approach for an evidence review that combines two different evidence gap maps into one learning exercise on transformational change. The review assesses the evidence in two sectors where there has been demonstrable progress in terms of scale, depth and permanence of change: in energy and in behavioural change in public health. The review uses causal evidence to assess how lessons about transformational change in these two sectors may inform climate change mitigation and adaptation investments in low- and middle-income countries. The evidence gap maps will inform the scope of systematic reviews and meta-analyses.

Keywords Transformational change · Systematic review · Energy · Behavioural change in public health · Scale · Depth · Sustainability

Resumen

Este artículo describe el enfoque de una revisión de pruebas. Ese enfoque combina dos mapas de brechas de evidencia diferentes, en un único ejercicio de aprendizaje alrededor del cambio transformacional. La revisión evalúa la evidencia en dos sectores donde ha habido un progreso demostrable en términos de escala, profundidad y permanencia de cambio: en energía, y en cambios de comportamiento en salud pública. La revisión utiliza evidencia causal para evaluar cómo las lecciones sobre

✉ Martin Prowse
prowsemartin@hotmail.com

¹ Center for Evaluation and Development (C4ED), Mannheim, Germany

² Climate Investment Funds, Washington, DC, USA

³ Independent Evaluation Unit, Green Climate Fund, Songdo, Incheon, South Korea

⁴ University of East Anglia, Norwich, UK

⁵ International Fund for Agricultural Development, Rome, Italy

⁶ Columbia University, New York, USA



cambio transformacional en estos dos sectores pueden informar las inversiones de mitigación y adaptación al cambio climático en países de ingresos bajos y medios. Los mapas de brechas de evidencia informarán el alcance de las revisiones sistemáticas y de meta-análisis.

Résumé

Cet article décrit l'approche d'une revue des preuves. Cet approche combine la documentation de deux schémas resumant les écarts des preuves dans un seul exercice d'apprentissage focalisé sur le changement transformationnel. La revue évalue les preuves dans deux secteurs où il y a eu un progrès démontrable en terme d'échelle, de profondeur et de permanence du changement: dans l'énergie, et dans le changement de comportement en santé publique. La revue utilise des preuves causales pour évaluer comment les leçons sur le changement transformationnel dans ces deux secteurs peuvent informer les investissements en matière d'atténuation et d'adaptation au changement climatique dans les pays à revenu faible et intermédiaire. Les cartographies des lacunes en matière de preuves informeront la portée des revues systématiques et des méta-analyses.

Introduction

The Intergovernmental Panel on Climate Change (IPCC) estimates that if greenhouse gas (GHG) emissions maintain their rise at the current business-as-usual rate, then by the end of the twenty first century the average temperature will increase by 2.6 to 4.8 °C and sea levels will rise by 0.45 to 0.82 m (Intergovernmental Panel on Climate Change 2018). The international community's response to such climate threats is not sufficient to meet the goals of the Paris Agreement. Although climate finance has risen considerably over the past years, it is still deemed too low compared to the level required for a 1.5 °C scenario (Climate Policy Initiative, 2019). For example, investments in low-carbon technologies fall short of what is required to meet the mitigation target, according to a report by the International Energy Agency (2019a, b). The same applies to insufficient adaptation finance (Global Commission on Adaptation 2019) which does not meet the needs expressed in nationally determined contributions (Neufeldt et al. 2018). To meet such targets, climate investments need to be increased and impact per dollar spent need to be considerably higher. Furthermore, the longer current mitigation and adaptation measures fall short, the higher the overall impact required by future interventions. It is therefore imperative to usher in interventions that have (1) a large impact that (2) is sustained over time and at (3) a large scale.

This article considers such changes in outcomes as being transformational not least as these three elements are common in how major multinational agencies operationalize transformational change.¹ For example, a recent publication from the

¹ For a growing literature on substantive definitions of, as well as the theoretic mechanisms behind, transformational change see Feola (2015), Few et al. (2017) and Kates et al. (2012). While these mechanisms enter neither our definition nor our analysis of the evidence for transformational change, they do inform the selection of interventions and thereby the types of studies included in our review.



Climate Investment Funds (2019) identified the dimensions of relevance, scale, systemic change and sustainability as integral to transformation. This work posits that all four dimensions must be in place (to a greater or lesser extent) for transformational change to be considered both real and lasting. As shown by recent interest, transformational change has become the Holy Grail in climate change and development assistance. While there is anecdotal evidence in brochures and examples of highly successful interventions in the academic literature, many interventions do not replicate when scaled up, or work well in one context, but fail elsewhere (Banerjee et al. 2017; Deaton 2010; Madrian 2014; Muralidharan and Niehaus 2017). In addition, there appears to be a lack of causal evidence for climate interventions as rigorous evidence has only recently started to grow (Prowse and Snilstveit 2010; Puri 2019; Ferraro 2009).

This article summarises one step that two agencies—the Green Climate Fund’s Independent Evaluation Unit and the World Bank’s Climate Investment Funds—are taking to contribute to this urgent goal. It does so by outlining our approach for an evidence review of causal evidence of transformational change and its drivers. On the one hand, the article approaches this directly by systematically reviewing the experimental and quasi-experimental literature with the potential to document transformational change across a broad set of interventions and outcomes. Here, the article focuses on the energy sector in low- and middle-income countries (LMICs) because of its key role in climate mitigation. The article also approaches this learning exercise indirectly by reviewing the evidence on behavioural change in public health. The public health literature has the longest tradition of long-term causal studies on behavioural change and the thickest evidence base using causal designs (see Nisa et al. 2019). Insights from this sector can offer insights into how to overcome the last-mile problem which so often stands in the way of realizing transformational change. The goal behind the evidence review is to assess how lessons about transformational change in energy and behavioural change in public health (in terms of interventions that led to large and sustained change at scale) may inform broader mitigation and adaptation investments. This review therefore combines, in a novel way, two different reviews into one learning exercise on transformational change with the aim of systematically mapping and meta-analysing multi-sector evidence. The primary research question guiding this review is: what are the attributes, determinants and contributors of transformational change in the energy and public health sectors? This article has six further sections. The second section offers some background for the evidence review. "[Interventions and Outcomes](#)" section outlines the intervention and outcome framework used in the review. The fourth section summarises the approach taken by the review and the methods used. "[Search Strategy](#)" section highlights the search strategy whilst "[Screening of Studies](#)" section details the screening of studies. The conclusion round off by suggesting that such cross-sector learning could contribute to understanding how we can meet one of the largest challenges in the coming decades.



Background

Due to the urgency of the climate challenge and the scale of climate finance required to limit temperature increases to close to 1.5 °C, climate finance needs to precipitate system-broad and long-term changes across sectors and societies. Yet, most interventions supporting adaptation and mitigation actions focus on what can be described as incremental changes (Termeer et al. 2017). For example, incremental interventions are often deployed in adaptation in an ex post manner and are framed as a process of adjustment instead proactive preparation for the scale of change that is required (Bassett and Fogelman 2013). Turning to mitigation, here a wider range of interventions are supporting the transition to lower-carbon societies (Markard et al. 2012) and the scale, longevity and degree innovation within such interventions are more closely aligned with transformational actions (Wienges et al. 2017). In this review, through our synthesis, we highlight robust and causal evidence across individual studies of transformational change in energy (that shows a direct connection to climate interventions) and behavioural change in the public health sector (from which lessons can be drawn to inform climate interventions). We now describe each of these two intervention areas in some detail.

Studying the energy sector in LMICs is key to future climate change mitigation efforts. Carbon dioxide, methane, nitrous dioxide and fluorinated gases are the key greenhouse gases emitted by human activities with 76% of these constituted by carbon dioxide (CO₂) emissions alone. Of overall emissions, fossil fuels and industrial processes account for 86%.² In terms of economic sectors, energy is the largest offender in terms of contribution to GHG emissions. It accounts for around 35%, including emissions that occur in the middle stages of energy production, e.g. fuel extraction, refining, processing and transportation (Pachuari and Meyer 2014). Energy contributes to the lagged, cumulative effect of greenhouse gas emissions. Such gases stay in the atmosphere for up to a century, such that on a per capita, historical basis, industrialised countries (that is, Annex 1 countries who are party to the UNFCCC) still bear the majority of the responsibility for such pollutants. That said, nearly all of the growth in energy demand, and consequently fossil-fuel use and GHG emissions, is predicted to come from LMICs (Wolfram et al. 2012). Part of this increase may in itself be driven by climate change. With rising temperatures, LMICs, for example, are expected to increase demand for residential air conditioning from 500 TWh in 2000 to around 4000 TWh in 2050 (World Energy Council 2015). The reliance of LMICs on fossil fuels for energy production means the projected increase in energy demand will, without strong countermeasures, result in even higher emissions (Ebinger and Vergara 2011a). For the period of 1994 and 2014, Falconí et al. (2019) already found considerably higher growth rates of per-capita CO₂ emissions in middle-income compared to high-income countries (HICs), with −0.2% for the latter compared to 2.8% for upper and 1.4% for lower

² Other gases are less dominant, however, still heavily present—16% of emissions are methane, 6%—nitrous dioxide and the remaining 2% consist of F-gases.



middle-income countries. Similarly, upper and lower middle-income countries have nearly 24 times (for upper) and 9 times (for lower) per-capita energy-use growth rate of HICs. The contrast between the responsibility of Annex 1 countries for historical emissions and the responsibility of non-Annex 1 countries for future emissions is why climate change is such an intractable problem. It also shows why the energy sector in LMICs is set to play such a key role for climate mitigation measures. At the same time, the energy sector itself is vulnerable to climate change. Changing precipitation and weather patterns directly affect renewable energy plants, which are dependent on natural activities—hydropower plants can suffer from drying rivers, wind power plants would produce less energy if there is a windless drought and solar panels suffer from higher precipitation and cloud cover (Ebinger and Vergara 2011a). Overall, LMICs are predicted to suffer the greatest impacts from climate change and often display limited adaptive capacity. That said, LMICs also have opportunities to implement effective mitigation and adaptation strategies. For example, according to the International Energy Agency (2019a, b), sub-Saharan Africa could achieve significant industrialisation and economic growth while keeping emissions relatively low by increasing the share of renewable energy in the energy mix. To achieve this, the IEA calls for investments on grid expansion, reinforcement and maintenance as well as on renewable energy generating capacity, in particular solar PV. As outlined in a report by the GCA (2019), investments in climate change adaptation could generate high rates of return and pay out a “triple dividend” of avoided losses, economic benefits (e.g. through reduced climate risk) as well as social and environmental co-benefits. Our review will illustrate which interventions show robust and causal evidence—across individual studies—of transformational change in energy.

This review also summarises indirect evidence on transformational change through focussing on behavioural change in public health. This is because we can use the long tradition of using causal methods in this sector to investigate the interventions that may produce large and sustained behavioural change. This review uses this tradition to highlight the key interventions that elicited sustained behaviour change in individuals within five areas—nutritional (and dietary) habits, physical activity, substance abuse, hygiene practices and utilization of health care services—due to the widespread use of behavioural science interventions within these five broad areas. So far, the relevant literature in public health, including systematic reviews, analyse either large-scale interventions or long-term effects and rarely consider behaviour change outcomes. Instead, they focus on changes in health outcomes, such as morbidity, weight, etc. (Loveman et al. 2011; Reiner et al. 2013; Zubala et al. 2017). Since health outcomes depend on many factors besides behaviour, health outcomes are poor proxies for health behaviour. Behavioural interventions may be particularly influenced with scalability or sustainability issues: A behaviour change may either not persist over long periods of time or does not espouse itself within a large swaths of the population that is targeted by the intervention (List 2022; Nisa et al. 2019). This issue becomes even more relevant given the shorter lifespan of intervention in low- and middle-income countries, where evidence is currently scarce. Consequently, our review intends to focus on studies that



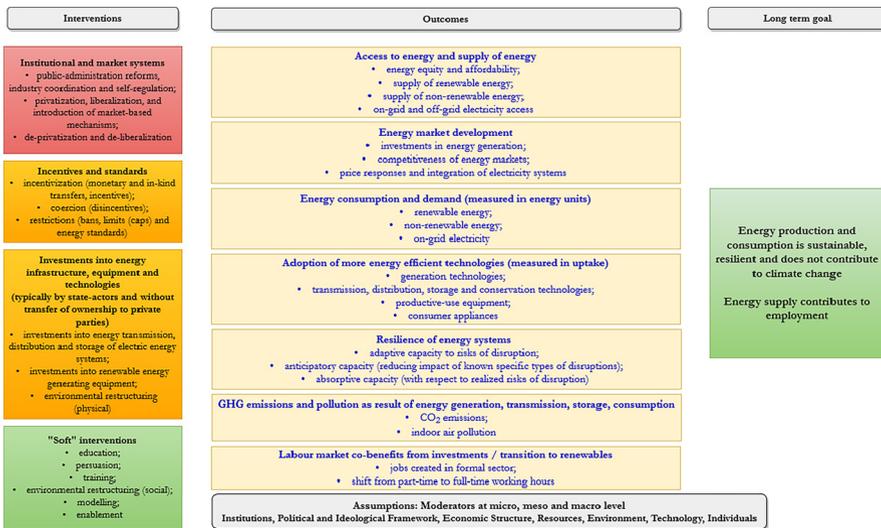


Fig. 1 Theory of change for energy. *Source* Authors

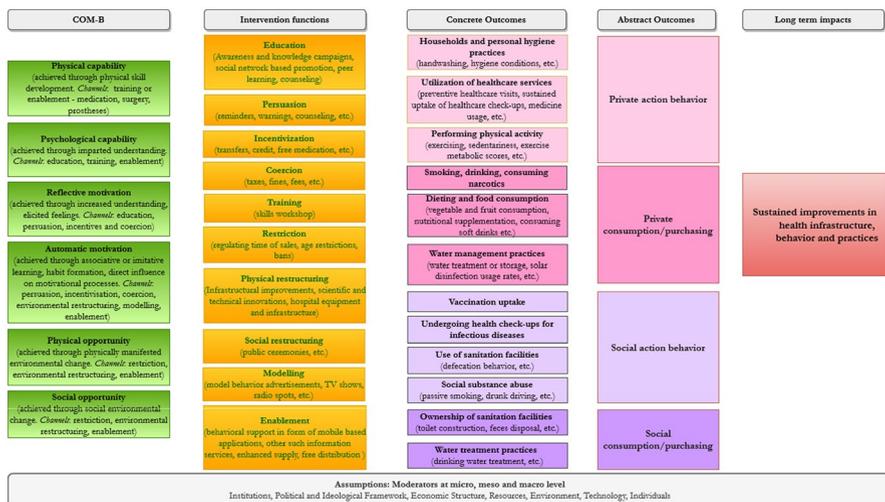


Fig. 2 Theory of change for behavioural change in public health. *Source* Authors

can help answer the question: what types of intervention yield long-term and large-scale effects on behaviour change outcomes in low and middle-income countries.

We now describe the interventions and outcomes that are covered in this review and how they are categorized within two broad theories of change (see Figs. 1, 2). These theories of change simultaneously structure and define the scope of this study.



Interventions and Outcomes

Transformational change, as such, is difficult to find directly for two reasons. First, since transformational change consists of several elements and still lacks an established definition, it is not the outcome measured in empirical studies. Instead, evidence for transformational change may be found across a wide range of possible outcomes. Second, restricting our search to studies that document transformational change, i.e. large effects, at scale and sustained over time, risks finding statistical outliers rather than an unbiased reflection of the available evidence. We therefore search for evidence across a wide range of interventions and outcomes in studies that have the potential to document transformational change—regardless of whether the individual study indeed found large effects over time. In a subsequent step, we synthesize the evidence across studies to identify those interventions that can produce transformational change.

For the energy sector, we cover a broad set of interventions that either target or could have effects for climate change mitigation and adaptation. These take place either at the level of institutional and market systems, through incentives and standards, through “soft” interventions (nudges), or in form of investments into infrastructure. Outcomes under the purview of this review capture either climate change mitigation, adaptation (resilience of energy systems), or labour-market co-benefits of investments or transition into renewable energy. These are described in more detail in the next section.

For the public health sector, and as detailed above, we include interventions targeting behavioural change in five broad areas—nutritional (dietary) habits, physical activity, substance abuse, hygiene practices and utilization of health care service—due to the use of behavioural science interventions within these five broad areas. The scope of the targeted areas for the interventions in the health sector is as shown in the inclusion and exclusion criteria (“[Appendix 1](#)”). As is common practice in systematic reviews, these already defined criteria may be further refined during the screening process in case they turn out not to be sufficiently precise for efficient and consistent coding. These interventions will then be coded following the behavioural change framework provided by Michie et al. (2011), which includes nine different intervention functions, as listed in the following section. The behavioural change framework outlines the intersection between the two sectors, to enable the comparison and cross-sectoral learning on behavioural change outcomes.

The Theories of Change also include the moderators and the assumptions which influence the overall relation between the interventions and their potential outcomes. Therefore, the existing institutions, the political and ideological framework, the economic structures, the available resources, the environmental and technological constraints, and finally the characteristics of the intervention population are all important variables that might moderate the effect or the nature of an intervention. These are therefore included within both Theories of Change. Specific assumptions on which causal chains between interventions and long-term goals rest are as follows:



- Individuals are responsive, receive the intervention as envisioned, and take up the intervention,
- Interventions are relevant for the context or have been contextualized appropriately,
- Institutions at all levels support the implementation of the interventions.

These assumptions are neither necessary nor sufficient conditions for large long-term impacts to occur. Rather, they are a set of ideal conditions that facilitate long-term effects. If one or several of these assumptions are not fulfilled, the causal chain may break down. As one example, think of an intervention in the form of an information campaign that promotes an effective vaccine. The first assumption stipulates that individuals are responsive to the information provided. However, if many individuals do not trust the content or the source of the information, then this assumption breaks down and the intervention is unlikely to be effective.³

Finally, the Theories of Change map out the outcomes and long-term goals that are targeted by the type of interventions listed above. The sector-specific Theories of Change are now described one by one.

Public Health

The purpose of the public health sector is to indirectly learn about transformational change. Hence the goal of the public health systematic review is at a higher level than any individual intervention or context in which the intervention is tested. Instead we examine the basis of theoretic mechanisms within a mid-level theory of behaviour change. Learning about theoretic mechanisms—why an intervention works—makes policy makers and development organizations much more flexible than focusing on the “what works” question, in particular when—as in our case—the learning goal targets different sectors. Knowing about the active ingredients of interventions allows to apply these to different, but theoretically similar problems, i.e. where the same obstacle to behaviour is present. The critical element for learning across concrete interventions is an appropriate mid-level theory of change, which is both sufficiently general to be transferable and detailed enough in order to be able to tackle conceptually different problems. Mid-level theories rest between high-level theories, which are too abstract to have empirical application, and project-level theories, which apply to a specific context (see Cartwright 2020; Cartwright et al. 2020; White 2023). Mid-level theory is an approach which helps assess the transferability of study findings from one setting to another.⁴

We choose the Behaviour Change Wheel (BCW; Michie et al. 2011) as the mid-level theory underlying the Theory of Change and to categorize interventions. Michie et al. (2011) rely on expert consultation as well as a review of a range of other behavioural frameworks to define a framework categorizing intervention and

³ We thank an anonymous referee for raising this point.

⁴ We thank Howard White for his input on mid-level theories of change.



policies that encompass all previous frameworks.⁵ This framework, which they call the “Behaviour Change Wheel”, groups interventions along nine intervention functions.⁶ Behind these functions lie three essential sources of behaviour change: capability, opportunity and motivation, or the COM-B system. These source functions are effectively the drivers of behaviour change, without one (or more) of these being targeted, behaviour change is not possible. Therefore, interventions targeting behaviour change in the long term and on a large scale should look towards explicitly including these sources within their design. Each of these sources is further broken into two additional categories. Within capability, we find psychological and physical capacity, to allow the individual to engage in the activity promoted/inhibited by the intervention. Similarly, without social and physical opportunity, which lies outside of the control of the individual, behaviour change might not be possible. Both capability and opportunity also provide the necessary stimulus to the cognitive processes that motivate behaviour change, either by reflection or automatically. All these sources inform the design of the intervention, as depicted within the causal chain. There are nine categories of intervention functions that are included within the Behaviour Change Wheel. They are meant to contribute towards long-term change in health behaviours, which are the targeted outcomes within the public health sector.

First are interventions under the category of education, such as awareness and knowledge campaigns, used to increase knowledge or understanding, not only to inspire a particular behaviour but also to provide knowledge about competing behaviours. The second category of interventions falls under persuasion, whereby through various methods of communication, such as reminders or warnings via phone or other ICTs, positive or negative feelings are induced to stimulate action. Incentivization in the form of monetary and in-kind rewards is the third category of interventions, meant to create reward expectations for following a particular behaviour or abstaining from it. The fourth category of interventions is coercion, the opposite of incentivization, which creates an expectation of punishment, such as by raising prices or increasing taxes. The fifth type of intervention is training, where individuals are imparted skills to encourage particular activities. Restriction, which prohibits engagement in target behaviour with the use of rules such as bans or regulated uses, is the sixth category of interventions. By discouraging competing behaviours, these can also be used to encourage a particular behaviour. Another set of interventions falls under the category of environmental restructuring, where, by modifying the physical context around an individual, such as improving infrastructure or technologies related to the targeted behaviour, behaviour change can be encouraged or discouraged. Another subset of interventions under this category captures the modification of the social context around the targeted behaviour, such as prompts that guide behaviour change. The penultimate category of interventions is modelling, where

⁵ One prominent alternative framework is MINDSPACE. MINDSPACE, though clearly useful—and much used—in its own right, mixes intervention functions and modes of delivery. By contrast, the BCW is systematically constructed and provides a mid-level theory of change.

⁶ They also categorize seven policy categories (communication/marketing, guidelines, fiscal, regulation, legislation, environmental/social planning and service provision) but these are not included in our theoretical causal chain of results.



behavioural change is stimulated by depicting what model behaviour should be. This is the method of leading by example, by showcasing the model behaviour. Finally, under the category enablement, any type of support that increases the means, reduces the barriers, or increase the capability to act on targeted behaviour (such as surgeries or prosthetics to increase physical activity) will be included. Within our review, we further divide the intervention function of environmental restructuring into its two categories, physical restructuring and social restructuring, giving us ten intervention functions in total.

Each intervention function affects one or more source functions, and thereby leads to the required modification of health behaviours, attitudes and practices, as depicted in the concrete outcomes in our Theory of Change. These interventions aim at changing behaviour in the outcome categories, which also define the scope of this sector. For the purpose of this review, we are not per se interested in all possible health outcomes, but rather in what we can learn from these health behaviours for behaviours related to climate change mitigation and adaptation. We therefore propose to define the scope of the health outcomes along the following dimensions: action/health-seeking behaviours and purchasing/consumption behaviour. These two dimensions can have a private benefit (quitting smoking), or might alternatively also affect health outcomes for other individuals (because of less exposure to passive smoke). Overall, these outcomes (and interventions) will lead us to observe sustained improvements in health behaviour, infrastructure and practices.

Energy

Many of the interventions in climate mitigation can be found in the energy sector. Due to the implementation of the Paris Agreement, 197 countries are required to have national GHG-emission reduction policies and plans for their post-2020 agenda (World Resources Institute 2018). Fostering low-carbon technologies is therefore projected to be a major issue for governments (Bouye et al. 2018). The long-term goal of the Theory of Change in the energy sector is that production and consumption is sustainable, resilient and does not contribute to climate change. Moreover, an increase in energy supply and demand also aims to contribute to higher employment, which is on the one hand a social co-benefit of energy investments but on the other hand a potential conflict with the goal of climate change mitigation. In this sector, we base the Theory of Change mainly on different assessment reports and systematic reviews concerning climate change mitigation and adaptation to it, especially the IPCC's Synthesis Report on Climate Change (Pachuari and Meyer 2014), the 3ie scoping report by Robalino et al. (2014), frameworks and reports by the World Bank Group, the International Energy Agency and the European Union Energy Initiative Partnership Dialogue Facility (Ebinger and Vergara 2011b), as well as on extensive discussions with the Climate Investment Fund's Evaluation and Learning Initiative and the Green Climate Fund's Independent Evaluation Unit.

In the spirit of Arnott et al. (2014), energy sector interventions are coded as „behavioural“ or „structural“. In light of cross-sector learnings, behavioural interventions are those that directly target behaviour change of individuals or household and



measure a behavioural outcome. These will be classified according to the Behaviour Change Wheel (BCW, Michie et al. 2011). Structural interventions are those that do not or only indirectly lead to individual or household-level behavioural change.

We group interventions into four broad categories, which, according to the ToC, will jointly contribute to achieving the long-term goals. The first category is institutional and market systems (structural), i.e. interventions that change the institutional structure of energy systems or markets. The sub-categories are public-administration reforms, industry coordination and industry self-regulation, privatization, liberalization and introduction of market-based mechanisms as well as de-privatization and de-liberalization. The second category is incentives and standards (behavioural). This category consists of three sub-categories that directly link to the behavioural framework from Michie et al. (2011), as described in the public health sector above: incentivization (such as transfers), coercion (such as taxes and fees) and restrictions (such as bans and limits). The third category is “soft” interventions, which do not change incentives (behavioural). The sub-categories therein are again taken from Michie et al. (2011): education, persuasion, training, environmental restructuring (such as social norms), modelling (such as presenting model behaviour in TV shows) and enablement (such as defaults). Lastly, the fourth category includes investments into energy infrastructure, equipment and technologies (structural). Sub-categories are investments into energy transmission, distribution and storage of electric energy systems as well as investments into renewable energy generating equipment.

These interventions may lead to outcomes grouped into seven categories. First, mainly through investments into energy infrastructure, such as grid-extension, access to energy and the supply of (renewable) energy may increase. Second, energy market development may be spurred through institutional and market-systems interventions (International Finance Corporation 2019). Third, energy consumption and demand (differentiated between renewable, non-renewable and on-grid electricity) and fourth, adoption of more energy-efficient technologies (including the transition to renewables), may change due to targeted interventions in all intervention categories. Fifth, the resilience of energy systems to climate change may increase due to investments into energy systems, such as smart grids and energy storage capacities (Ebinger and Vergara 2011a; Stuart and Escudero 2017). Sixth, as a result of incentives and standards (such as energy-efficiency standards), as well as cleaner energy supply and demand and adoption of more energy-efficient technologies (such as improved cookstoves), GHG emissions and indoor air pollution may decrease. Lastly, as a labour-market co-benefit from investments into renewable energy, formal employment may increase.

In order to facilitate cross-sector learnings from the public health sector, all behavioural outcomes within these seven outcomes will be coded in terms of whether they are action behaviours or purchasing/consumption behaviours.

Approach and Methods

To our knowledge, there appears to be an absence of systematic evidence on the causal drivers of transformational change in general, and in particular in relation to climate change mitigation and adaptation. The study that is closest to our review is



Lee et al. (2013), who systematically review the literature on organizational transformation, mainly in health care. Their definition of transformational change is, however, focussed on organizational practices, whereas we look at a broad range of outcomes. Furthermore, most included studies in their review are qualitative and thereby not able to show causal drivers of transformational change. This review will reduce this gap within the literature in order to inform governments, donors and other policy makers on the available evidence on a broad set of interventions and their effects on climate change mitigation and adaptation outcomes in the energy sector. We contribute to the literature on the drivers of transformational change in the following seven ways.

First, we discuss attributes of transformational change by offering a precise definition of transformational change (see below), which will form the basis of this review. Second, in order to learn about causal evidence on transformational change, we select only quantitative studies with an experimental or quasi-experimental study design. Furthermore, our inclusion criteria are based on a precise definition of transformational change. More specifically, we only include studies that have the potential to document transformational change according to these criteria. For instance, we intend to only include studies where data-collection was done at least 1 year after the intervention. Whether transformational change indeed happened is the empirical question to be answered through our meta-analysis. Third, while our review is broad in scope, and as illustrated above, we have a precise but extensive list of interventions and outcomes within each sector, within clearly structured categories. This allows us to search for evidence for transformational change across fields of studies while at the same time keeping the scope of the review manageable. Fourth, as a first step we provide a framework of reviewed evidence in the form of an evidence gap map (EGM) of interventions in the specified sectors. EGMs are a convenient and simple-to-use tool for policymakers to quickly inform themselves about existing evidence. This exercise will highlight where research is comprehensive and where there appears to be a lack of evidence. Moreover, it enables policymakers and practitioners to make informed decisions about project prioritization and further research activities. Fifth, we then conduct meta-analyses with the data extracted from the selected quantitative studies for sufficiently populated cells of the EGMs (i.e. at least 10 studies for the same intervention and outcome combination). This is another exercise that has not been found to be common in the literature on transformational change. Sixth, the results of the meta-analysis are important to determine where robust evidence exists, i.e. across individual studies and contexts, for transformational change. Doing so will minimize the risk that large effects of interventions are simply statistical outliers. By using the results from this meta-analysis, we intend to produce “transformational change maps” (TCMs), i.e. infographics that only show those intervention and outcome combinations where evidence for large effects at scale and over time exists. The TCMs will show the determinants of transformational change.

And seventh, in order to identify contributors of transformational change, we will search for common characteristics between populated cells in the TCMs, i.e. those intervention and outcome combinations where we find evidence for transformational change. We will also run, where applicable, meta-regressions across these cells in



the TCMs and across sectors in order to explain heterogeneity in study results. This way we might learn which characteristics of interventions which contribute to transformational change and to do so we have to operationalize it into clearly measurable criteria. The first of these is a large depth of change: transformational change requires a sizeable change. This is measured in terms of a large effect size an intervention produced on the outcomes. To define what a large impact is, we rely on previous literature that has attempted to standardize these definitions. Sawilowsky (2009) defines rule of thumb effect sizes for Cohen's d as large if $d=0.8$ (and very large if $d=1.2$ and huge for a $d=2$), based on a review of literature and contextualization of effect sizes (Cohen 1988). For relative risk, common in the medical literature, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) uses a scale separating relative risks of at least 2 as large (and those greater than 5 as very large, Guyatt et al. 2011). We will use these two definitions to define large impacts based on effect sizes in outcomes. Second, a large scale of change: Even with a large effect size, an intervention only becomes an important contributor of transformational change if it has sufficiently large scale, i.e. targeting many beneficiaries or covering large areas. Given the variety of interventions within the two sectors, we consider interventions as large scale if there are at least 1000 individual beneficiaries (effect being measured here is the treatment effect on the treated) or if they target an entire administrative area larger than a village (e.g. district, region, state). Third, sustained change. For a change to be transformational, it has to persist over time. The definition of sustained is found to vary considerably over the literature we reviewed (between 6 months and several years). In order to maintain coherence across the results, we consider an effect sustained if it persists at least 1 year after first full implementation of the intervention. Note that this is a lower bound so that impacts that arrive later than 1 year also pass this threshold. The question whether impacts are likely to wane or increase over time may be different between the two sectors. In public health, the time required to form a habit may be relatively short (Lally et al. 2010), while dis-adoption of behaviours is a strong concern. Therefore, behaviour change is unlikely to happen if it is not already present after 1 year. By contrast, many energy interventions may instead take a long time to demonstrate impacts. We therefore acknowledge that absence of large impacts in the energy sector after 1 year does not imply that large impacts may not arrive later. This is a concern for studies that measure outcomes only until 1 year after the onset of the intervention. While studying longer time-frames than 1 year would certainly be useful, setting the threshold higher would risk leaving out many studies.

The three criteria above only help finding studies that have the potential to document transformational change. Crucially, individual studies will also be selected for inclusion into the review if the evaluated intervention did not lead to large effects over time. If we only included studies with large effects over time, we would run a strong risk of picking statistical outliers instead of finding an unbiased picture of the available evidence. Study results will only become a selection criterion after meta-analysis and therefore always at the level of a group of closely related studies (with the same intervention and outcome combination). We will describe in the analysis subsection how the fulfilment of the three criteria of transformational change is reflected in the analysis that leads to the TCMs.



Inclusion/Exclusion Criteria

Following Petticrew and Roberts (2006), we use the PICOS model to precisely describe the inclusion and exclusion criteria. The tables including the summary of the inclusion and exclusion criteria for both sectors can be found in “[Appendix 1](#)”. Pilot screening may lead to adjustments in these tables to make sure that the categorization and coding of studies are sufficiently clear.

Population

We include interventions rolled out in LMICs, as defined by the current World Bank categorization (financial year 2020).⁷ Thus, we exclude studies of interventions in high-income countries or that include LMICs but do not separately report results for those. In the energy sector, we exclude interventions targeted at children (below the age of 12) because generally they are not main agents of climate change mitigation and adaptation. In the health sector, interventions that target behavioural change of adolescents or children (below the age of 18) are excluded because we study long-lasting behavioural change, about which we may learn more general lessons from adults, who have more solidified personalities than adolescents. While it would be interesting to compare adults and adolescents, this would be beyond the scope of this review.

Interventions

The types of interventions we study are informed by the sector-specific Theories of Change described above and in Figs. 1 and 2. We focus on studies which seek to evaluate the causal effect of an intervention that was purposefully implemented. We focus only on interventions that are sufficiently large in scale in order to draw meaningful conclusions. Results need to be representative of a large-scale intervention through two ways, following (Muralidharan and Niehaus 2017). First, the scale of the intervention: there need to be at least 1000 treated beneficiaries (automatically fulfilled if there are more than 1000 treated individuals in the study sample). If the number of beneficiaries are not given or in case the intervention is disseminated through radio and other media, it needs to target an entire administrative area larger than a village (e.g. district, region, state). Second, the scale of the population represented the sample of treated individuals must be representative of a sampling frame of at least 1000 treated individuals or of an administrative area larger than a village. While truly large scale (to reduce the familiar upward “bias” of small-scale interventions and studies) would mean a higher threshold than 1000 beneficiaries, it is

⁷ The population part of the search terms includes the list of all LMIC as defined by the World Bank in financial year 2020 (see “[Appendix 4](#)” section). The country of implementation is almost always named in the title, especially in the case of experimental and quasi-experimental studies that are rarely done at regional (supranational) level. For sub-national studies, keywords usually include the country name. Within the Web of Science database, we use the topic field and therefore the search is done on title, abstract and keywords.



purposefully set low initially in order not to risk the exclusion of too many studies. Depending on the number of studies passing this threshold, it might therefore still be raised later.

Comparison

We consider only quantitative studies that aim to evaluate the causal effect of an intervention on the outcome, i.e. experimental or quasi-experimental studies further defined later. We include studies that have a clearly defined comparison group for evaluation of the treatment effect. The nature of the comparison group depends on the type of research design used in the study and can include both active and passive comparison groups.

Outcomes

Since our major outcome, transformational change, cannot be directly measured, we look at a range of outcomes and measure change therein, which could reflect transformative processes in the two sectors. Our list of outcomes is described above and in Figs. 1 and 2. When baseline values are used for identifying treatment effect, then time between baseline and endline needs to be at least 1 year. While collecting data 1 year after the intervention is in many cases not sufficiently long to be certain of a sustained change, e.g. by enduring over changes in political or administrative leadership, a higher threshold may lead to the exclusion of too many studies. This threshold may be adjusted depending on the number of studies we find, possibly differentiating between types of outcomes. As an example, the 1-year threshold is likely to be too short for key behavioural outcomes in public health, such as smoking and alcohol consumption.

Study Design

Based on the research design, we categorize the studies into two major groups: (a) Experimental designs—this type of study specifically uses random assignment of intervention to the treatment group and evaluates the effect by comparing the outcome with the control group and by using an appropriate methodology; (b) Non-experimental designs—in cases when the assignment of treatment is not random, various quasi-random designs are used to evaluate the treatment effects. These methods include and will be restricted to regression discontinuity design (RDD), instrumental variable (IV), difference-in-differences (DID) and propensity score matching (PSM). For the health sector, in addition to these aforementioned methods, we will also use interrupted time series (ITS) and controlled before after (CBA), given their relevance in the health literature. Of both these design types, the finalized studies would be critically appraised in order to identify their strength and weaknesses. We will explore the possibility of using Robot Reviewer for a (semi-) automated risk of bias assessment.



Exclusion Criteria

We will exclude studies that are conducted outside the time frame of 1990–2020 and before 2000 in the public health sector or not including a separate sample from LMICs. We also exclude studies that do not attempt to evaluate causal effects of the intervention on the outcome, in particular, that do not follow the methods explained in the study design. As mentioned above, we will exclude studies that are not sufficiently large-scale or long-term (as defined before). In addition, all studies that are not included within our interventions, even if they measure relevant outcomes, or vice versa, measure relevant outcomes but are not capturing relevant interventions will also be excluded.

Search Strategy

The search strategy aims to find both published and unpublished studies. A three-stage search strategy will be utilized in this review. In the first stage, studies will be searched using text in the title, abstract and the keywords (see “Appendix 2” for a full list of databases). For searching the databases, we decided on search terms for each sector as described in the following subsection on combinations of search terms.⁸ In addition to these databases, the Cochrane, Campbell Collaboration, Collaboration for Environmental Evidence and 3ie libraries will be searched for impact evaluation studies and systematic reviews in the area of the above sectors. Further searches for grey literature in the energy sector will be conducted on institutional websites. In case there are less than 12,000 search results in the public health sector, a further search in Epistemonikos database will be conducted for impact evaluation studies and systematic reviews in the public health area. A record will be maintained describing the databases searched, the keywords used, and search results from each search engine. As the final (optional) step, in case there are less than 50 studies (in any of the sectors) selected for data extraction from the full text screening, we will do backward snowballing of the studies that have been selected from the full text screening, as a follow up on the initial search.⁹ We are planning to run searches on the most appropriate databases for published literature, and websites of agencies and research institutes for grey literature. The choice of databases was guided by relevance and comprehensiveness of their coverage of the sectoral literature as well as the technical possibility to apply advanced search filters allowing to increase search precision. We also validated the selection by running preliminary searches following the expert advice from our search expert.¹⁰ Given the scope of the review in terms

⁸ The search terms for the health sector are preliminary and are to be finalized based on the benchmark studies, and the number of final papers to be screened.

⁹ Backward snowballing is the process of identifying articles from reference lists of studies found in database searches and included into review as studies satisfying all inclusion criteria.

¹⁰ The initial list of databases was retrieved from various sources, such as <https://libguides.mit.edu/energy> and https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines. Running initial broad searches for the energy sector allowed us to exclude databases that contained mostly technical and engineering papers.



of the range of topics as well as the time period covered, we will not perform hand searches of key journals. Instead, we will run a database search in the Web of Science platform with the simplified set of search terms in the three relevant energy journals with the highest impact factors.¹¹ We will not hand search specific journals in public health, as the relevant studies for LMICs are dispersed across a large number of journals, and we expect to capture a large number of studies already through the database searches.

Search Terms

The search terms are organized in six different categories that reflect the inclusion criteria and the sector-specific theories of change. The search terms within each category are combined with the OR operator, whereas the AND operator is used to combine the different categories of search terms.

- (1) *Long-term or large-scale* this category encompasses terms used to describe studies carried out over a longer time span or at a large scale.
- (2) *Methodology* these terms capture the experimental and quasi-experimental methods (for more details see the inclusion/exclusion tables in “[Appendix 1](#)”).
- (3) *Countries* all lower- and middle-income countries as well as general terms describing LMICs are listed here.
- (4) *Interventions* terms are based on sector-specific ToCs.
- (5) *Outcomes* terms are based on sector-specific ToCs.
- (6) *Sector-specifying terms* this category contains terms used to describe the respective sectors.

The following combinations of categories are used:

- *Energy* the five categories (2–6 above) are combined through the AND operator. In case the total number of studies to be screened exceeds 7500, the long-term or large-scale terms (category 1 above) will be applied with the AND operator.
- *Public Health* the four categories (2–5 above) are combined with the AND operator. In case the total number of studies to be screened exceeds 15,000, the long-term or large-scale terms (category 1 above) will be applied with the AND operator.

The search strategies are tested against a set of benchmark studies in each sector. If more than two thirds of the benchmark papers can be retrieved through the database searches, the search strategy is deemed satisfactory. This threshold has already

¹¹ The highest impact journals relevant for this review were selected from the list available at Scimago Journal and Country Rank for energy. These more relaxed restrictions (in terms of outcomes, and long-term or large-scale) will ensure that only three of the six categories below (countries, methodology and interventions) are combined using the AND operator with the following Publication Name terms string: SO=(“energy” OR “energy economics” OR “energy journal” OR “energy policy”).



been passed in the energy sector with the total number of studies within the target of 7500 (for the list of benchmark papers, see “[Appendix 3](#)”).

In public health, the list of benchmark papers has already been determined (also see “[Appendix 3](#)”), while the search strategy is being finalized by testing over two databases: PubMed and Web of Science. Currently, 68–90% of the benchmark studies for health (9 out of 10 in the Web of Science database and 11 out of 16 in the PubMed database) were part of the search results. The number of studies found for screening in these two databases stands at 30,120 papers (12,448 in the Web of Science database and 17,672 in the PubMed database). Given this is much larger than the targeted 15,000 papers, our current strategy is to streamline the search so as to not lose benchmark papers but remove most of the non-relevant papers within the search results. Several iterations, requiring various permutations and combinations of the search categories, are required. The current list of search terms (in Boolean format) is provided in “[Appendix 4](#)”.

Screening of Studies

The screening process of the two populations of studies, which we found through the literature searches described in the previous section, will be carried out in several steps. Note that these steps will be done for each sector such that there are two separate screening processes. First, pilot screening will make sure that the coding tools are well understood or revised. Two independent screeners will each screen 200 studies. The results of pilot screening are considered satisfactory when the overlap between the inclusion decisions of both screeners after reconciliation is above 80%. Second, titles, abstracts and keywords will be screened to exclude any irrelevant studies. In order to save time given the wide scope of the literature search, this stage will be assisted by the machine-learning algorithm embedded in EPPI Reviewer 4. We propose the following procedures to achieve both speed and quality of screening. The machine learning algorithm will be fed the results of the pilot screening of 200 studies. The software will then sort the entire population of studies by prioritizing them according to relevance. The first 50% of studies, sorted by relevance, will be screened by two independent screeners. The next 25% of studies will be screened only by a single screener and the last 25% directly excluded from the review. The screening process will stop earlier in case 100 continuous studies, sorted by priority, are all excluded. Third, we will apply the specified inclusion/exclusion criteria to the full text and determine whether the study should be included for analysis. We will record all search results, including the reasons for exclusion at the full-text screening stage. These results will be presented in the PRISMA diagram. At least 20% of studies will be double-screened by a second reviewer. Disagreement will be resolved through discussion and third-member involvement.

Systematic reviews will be screened on the basis of their inclusion criteria. If the inclusion criteria of a systematic review meet all of our inclusion criteria, it passes on to data-extraction to be shown in our own EGMs. Depending on the number of studies found, either one or two people working independently will extract information from each study included in the review. In this step, data will be extracted



and summarized using a pre-piloted extraction form by two people. Disagreements in coding will be resolved through discussion and third-member involvement. The goal of the analysis is to document evidence for transformational change. The analysis will proceed in several steps for each sector as described below, with technical details following in later paragraphs.

- (1) We will use simple EGMs, with interventions listed along the *Y*-axis and outcomes along the *X*-axis, to document evidence and gaps within the scope of each sector.
- (2) We will then concentrate on the sufficiently populated cells (at least 10 individual studies) within the map to run meta-analyses on the available evidence and estimate average effect sizes.
- (3) We will then map only those combinations of interventions and outcomes where evidence of transformational change is found. That is to say, we will only show those combinations of interventions and outcomes where there is a large effect size at least 1 year after the intervention, following the thresholds defined before. It is this step where the results of the studies, i.e. depth of change and sustained change, are used as selection criteria. However, selection is not done at the level of the individual study but rather at the level of intervention–outcome combinations (cells in the EGM). Based on the simulation results of our meta-analysis expert (Frank Renkewitz), 10 studies are a lower bound to test for heterogeneity and therefore to assess the generalizability of the results. The results of this exercise will be shown in “transformational change maps” (TCMs) and discussed.
- (4) In order to identify contributors of transformational change, we will search for common characteristics between populated cells in the TCMs, i.e. those intervention and outcome combinations where we find evidence for transformational change. We will also run, where applicable, meta-regressions across these cells in the TCMs and across sectors in order to explain heterogeneity in study results. This way we might learn which characteristics of interventions contribute to transformational change.

In order to draw the EGMs, the following procedure will be applied. For the categorization of studies, we intend to follow Rankin et al. (2016) to determine the categorization of studies in the EGMs. In case several different interventions were grouped together, each intervention would be coded separately in order to be able to show all available evidence related to a particular intervention. For example, a study may look at the effects of a programme that includes a cash transfer intervention and an awareness intervention on two different outcomes. In this case, the two associated outcomes would be coded separately for each intervention. In some studies, it might be that only some elements of the programme or evaluation were relevant to this EGM (e.g. specific intervention or outcome) and only these aspects would then be extracted and coded. Systematic reviews will be coded based on the PICOS of the review. If a systematic review covers more than one intervention and outcome, it will appear in each cell that applies. In terms of generating categories based on the outcomes related to each sector, the outcomes are presented on the *x* axis of the map



(every column) and indicate a cluster of multiple studies. These categories would be generated on the basis of the outcomes as described in the sector-specific theories of change. In terms of generating categories for the interventions and outcomes related to each sector, the Y axis (each row) of the EGM lists all the specific interventions that were found as part of the review. These would be listed under the nature of the intervention. For instance, if the aim is to reduce CO₂ emissions via carbon taxation, then incentivization would be the category for that particular row. All impact evaluations/systematic reviews that use carbon taxation as an intervention would be included within that row. In terms of ranking the systematic reviews on the basis of their quality, following the categorization of each systematic review, we will sort it according to the confidence with which one can attribute the particular outcomes to the given intervention. This ranking code can be based on the SURE (2011) ranking, which was used in the Snilsveit et al. (2013) paper. The checklist ranks systematic reviews on the basis of methods that were used to identify, include and critically appraise studies in the systematic review, as well as the methods used to analyse the findings.¹² To assess the risk of bias in primary studies, we will use the Cochrane tool by Higgins et al. (2011).

We will attempt to do a meta-analysis for the studies with comparable variables and coefficients. We will make the studies comparable by calculating the same standardized effect sizes across the studies. We will also attempt to detect publication bias and subsequently run sensitivity analyses of the distribution of the effects, after comparing outcomes of different correction methods. We exclude studies that do not provide sufficient information to do this or which are not exclusively based on experimental or quasi-experimental methods. Synthesis of the evidence from the included studies will be presented through narrative and statistical analysis of comparable effect sizes using meta-analysis. Meta-analysis is useful in synthesizing quantitative evidence as it takes into account the statistical power of the estimated effect. Calculations of standardized mean difference, or the risk ratios are appropriate for similar type of treatment effects, hence they can be widely used for studies that apply randomized control trials. However, in case of quasi-experimental studies, the treatment effects may not be strictly comparable. For instance, studies that use a regression discontinuity design or instrumental variable method typically estimate local average treatment effect (LATE), while those using propensity score matching would estimate the average treatment effect on treated (ATT). Therefore, we shall conduct meta-analysis where it is possible to convert the treatment effects into comparable measures (Duvendack et al. 2012). Specifically, we shall carry out meta-analysis if the following conditions are met:

- (a) the interventions are sufficiently similar to be comparable,
- (b) the effect sizes can be computed for comparison,
- (c) the outcome measures are sufficiently similar,

¹² Besides SURE, there are other critical appraisal tools we will consider, including, among others, AMSTAR (Shea et al., 2007), Critical Appraisal Skills Programme (The Joanna Briggs Institute Critical Appraisal tools Aromataris et al., 2015 and ROBIS Whiting et al., 2016).



- (d) there are at least 10 different studies available that meet these criteria for the same intervention and outcome combination.

These results will be presented using conventional methods such as forest plots. In terms of software, we will use Stata or R for this purpose. In case there are less than 25 studies that enter meta-analysis in each sector, we will at that point explore the possibilities of combining intervention (sub-)categories (combining cells), rearranging cells so as to conduct further meta-analysis or alternative evidence aggregation methods that can provide suggestive evidence on transformational change. When meta-analysis is possible (see above), we shall test for heterogeneity across studies and assess the amount of heterogeneity by the tau statistic as well as the I^2 statistic. Tau denotes the standard error of true effect sizes in the original units, whereas I^2 measures the percentage of variability across studies that is not due to sampling error but rather to differences in study population, intervention and implementation. Thus, tau indicates the stability of an average true effect size across studies, while the I^2 allows for a rough categorization of heterogeneity (Borenstein et al. 2011). We will follow the corresponding rule of thumb that if the I^2 statistic hits the threshold of 75% then there is high heterogeneity, with 50% there is moderate heterogeneity and with 25% the extent of heterogeneity is low. We can also use the Q -statistic to test for statistical heterogeneity in the outcome variables. If high heterogeneity is present, we shall investigate what factors explain it by conducting moderator analysis, including sub-group meta-analysis and meta-regression, if possible. For sufficient statistical power in meta-regressions, we follow Borenstein et al. (2011), who recommend that each covariate (the coding of studies) contains at least ten studies. Where studies are sufficiently similar to be comparable, we will run meta-regressions across sufficiently populated cells in both EGMs. This method will enable us to examine which factors contribute to transformational change. To check if the results are sensitive to the quality of data and approaches to analysis, we shall report at sub-group based results levels, assuming at least 10 studies per sub-group, particularly based on study design. We shall use funnel plots and corresponding regression methods (Stanley and Doucouliagos 2014) and sub-group analysis comparing published versus unpublished studies to assess potential publication bias.

Conclusion

This review combines two different reviews into one learning exercise on transformational change. As part of this, we will discuss what can be learned from transformational change in the public health sector for climate change mitigation and adaptation in the energy sector. Given the focus of the health sector, those learnings will concentrate on which type of interventions may lead to, predominantly individual, behavioural change. The categorization and coding frameworks in both sectors are designed to facilitate these lessons by making intervention and outcome categorization as similar as it is possible given the different natures of the two sectors. For instance, the intervention framework by Michie et al. (2011) is applied both in the health sector and to three intervention categories in the energy sector.



Cross-sector learning will not be a statistical exercise per se, but a discussion informed by the data synthesis. We propose the following steps. First, we will discuss which determinants and contributors of transformational change identified in our review are similar between the two sectors. Second, we will discuss potential reasons for areas of conflicting evidence between the sectors. Third, assuming a larger body of evidence in the health sector and consequently gaps in evidence in the energy sector, we will discuss which determinants of transformational change in the health sector could also apply to the energy sector. This step will be guided by thinking about the theoretic mechanisms behind long-term behavioural change in the health sector (as one example: commitment devices that can narrow the divide behind an intention to exercise and actually exercising). Then, we will ask about which outcomes in the energy sector the same theoretic mechanisms may also apply and lastly which interventions are therefore promising to achieve the same transformational change in the energy sector.

Climate change is one of the most pressing global priorities of the twenty first century. To achieve the necessary mitigation and adaptation activities, transformational changes are needed across systems and individual behaviour. This joint evidence review by the Green Climate Fund—Independent Evaluation Unit and the Climate Investment Funds, completed by the Center for Evaluation and Development with the assistance of the Africa Centre for Systematic Reviews and Knowledge Translation and with the advice of the Campbell Collaboration, will map out the landscape of evidence on transformational change in two sectors (for the distribution of roles, see “[Appendix 5](#)”). The lessons from that landscape could contribute to making the globe a more habitable planet in the twenty first century and beyond.

Appendix 1: Inclusion and Exclusion Criteria

Energy

Included	Clarificatory examples of what will be captured	Excluded
1 Population		
Adults and adolescents as well as social systems in low and middle income countries. Studies combining a population in LMICs and high-income countries (HIC) if the analysis distinguishes the two samples		Interventions targeting children (under age of 12) and studies in high-income countries (HIC), combination of both LMICs and HICs if effects are not reported separately



Included	Clarificatory examples of what will be captured	Excluded
2 Interventions in the energy sector¹³		
Institutional and market systems <i>(Interventions that change the institutional structure of energy systems or markets)</i>		<ul style="list-style-type: none"> • Interventions outside the energy sector or that do not target climate change mitigation or adaptation through energy production or usage are typically excluded: e.g. afforestation programs (e.g. REDD+), labor market reforms, labor-market trainings, agronomic trials, engineering, geological, geographical, investments into infrastructure other than energy infrastructure • Economic growth is not an intervention. Studies only relevant if economic growth is a mediator of the effects of an intervention in the energy sector on a relevant energy outcome (e.g. GHG emissions) to be causally identified • Investments and distributions of small energy-generating appliances (smaller than 200 W), like solar lanterns, solar household-systems • Investments into non-renewable and nuclear
<ul style="list-style-type: none"> • Public-administration reforms, industry coordination and industry self-regulation • Privatization, liberalization, and introduction of market-based mechanisms • De-privatization and de-liberalization 	<p>Technical assistance, re-structuring of government units, changes in management practices; business associations, industry bodies</p> <p>Energy and emissions trading platforms, frameworks for private sector involvement (PPPs)</p>	
Incentives and standards		

¹³ As mentioned above, interventions are coded as “behavioral” or “structural”. Behavioral interventions fall mainly in the category of “soft” interventions. The sub-categories within this category are therefore classified according to seven of the “invention functions” of the BCW. Structural interventions fall mainly in the category of “institutional and market systems”. The remaining two categories, “incentives and standards” and “investments into energy infrastructure” are mixed, with both structural interventions and behavioral interventions expected. The sub-categories within “incentives and standards” as well as the sub-category “physical environmental restructuring” within “investments into energy infrastructure” reflect the BCW intervention functions.



Included	Clarificatory examples of what will be captured	Excluded
<ul style="list-style-type: none"> • Incentivization (monetary and in-kind transfers, incentives) 	Distribution of more energy efficient consumer appliances, subsidies, block tariffs, tax rebates, feed-in tariffs	
<ul style="list-style-type: none"> • Coercion (disincentives) 	Taxes and fees, permits, green quotas	
<ul style="list-style-type: none"> • Restrictions (bans, limits (caps) and energy standards) 	Command and control, damage control, prohibition	
“Soft” interventions		
<i>(Interventions that do not change the incentives of actors)</i>		
<ul style="list-style-type: none"> • Education 	Awareness and knowledge campaigns	
<ul style="list-style-type: none"> • Persuasion 	Reminders or warnings	
<ul style="list-style-type: none"> • Training 	Community skills building workshops on efficient energy use	
<ul style="list-style-type: none"> • Social environmental restructuring 	Social norms, peer pressure, feedback	
<ul style="list-style-type: none"> • (Role) modelling 	Model behaviour advertisements, TV shows or posters	
<ul style="list-style-type: none"> • Enablement 	Setting defaults, making options salient, creating options	
Investments into energy infrastructure, equipment and technologies (typically by state-actors and without transfer of ownership to private parties) and other physical environmental restructuring		
<ul style="list-style-type: none"> • Investments into energy transmission, distribution and storage of electric energy systems 	Batteries for storage, pumped-storage hydroelectricity	
<ul style="list-style-type: none"> • Investments into renewable energy generating equipment 	Dams for hydro-power	
<ul style="list-style-type: none"> • Other physical environmental restructuring 	Changes to the physical environment other than investments into infrastructure and distribution of appliances, such as changing traffic signals	

3 Outcomes¹⁴

¹⁴ In light of cross-sector learnings, behavioral outcomes at the level of individuals and households are coded according to the framework in health, i.e., along the four categories “action behavior—privat”, “action behavior—social”, “consumption / purchasing behavior—privat” and “consumption / purchasing behavior—social”. Behavioral outcomes are mainly expected in the categories “Energy consumption and demand”; “Adoption of more energy efficient technologies”, sub-category “consumer appliances”.



Included	Clarificatory examples of what will be captured	Excluded
<p>Access to energy and supply of energy</p>		<ul style="list-style-type: none"> ● Implementation of new business model in energy sector; engineering or technological innovations other than related to energy-efficiency, economic growth, biomass production, agricultural yields ● Carbon sequestration and (de)forestation ● Labor market outcomes that are measured in response to an intervention targeting climate-change mitigation or energy access ● Disaster risk reduction measures (that only indirectly affect energy system resilience) ● Time use other than for formal employment
<ul style="list-style-type: none"> ● Energy equity and affordability 	<p>Energy inclusiveness, energy affordability (widening of access), energy cost reductions</p>	
<ul style="list-style-type: none"> ● Supply of renewable energy (measured in units of energy) (traditional and next generation) 	<p>Generation and supply of traditional renewable energy, next generation renewable energy</p>	
<ul style="list-style-type: none"> ● Supply of non-renewable energy (measured in units of energy) 	<p>Generation and supply of energy from coal, oil, gas, LPG, LNG, kerosene, petrol, diesel, nuclear</p>	
<ul style="list-style-type: none"> ● On-grid and off-grid electricity access 	<p>Generation and supply of electricity, electricity coverage; adoption of grid access</p>	
<p>Energy market development</p>		
<ul style="list-style-type: none"> ● Investments (measured in monetary terms) ((typically private) in energy generation (renewable energy (traditional and next generation), fossil fuels, nuclear energy, electricity) 		
<ul style="list-style-type: none"> ● Competitiveness of energy markets (market power of energy suppliers, composition) 	<p>Number of suppliers of energy products and services, concentration indices, service quality standards, power outages, variation in voltage, use of technological innovations</p>	
<ul style="list-style-type: none"> ● Price responses and integration of electricity systems 	<p>Liquidity, pricing regulation and instruments, price adjustments, spatial connectivity of electricity systems, linkages within the power supply chain</p>	



Included	Clarificatory examples of what will be captured	Excluded
Energy consumption and demand for (measured in energy units)		
<ul style="list-style-type: none"> • Renewable energy (traditional and next generation) 	Consumption of and demand for traditional renewable energy, next generation renewable energy	
<ul style="list-style-type: none"> • Non-renewable energy 	Consumption of and demand for energy from coal, oil, gas, LPG, LNG, kerosene, petrol, diesel, nuclear; diesel-generators (off-grid)	
<ul style="list-style-type: none"> • On-grid electricity 	Consumption of and demand for on-grid electricity (produced by a mix of energy sources; otherwise, it is categorized under renewable or non-renewable)	
Adoption of more energy efficient technologies (measured in uptake, not in monetary units or through demand for energy)		
<ul style="list-style-type: none"> • Generation technologies 	Adoption of energy-efficient generation technologies	
<ul style="list-style-type: none"> • Transmission, distribution, storage and conservation technologies 	Adoption of energy efficient transmission, distribution, storage and conservation technologies	
<ul style="list-style-type: none"> • Productive-use equipment 	Adoption of energy efficient technologies and equipment in manufacturing, infrastructure, services	
<ul style="list-style-type: none"> • Consumer appliances 	Adoption of more energy efficient consumer appliances (lighting, transportation, cooking)	
Resilience of energy systems (adaptation)		
<ul style="list-style-type: none"> • Adaptive capacity to risks of disruption (e.g. security of energy supply through diversification of energy sources, lower energy imports—e.g., forecasting) 	Energy security, reliance on energy imports, excess generating capacity, oil, gas, and LNG storage reservoirs	



Included	Clarificatory examples of what will be captured	Excluded
<ul style="list-style-type: none"> • Anticipatory capacity (reducing impact of known specific types of disruptions through preparedness and planning—e.g., proactive action to reduce vulnerability) • Absorptive capacity (with respect to realized risks of disruption—exercised during and after a disturbance) 	<p>Energy use planning, peak energy use, smoothing of energy consumption, decentralization of energy systems, integrating energy resilience into systems planning (heat, power, transportation systems)</p> <p>Ability of households to cope with energy production side blackouts, power quality, reliability of energy systems</p>	
<p>GHG emissions and pollution as result of energy generation, transmission, storage, consumption</p>		
<ul style="list-style-type: none"> • GHG emissions (e.g. carbon capture at power plants) • Indoor air pollution (e.g. from cookstoves) 		
<p>Labor market co-benefits (jobs creation, demand for workforce) from investments into energy infrastructure, equipment and technologies</p>		
<ul style="list-style-type: none"> • Jobs created in formal sector • Shift from part-time to full-time working hours • Working hours in salaried/formal employment 	<p>Employment, unemployment, number of new jobs, local-level multiplier effects on labor market</p>	
<p>4 Study design</p>		
<ul style="list-style-type: none"> • Impact evaluation methods (experimental and quasi-experimental), Heckman selection, fixed effects with established control group • Systematic reviews 		<ul style="list-style-type: none"> • Correlation analysis and random effects without a clear control group (as often the case with time series analysis), input–output models, general-equilibrium models and other methods that are not based on control groups • Non-systematic reviews
<p>5 Publication date</p>		
<p>After 01.01.1990</p>		<p>Before 01.01.1990</p>
<p>6 Number of beneficiaries (scale)</p>		



Included	Clarificatory examples of what will be captured	Excluded
<p>Results need to be representative of a large-scale intervention, through two ways (Muralidharan and Niehaus 2017)</p> <ul style="list-style-type: none"> • Scale of intervention: at least 1000 individual beneficiaries (ToT) (automatically fulfilled if more than 1000 treated individuals in sample) <p>If number of beneficiaries not given or for radio and other media-based interventions, intervention needs to target an entire administrative area larger than villages (e.g. districts, regions, ...)</p> <ul style="list-style-type: none"> • Scale of population represented: sample of treated individuals randomly drawn from a sampling frame of at least 1000 treated individuals (ToT) or from an administrative area larger than a village <p>7 Timing of data collection</p> <p>Outcomes were measured at least 1 year after first full implementation of the causally identified intervention component of interest. When baseline values are used for identifying treatment effect, then time between baseline and endline needs to be at least 1 year</p>		<ul style="list-style-type: none"> • Interventions targeting less than 1000 beneficiaries; single household/village/community/firm interventions • Interventions targeting small group of individuals (e.g. a training program of energy sales agents in a small city) • Studies that are not representative of at least a population of 1000 treated individuals or an administrative area larger than a village <p>Outcomes are only measured for a period less than 1 year</p>

Public Health

Included	Clarificatory examples of what will be captured	Excluded
<p>1 Population</p> <p>Adults (> 18 years) in low and middle income countries. Studies combining a population in LMICs and high-income countries (HIC) if the analysis distinguishes the two samples</p>		<p>Interventions targeting children (under age of 18) and studies in high-income countries (HIC), combination of both LMICs and HICs if effects are not reported separately</p>
<p>2 Interventions</p>		



Included	Clarificatory examples of what will be captured	Excluded
<p>Interventions targeting the five broad areas of:</p> <p>(1) <i>Substance abuse</i> (2) <i>Utilization of health services</i> (3) <i>Hygiene practices</i> (4) <i>Nutrition and dietary habits</i> (5) <i>Physical activity</i></p> <p>Following the behavioural framework under Michie et al. (2011) as defined below</p> <ul style="list-style-type: none"> • Education (awareness and knowledge campaigns to promote frequent handwashing) • Persuasion (reminders or warnings per phone to not drink and drive or harms of smoking) • Incentivization (monetary and in-kind rewards such as free or subsidised medication, consultation, vaccines, etc.) • Coercion (price increases for alcohol or tobacco, or punishment) • Training (community skills building workshops for hygiene practices) • Restriction (prohibiting sales of alcohol to under 18, at only at particular times, smoking only in particular areas in a bar, etc.) • Environmental restructuring <ul style="list-style-type: none"> – Physical (improvements or creation of infrastructure to facilitate access to health services or use of health facilities or reduce unhealthy practices) – Social (setting defaults, prompts) • Modelling (model behaviour advertisements, TV shows or posters) • Enablement¹⁵ (behavioural support for smoking cessation such as mobile based applications and services, or those that encourage health check-ups) 	<p>Within each broad intervention area</p> <p>(1) <i>Substance abuse</i> (a) Narcotics and Stimulant drug use (b) Alcohol consumption (c) Smoking</p> <p>(2) <i>Utilization of health services</i> (a) Uptake of health care related services (participation in health care counseling, i.e. for ANC or PNC, vaccination, etc.) (b) Uptake of health care check-ups (HIV testing, ANC, PNC) (c) Compliance with health care service offered (use of ORS, malarial profilax, insecticide treated bednets, institutional delivery, vaccination, immunization, etc.)</p> <p>(3) <i>Hygiene practices</i> (a) Drinking water treatment and storage (b) Use of sanitation facilities (c) Handwashing and personal hygiene</p> <p>(4) <i>Nutrition and dietary habits</i> (a) Nutritional supplementation (b) Dieting (c) Nutritional equity/ food equity (d) Diet diversification (e) Food fortification</p> <p>(5) <i>Physical activity</i> (a) Exercise or fitness (b) Sedentariness</p>	<ul style="list-style-type: none"> • Within each of the five broad intervention areas (1) <i>Substance abuse</i> (a) Violence due to substance abuse (2) <i>Utilization of health services</i> (a) Hospital/health care quality improvement (b) Insurance uptake (3) <i>Hygiene practices</i> (a) Food storage (b) Waste disposal and treatment (c) Menstrual hygiene (4) <i>Nutrition and dietary habits</i> (a) Agricultural and related food fortification (biofortification) (5) <i>Physical activity</i> (a) Exercise among athletes • Natural interventions (those that are not in the control of humans) such as due to sudden climate related shocks/natural disasters/migration • Laboratory/clinical trial/field lab interventions targeting a hospital, clinic or laboratory (e.g. quality of care, PBF, health professional training/education) • Following policy changes from governmental authorities <ul style="list-style-type: none"> (1) (De)Regulation/decentralization/privatization/simplification of procedures (deregulation and liberalization of health sector) (2) Governmental change/laws or regulation (legislature, bills or policies, coordination of government at different levels—national, sub-national, etc.)

3 Outcomes

¹⁵ Capability beyond training and education; opportunity beyond environmental restructuring.



Included	Clarificatory examples of what will be captured	Excluded
<p>Outcome related to the five sectors of substance abuse, utilisation of health services, hygiene practices, nutrition and dietary practices and physical activity will be categorised under</p>		
<p>Action behavior (actions taken by individual to improve their health status)</p>		
<ul style="list-style-type: none"> • Social (HIV testing, open defecation, drunk driving, passive smoking, etc.) • Private (ANC, PNC, institutional delivery, handwashing, etc.) 		
<p>Consumption/purchasing (purchasing drugs or alcohol, nutritious food)</p>		
<ul style="list-style-type: none"> • Social (toilet construction, etc.) • Private (expenditure on drugs, alcohol bought, spending on tobacco/cigarettes, etc.) 		
4 Study design		
<ul style="list-style-type: none"> • Impact evaluation methods (experimental and quasi-experimental), Heckman selection, fixed effects with established control group • Systematic reviews 		<ul style="list-style-type: none"> • Correlation analysis and random effects, input–output models, general-equilibrium models and other methods that are not based on control groups • Time series analysis of indicators at the aggregated (macro) level • Non-systematic reviews
5 Publication date		
After 01.01.2000		Before 01.01.2000
6 Number of beneficiaries (scale)		
<p>Results need to be representative of a large-scale intervention, through two ways (Muralidharan and Niehaus 2017)</p>		<ul style="list-style-type: none"> • Interventions targeting less than 1000 beneficiaries; single household/village/community/firm interventions • Interventions targeting small group of individuals (e.g. a training program of energy sales agents in a small city) • Studies that are not representative of at least a population of 1000 treated individuals or an administrative area larger than a village
<ul style="list-style-type: none"> • Scale of intervention: at least 1000 individual beneficiaries (ToT) (automatically fulfilled if more than 1000 treated individuals in sample) 		
<p>If number of beneficiaries not given or for radio and other media-based interventions, intervention needs to target an entire administrative area larger than villages (e.g. districts, regions, ...)</p>		
<ul style="list-style-type: none"> • Scale of population represented: sample of treated individuals randomly drawn from a sampling frame of at least 1000 treated individuals (ToT) or from an administrative area larger than a village 		
7 Language		



Included	Clarificatory examples of what will be captured	Excluded
English		Other languages
8 Timing of data collection		
Outcomes were measured at least 1 year after first full implementation of the causally identified intervention component of interest. When baseline values are used for identifying treatment effect, then time between baseline and endline needs to be at least 1 year		Outcomes are only measured for a period less than 1 year

Appendix 2: Database Searches

Energy Sector

Databases

- Academic Search Complete (via EBSCO)
- CAB Abstracts (via EBSCO)
- EconLit (via EBSCO)
- GreenFILE (via EBSCO)
- Web of Science (Social Sciences Citation Index, Science Citation Index

Expanded, Emerging Sources Citation Index)

- World Bank eLibrary (via EBSCO)¹⁶

Websites of agencies and research institutes

- African Development Bank (AfDB) <https://www.afdb.org/en>
- Asian Development Bank <https://www.adb.org/>
- Campbell Collaboration <https://campbellcollaboration.org/>
- Collaboration for Environmental Evidence <https://www.environmentalevidence.org/>
- International Initiative for Impact Evaluation: 3ie Development Evidence Portal <https://developmentevidence.3ieimpact.org/>
- National Bureau of Economic Research <https://www.nber.org/>
- Swedish International Development Cooperation Agency (SIDA)¹⁷

¹⁶ We will conduct the search in Ideas RePEc and World Bank eLibrary databases assuming that the technical problem on the EBSCO database will be solved by EBSCO. We are in contact with the EBSCO technical support service on this matter.

¹⁷ We included SIDA as a bilateral agency website because during preliminary searches we identified it as having potentially relevant impact evaluation studies. We excluded websites of other shortlisted bilateral agencies due to a lack of relevant studies.



Public Health

Databases

- EconLit (via EBSCO)
- Global Health (CAB-Ovid)
- Medline (Ovid)
- Web of Science (Social Sciences Citation Index)

Websites of agencies and research institutes

- Campbell Collaboration <https://campbellcollaboration.org/>
- Cochrane
- + Central Registry of Controlled Trials
- + Cochrane Database of Systematic Reviews <https://www.cochranelibrary.com/cdsr/>
- Collaboration for Environmental Evidence <https://www.environmentalevidence.org/>
- International Initiative for Impact Evaluation: 3ie Development Evidence Portal <https://developmentevidence.3ieimpact.org/>

Appendix 3: Benchmark Papers

Benchmark Studies in the Energy Sector

Agurto Adrianzén, M. 2013. Improved cooking stoves and firewood consumption: Quasi-experimental evidence from the Northern Peruvian Andes. *Ecological Economics* 89: 135–143. <https://doi.org/https://doi.org/10.1016/j.ecolecon.2013.02.010>.

Akpanđjar, G., and C. Kitchens. 2017. From darkness to light: The effect of electrification in Ghana, 2000–2010. *Economic Development and Cultural Change* 66 (1): 31–54. <https://doi.org/https://doi.org/10.1086/693707>.

Dinkelman, T. 2011. The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review* 101 (7): 3078–3108. <https://doi.org/https://doi.org/10.1257/aer.101.7.3078>.

Du, L., Y. He, and J. Yan. 2013. The effects of electricity reforms on productivity and efficiency of China's fossil-fired power plants: An empirical analysis. *Energy Economics* 40: 804–812. <https://doi.org/https://doi.org/10.1016/j.eneco.2013.09.024>.

Gao, H., and J. Van Biesebroeck. 2014. Effects of deregulation and vertical unbundling on the performance of China's electricity generation sector: Effects of deregulation and vertical unbundling. *The Journal of Industrial Economics* 62 (1): 41–76. <https://doi.org/https://doi.org/10.1111/joie.12034>.

Grogan, L. 2016. Household electrification, fertility, and employment: Evidence from hydroelectric dam construction in Colombia. *Journal of Human Capital* 10: 109–158. <https://doi.org/https://doi.org/10.1086/684580>.

Grogan, L. 2018. Time use impacts of rural electrification: Longitudinal evidence from Guatemala. *Journal of Development Economics* 135: 304–317. <https://doi.org/https://doi.org/10.1016/j.jdeveco.2018.03.005>.



Grogan, L., and A. Sadanand. 2013. Rural electrification and employment in poor countries: Evidence from Nicaragua. *World Development* 43: 252–265. <https://doi.org/https://doi.org/10.1016/j.worlddev.2012.09.002>.

Hanna, R., E. Dufo, and M. Greenstone. 2016. Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *American Economic Journal: Economic Policy* 8 (1): 80–114. <https://doi.org/https://doi.org/10.1257/pol.20140008>.

Khandker, S.R., H.A. Samad, R. Ali, and D.F. Barnes. 2012. Who benefits most from rural electrification? Evidence in India. Policy Research Working Paper Series 6095, 1–39. The World Bank.

Lenz, L., A. Munyehirwe, and J. Peters. 2015. *Does large scale infrastructure investment alleviate poverty? Impacts of Rwanda's Electricity Access Roll-Out Program*. <https://doi.org/https://doi.org/10.4419/86788636>.

Lin, B., and X. Chen. 2018. Is the implementation of the Increasing Block Electricity Prices policy really effective? Evidence based on the analysis of synthetic control method. *Energy* 163: 734–750. <https://doi.org/https://doi.org/10.1016/j.energy.2018.08.178>.

Lipscomb, M., A.M. Mobarak, and T. Barham. 2013. Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil. *American Economic Journal: Applied Economics* 5 (2): 200–231. <https://doi.org/https://doi.org/10.1257/app.5.2.200>.

Salmon, C., and J. Tanguy. 2016. Rural electrification and household labor supply: Evidence from Nigeria. *World Development* 82: 48–68. <https://doi.org/https://doi.org/10.1016/j.worlddev.2016.01.016>.

Sen, A. 2012. Diversity in unity: An empirical analysis of electricity deregulation in Indian states. *The Energy Journal* 33 (1). <https://doi.org/https://doi.org/10.5547/ISSN0195-6574-EJ-Vol33-No1-4>.

Urpelainen, J., and S. Yoon. 2017. Can product demonstrations create markets for sustainable energy technology? A randomized controlled trial in rural India. *Energy Policy* 109: 666–675. <https://doi.org/https://doi.org/10.1016/j.enpol.2017.07.036>.

van de Walle, D., M. Ravallion, V. Mendiratta, and G. Koolwal. 2015. Long-term gains from electrification in rural India. *The World Bank Economic Review* 19(1): 1–24. <https://doi.org/https://doi.org/10.1093/wber/lhv057>.

Yao, X.-L., Y. Liu, and X. Yan. 2014. A quantile approach to assess the effectiveness of the subsidy policy for energy-efficient home appliances: Evidence from Rizhao, China. *Energy Policy* 73: 512–518. <https://doi.org/https://doi.org/10.1016/j.enpol.2014.06.010>.

Benchmark Studies in the Public Health Sector

Arnold, B., B. Arana, D. Mäusezahl, A. Hubbard, and J.M. Colford. 2009. Evaluation of a pre-existing, 3-year household water treatment and handwashing intervention in rural Guatemala. *International Journal of Epidemiology* 38 (6): 1651–1661. <https://doi.org/https://doi.org/10.1093/ije/dyp241>.



Banerjee, A.V., E. Duflo, R. Glennerster, and D. Kothari. 2010. Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *The British Medical Journal* 340: 1–9. <https://doi.org/https://doi.org/10.1136/bmj.c2220>.

Biran, A., W.-P. Schmidt, K.S. Varadharajan, D. Rajaraman, R. Kumar, K. Greenland, et al. 2014. Effect of a behaviour-change intervention on handwashing with soap in India (SuperAmma): A cluster-randomised trial. *The Lancet Global Health* 2 (3): 145–154. [https://doi.org/https://doi.org/10.1016/S2214-109X\(13\)70160-8](https://doi.org/https://doi.org/10.1016/S2214-109X(13)70160-8).

Boudreaux, C., P. Chanthala, and M. Lindelow. 2014. Assessing the elimination of user fees for delivery services in Laos. *Public Library of Science ONE* 9 (3): 1–8. <https://doi.org/https://doi.org/10.1371/journal.pone.0089784>.

Chao, J., Y. Wang, H. Xu, Q. Yu, L. Jiang, L. Tian, et al. 2012. The effect of community-based health management on the health of the elderly: A randomized controlled trial from China. *BioMed Central Health Services Research* 12 (1): 449. <https://doi.org/https://doi.org/10.1186/1472-6963-12-449>.

Dong, W., J. Gao, Z. Zhou, R. Bai, Y. Wu, M. Su, et al. 2018. Effects of China's urban basic health insurance on preventive care service utilization and health behaviors: Evidence from the China Health and Nutrition Survey. *Public Library of Science ONE* 13 (12): 1–14. <https://doi.org/https://doi.org/10.1371/journal.pone.0209890>.

Galiani, S., P. Gertler, N. Ajzenman, and A. Orsola-Vidal. 2016. Promoting handwashing behavior: The effects of large-scale community and school-level interventions. *Health Economics* 25 (12): 1545–1559. <https://doi.org/https://doi.org/10.1002/hec.3273>.

Guiteras, R., J. Levinsohn, and A.M. Mobarak. 2015. Encouraging sanitation investment in the developing world: A cluster-randomized trial. *Science* 348 (6237): 903–906. <https://doi.org/https://doi.org/10.1126/science.aaa0491>.

Hoddinott, J., I. Ahmed, A. Ahmed, and S. Roy. 2017. Behavior change communication activities improve infant and young child nutrition knowledge and practice of neighboring non-participants in a cluster-randomized trial in rural Bangladesh. *Public Library of Science ONE* 12 (6): e0179866. <https://doi.org/https://doi.org/10.1371/journal.pone.0179866>.

Hotz, C., C. Loechl, A. Brauw, P. de Eozenou, D. Gilligan, M. Moursi, et al. 2012. A large-scale intervention to introduce orange sweet potato in rural Mozambique increases vitamin A intakes among children and women. *British Journal of Nutrition* 108 (1): 163–176. <https://doi.org/https://doi.org/10.1017/S0007114511005174>.

Jafar, T.H., J. Hatcher, N. Poulter, M. Islam, S. Hashmi, Z. Qadri, et al. 2009. Community-based interventions to promote blood pressure control in a developing country. *Annals of Internal Medicine* 151 (9): 1–11. <https://doi.org/10.7326/0003-4819-151-9-200911030-00004>.

Khetan, A., M. Zullo, A. Rani, R. Gupta, R. Purushothaman, N.S. Bajaj, et al. 2019. Effect of a community health worker-based approach to integrated cardiovascular risk factor control in India: A cluster randomized controlled trial. *Global Heart* 14 (4): 355–365. <https://doi.org/https://doi.org/10.1016/j.gheart.2019.08.003>.



Mendis, S., S.C. Johnston, W. Fan, O. Oladapo, A. Cameron, and M.F. Faramawi. 2010. Cardiovascular risk management and its impact on hypertension control in primary care in low-resource settings: A cluster-randomized trial. *Bulletin of the World Health Organization* 88 (6): 412–419. <https://doi.org/https://doi.org/10.2471/BLT.08.062364>.

Mosler, H.-J., S.M. Kraemer, and R.B. Johnston. 2013. Achieving long-term use of solar water disinfection in Zimbabwe. *Public Health* 127 (1): 92–98. <https://doi.org/https://doi.org/10.1016/j.puhe.2012.09.001>.

Nguyen, Q.N., S.T. Pham, V.L. Nguyen, L. Weinehall, S. Wall, R. Bonita, and P. Byass. 2012. Effectiveness of community-based comprehensive healthy lifestyle promotion on cardiovascular disease risk factors in a rural Vietnamese population: A quasi-experimental study. *BioMed Central Cardiovascular Disorders* 12 (1): 56. <https://doi.org/10.1186/1471-2261-12-56>.

Patil, S.R., B.F. Arnold, A.L. Salvatore, B. Briceno, S. Ganguly, J.M. Colford, and P.J. Gertler. 2014. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *Public Library of Science Medicine* 11 (8): 1–17. <https://doi.org/https://doi.org/10.1371/journal.pmed.1001709>.

Appendix 4: Final Search Terms for the Public Health Sector for the Web of Science Database

Search	Results #	Search expression
# 1	1,428,140	(TS=("long-term" OR "long term" OR longitudinal OR "over time" OR "multiple wave*" OR ((two OR three OR four OR five OR six OR seven OR eight OR nine OR ten) NEAR/0 wave*) OR transformat* OR prolonged OR wane* OR waning OR sustain* OR unsustain* OR "not sustainable" OR ("year*-long" NEAR/0 (stud* OR data*)) OR ("month*-long" NEAR/0 (stud* OR data*)) OR ((panel OR longitudinal) NEAR/2 (study OR analysis OR data)) OR ((paradigm* OR "ground-breaking" OR "ground breaking") NEAR/2 (shift* OR chang*)) OR large-scale OR largescale OR "large scale" OR year* OR month* OR time* OR "long-period" OR "long time period" OR "over a long period" OR "long run" OR "long-run" OR "follow*up")) AND LANGUAGE: (English) Indexes = SSCI Timespan = 2000–2020



Search Results # Search expression

- # 2 474,013 (TS=(("quasi experiment*" OR quasi-experiment* OR quasiexperiment* OR "random* control* trial*" OR "random* trial*" OR RCT OR randomi* OR (matching NEAR/2 (study OR procedure OR "using" OR use* OR observable*)) OR "propensity score" OR psm OR "regression discontinuity" OR "regression kink" OR "fuzzy regression" OR "sharp regression" OR "discontinuous design" OR rdd OR "difference in difference*" OR "difference-in-difference*" OR "diff in diff" OR "diff-in-diff" OR (random* NEAR/1 (allocat* OR assign* OR select*)) OR "research synthesis" OR "fixed effect*" OR "synthetic control" OR "rapid evidence assessment*" OR "systematic literature review*" OR "systematic* review*" OR metaanaly* OR "meta analy*" OR meta-analy* OR "control* evaluation" OR "control* treatment" OR "instrumental variable*" OR (as NEAR/2 instrument) OR (heckit NEAR/2 (model* OR estimat* OR procedure OR method)) OR (heckman* NEAR/5 (sample OR selection OR model OR correction)) OR ((treatment OR intervention OR comparison OR control OR subsidy) NEAR/0 group) OR ((counterfactual OR "counter factual" OR "counterfactual" OR random*) NEAR/2 (stud* OR analysis OR experiment*)) OR ((counterfactual OR "counter factual" OR "counter-factual" OR random*) NEAR/2 (outcome*)) OR causal* OR "control group*" OR "comparison group*" OR ((control OR treatment) NEAR/0 (communit* OR village*)) OR (experiment* NEAR/1 (stud* OR analysis OR design*)) OR IV OR ITT OR ((treatment OR intervention) NEAR/2 effect*) OR "intention-to-treat" OR "intention to treat" OR ("econometric analysis") OR (impact* NEAR/1 (evaluation OR stud*)) OR "controlled before?and?after" OR "controlled before?after" OR "quasi?experimental time series" OR "interrupted time series") NOT (granger OR "kuznets curve" OR "unidirectional causality" OR "uni-directional causality" OR "bidirectional causality" OR "cointegrated equation" OR cointegration OR "panel causality" OR "co-integration" OR "causality test*" OR "wavelet coherence" OR "spatial econometric" OR nexus OR "response surface" OR "choice experiment*" OR ((root OR "cross-section* dependence" OR "cross-sectional augmented") NEAR/0 test))))
AND LANGUAGE: (English)
Indexes = SSCI Timespan = 2000–2020
- # 3 561,959 (TS=((Africa OR Caribbean OR "West Indies" OR "Middle East" OR "Central America" OR "Pacific Islands" OR Micronesia OR Polynesia OR Melanesia) OR (Asia NOT (Japan OR Korea OR "Hong Kong" OR Hong-Kong)) OR ("South America" OR "Latin America")) OR (Afghanistan OR Albania OR Algeria OR "American Samoa" OR Angola OR Argentina OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Byelarus OR Byelorussian OR Belarus OR Belorussian OR Belorussia OR Belize OR Benin OR Bhutan OR Bolivia OR Bosnia OR Herzegovina OR Hercegovina OR Botswana OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Faso" OR "Upper Volta" OR Burundi OR Urundi OR "Cabo Verde" OR Cambodia OR "Khmer Republic" OR Kampuchea OR Cameroon OR Camerouns OR Cameron OR Camerons OR "Cape Verde" OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Mayotte OR Congo OR Zaire OR "Costa Rica" OR "Cote d'Ivoire" OR "Côte d'Ivoire" OR "Ivory Coast" OR Cuba OR Djibouti OR "French Somaliland" OR Dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR "Equatorial Guinea" OR Eritrea OR "Eswatini" OR Ethiopia OR Fiji OR Gabon OR "Gabonese Republic" OR Gambia OR Gaza OR Georgia OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR



Search Results # Search expression

Guatemala OR Guinea OR Guiana OR Guyana OR "Guinea-Bissau" OR
Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Jamaica
OR Jordan OR Kazakhstan OR Kazakh OR Kenya OR Kiribati OR Kosovo
OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kir-
gizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Basutoland
OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy
Republic" OR Malaysia OR Malaya OR Malay OR Sabah OR Sarawak OR
Malawi OR Maldives OR Mali OR "Marshall Islands" OR Mauritania OR
Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova
OR Moldovia OR Moldovian OR Mongolia OR Montenegro OR Morocco
OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR
Namibia OR Nauru OR Nepal OR "Netherlands Antilles" OR Nicaragua
OR Niger OR Nigeria OR Muscat OR Pakistan OR Palestine OR Paraguay
OR Peru OR Philippines OR Philipines OR Phillipines OR Phillipines OR
"Papua New Guinea" OR Romania OR Rumania OR Roumania OR Russia
OR Russian OR Rwanda OR Ruanda OR "Saint Lucia" OR "St Lucia"
OR "St. Lucia" OR "Saint Vincent" OR "St Vincent" OR "St. Vincent"
OR Grenadines OR Samoa OR "Samoan Islands" OR "Navigator Island*"
OR "Sao Tome" OR "São Tomé and Príncipe" OR Senegal OR Serbia
OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia
OR Sudan OR Suriname OR Surinam OR Swaziland OR "South Africa"
OR Syria OR Syrian OR Tajikistan OR Tadjikistan OR Tadjikistan OR
Tadzhik OR Tanzania OR Thailand OR Togo OR "Togolese Republic" OR
Tonga OR Tunisia OR Turkey OR Turkmenistan OR Turkmen OR Tuvalu
OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New
Hebrides" OR Venezuela OR Vietnam OR "Viet Nam" OR "West Bank"
OR Gaza OR Yemen OR Zambia OR Zimbabwe) OR ((developing OR
"less* developed" OR "less-developed" OR "under developed" OR under-
developed OR "middle income" OR "middle-income" OR "low* income"
OR "low*-income" OR underserved OR "under served" OR deprived OR
poor*) NEAR/0 (countr* OR nation OR nations OR population* OR world
OR state*) OR (((developing OR "less* developed" OR "under developed"
OR underdeveloped OR "middle income" OR "low*-income" OR "low*
income" OR underserved OR "under served" OR deprived OR poor*)
NEAR/0 (economy OR economies)) OR (low* NEAR/0 (gdp OR gnp OR
"gross domestic" OR "gross national")) OR (low NEAR/3 middle NEAR/3
countr*) OR (lmic OR lmics OR "third world" OR "lami countr*" OR
"global south") OR "former soviet" OR "post-soviet" OR "commonwealth
of independent states" OR "non-OECD" OR ((transition* OR cis) NEAR/0
(countr* OR state* OR economy OR economies))))

AND LANGUAGE: (English)

Indexes = SSCI Timespan = 2000–2020



Search Results # Search expression

4 1,530,495 (TS = (((prevent* OR health OR primary OR community OR "peer group" OR group) NEAR/2 (care OR service OR program* OR session OR educat* OR re-educat* OR reeducat* OR intervention OR train* OR retrain* OR re-train* OR check* OR knowledge OR support)) OR ((physical* OR game* OR leisure* OR fitness OR wellness OR health OR care) NEAR/2 (event* OR setting* OR program* OR venue* OR site* OR center OR centre OR check OR check-up OR checkup)) OR ((prevent* OR intervention* OR campaign* OR initiative*) NEAR/2 (diabetes OR obesity OR cardiac)) OR ((aquatic OR resistance OR physical) NEAR/2 (training OR exercis* OR exert*)) OR ((lifestyle OR "life style" OR life-style) NEAR/2 (intervention* OR change OR improv* OR better* OR campaign*)) OR ((activity OR movement OR fitness) NEAR/2 (track* OR sens* OR monitor*)) OR pedometr* OR step count* OR mhealth OR (self NEAR/2 (help* OR manag* OR monitor* OR track*)) OR (walk* OR run* OR jog* OR swim* OR danc* OR garden* OR cycl* OR bicycl* OR bike* OR recreation*) OR ((food* OR diet* OR nutritio* OR nutrient*) NEAR/2 (choice OR mediterranean OR health* OR balance* OR fat OR fats OR salt* OR sugar* OR unhealthy OR therapy OR polic* OR diversi* OR balanc* OR prepar*)) OR ((beverage* OR drink* OR liquid*) NEAR/2 (sweet OR sweetened OR carbonated OR cola OR sugar OR caloric OR energy)) OR ((physical* OR game* OR leisure* OR fitness) NEAR/2 (event* OR setting* OR program* OR venue* OR site* OR center OR centre)) OR ((media OR community OR school OR family OR parent*) NEAR/2 (intervention* OR program* OR campaign* OR initiative*)) OR ((lifestyle OR "life style" OR life-style) NEAR/2 (intervention OR change OR improv* OR better*)) OR ((food NEAR/2 (ration* OR supplement* OR fortif*)) NEAR/2 (program* OR intervent* OR campaign* OR initiative*)) OR ((vitamin* OR mineral* OR iodin* OR iron OR zinc OR micronutrient* OR nutrient*) NEAR/2 (suppl* OR capsule* OR inject* OR deficiency*)) OR (food NEAR/2 label*) OR ((weight OR "weight control" OR "weight reduction") NEAR/2 (program* OR intervent* OR campaign* OR initiative*)) OR (water NEAR/2 (drink* OR provide OR provis* OR filter* OR sanitiz* OR sanitis* OR purifi* OR treat* OR guard OR manage* OR disinfect* OR steriliz* OR sterilis* OR boil* OR sedimentation OR biofilter* OR "anti-bacterial agent*" OR antimicrobial*)) OR "sodium hypochlorite" OR SODIS OR "water management" OR sanitizer OR sanitiser OR "sanitary engineering" OR ("household water treatment" NEAR/2 "safe storage") OR hwts OR "water safety plan*" OR "water supply" OR ((chemical OR heat) NEAR/2 treatment) OR ((ultraviolet OR UV) NEAR/2 (radiation OR treatment)) OR (chlorine NEAR/2 (dispenser OR filter)) OR ((sanitation OR handwash* OR "WASH" OR hwipc OR "toilet construction" OR "sewage construction" OR hygiene OR wastewater) NEAR/2 (facilit* OR station OR intervention* OR program* OR campaign* OR initiative*)) OR "pour toilet*" OR "flush toilet*" OR "pit latrine*" OR "composting toilet*" OR "on?site system*" OR "off?site system*" OR sewerage OR "septic tank*" OR "Community Led Total Sanitation" OR CLTS OR "WASH" OR ("water sanitation" NEAR/2 hygiene) OR ((water OR sanitation) NEAR/2 hygiene) OR ("Participatory Hygiene" NEAR/2 "Sanitation Transformation") OR SARAR OR "Urban Led Total Sanitation" OR "community approach*" OR "supply side improvements" OR "hygiene promot*" OR "water closet*" OR ((hygiene OR mother OR mothers OR health) NEAR/2 club*) OR (("nicotine replacement" OR "smoking cessation" OR replacement OR anti-smoking



Search Results # Search expression

OR antismoking OR no-smoking OR "no smoking" OR "non smoking" OR non-smoking OR nonsmoking OR drug OR anti-drug OR "substance abuse" OR alcohol) NEAR/2 (therap* OR intervention* OR program* OR initiative* OR campaign* OR counsel* OR treat* OR polic*) OR ((smoking OR tobacco OR cigar* OR bidi* OR beedi* OR hooka* OR waterpipe* OR kretek* OR shisha* OR chutta* OR dhumti* OR hookli* OR chillum* OR alcohol OR liquor* OR drug* OR marijuana OR cocaine OR crack OR heroin OR meth OR methamphetamine OR amphetamine OR narcotic OR opium OR MDMA OR LSD) NEAR/2 (restriction* OR ban* OR prevent* OR polic* OR prohibit*) OR (nicotine NEAR/2 (patch* OR spray* OR inhaler* OR lozenge* OR gum*)) OR bupropion OR ((label* OR pack* OR packet* OR package*) NEAR/2 warn*) OR ((behavioral OR cessation) NEAR/2 (support OR aid OR therap* OR instruct*)) OR "cigarette* price" OR ((indoor OR workplace OR work-place OR office OR hospital OR employee*) NEAR/2 ((restriction* OR ban* OR polic* OR prohibit*) NEAR/2 smok*)) OR (("door to door" OR home OR clinic OR "preventive healthcare" OR monthly) NEAR/2 (visit OR checkup* OR check-up* OR test OR tests OR testing)) OR ((health OR healthcare OR "health care" OR medical) NEAR/2 (availab* OR provision* OR provid* OR promot* OR prevent* OR barrier* OR constrain* OR imped* OR facilitat* OR hinder* OR block* OR obstacle OR restrict* OR optimiz* OR optimis* OR adher* OR access* OR motivat* OR accept* OR availabl*)) OR ("insecticide treated" NEAR/2 (net* OR bednet* OR "bed net*)) OR "insurance provi*" OR "facility based delivery" OR "institutional* delivery" OR "village council meeting" OR (elimination NEAR/2 ("user fee" OR fee OR charge)) OR checkup* OR ((vac-cin* OR immuniz* OR innocul* OR "antenatal care" OR ANC OR "postnatal care" OR PNC OR health OR "health care" OR healthcare OR medical) NEAR/2 (therap* OR intervention* OR program* OR initiative* OR campaign* OR counsel* OR treat* OR polic* OR camp)) OR (train* NEAR/2 ("GPs" OR "service provider" OR "health worker" OR midwife)) OR ((education* OR information*) NEAR/2 (campaign* OR session* OR poster* OR leaflet* OR counseling)) OR (technical NEAR/2 information) OR "home based counselling" OR stickers OR broadcasting OR leaflets OR meetings OR "individual advice" OR "social mobilization" OR advocacy OR advocat* OR "behavior change communication" OR "family based home health education" OR ((messag* OR SMS OR "short message service" OR email* OR e-mail* OR "electronic mail*" OR television OR tv OR televised OR radio OR newspaper OR movie OR in-store OR "in store" OR magazine* OR internet OR web OR print) NEAR/2 (campaign* OR commercial OR commercials* OR display OR displays OR retail OR store OR "point of purchase" OR "point-of-purchase" OR "point of sale" OR "point-of-sale")) OR "cash transfer" OR ((monetary OR financial) NEAR/2 incentive*) OR "tax reform" OR subsid* OR voucher* OR microcredit OR micro-credit* OR loan* OR financ* OR price* OR prices OR tax OR taxes OR taxation OR training OR "technical information" OR "capacity building" OR restrict* OR forbid* OR warning* OR ban OR bans OR banning OR prohibiti* OR ordinance OR ((mobile OR "smart phone" OR smartphone OR phone OR cellphone OR "cell phone" OR tablet* OR electronic) NEAR/2 (app OR apps OR application* OR messag*)) OR community motivation OR "Participatory Rural Appraisal" OR "Community Action Planning" OR remind* OR prompt* OR cue* OR cuing OR advertis* OR brand* OR marketing OR mass distribution campaign OR enable* OR infrastrucur* OR "mass media campaign" OR "social media campaign" OR stickers OR broadcasting OR leaflet* OR promot* "social marketing" OR "role modelling" OR "technical assistance" OR "behavioral support session" OR "behavioral support" OR "self help material" OR aid)) AND LANGUAGE: (English)
 Indexes = SSCI Timespan = 2000–2020



Search #	Results #	Search expression
# 5	1,484,963	(TS=((physical* NEAR/2 (activ* OR inactiv* OR fit* OR train* OR exertion OR "effort exert*" OR "functioning score" OR condition)) OR "activ*" OR "walk" OR "walking" OR "run" OR "running" OR "jog" OR "jogging" OR "exercise" OR "fitness" OR (gym* NEAR/2 member*) OR "cycling" OR "sedentar*" OR "metaboli*" OR "inactiv*" OR (lifestyle NEAR/2 (change OR health*)) OR "sport*" OR ((sedentar*" OR "sitting" OR "seat*" OR "lying" OR "reclin*" OR "recumben*" OR screen OR computer OR TV OR television) NEAR/2 time) OR "screentime" OR "self track*" OR "fitness track*" OR (keep* OR cardio* OR aerobic OR fitness) NEAR/2 (fit* OR activ* OR train* OR "motor activity" OR "exercise" OR "nutrit*" OR "supplement*" OR "feed*" OR "nutrient" OR "intake" OR "diet*" OR "meal frequency" OR "vegetable consumption" OR "leafy" OR "fruit-based" OR "food fortification" OR "vitamin intake" OR "glucose tolerance" OR "soft drink consumption" OR "lycn knowledge" OR "appetite regulation" OR "eating" OR "energy consumption" OR "portion size" OR "ingestion" OR "*food" OR "(food OR fruit* OR vegetable* OR salt* OR fat* OR sugar*) NEAR/2 (intake* OR consum* OR eat* OR ate) OR "overweight" OR "over weight" OR "obesity" OR "overeat*" OR "over eat" OR "(waist* NEAR/2 (circumference* OR measur*)) OR "(weight OR bmi OR body mass index) NEAR/2 (gain* OR loss* OR lose* OR lost OR change*)) OR "body weight" OR "(diet* NEAR/2 (chang* OR education OR behavio*r OR pattern*)) OR (water NEAR/2 (increas* OR intake* OR consum*)) OR ("drinking water" NEAR/2 (increas* OR intake* OR consum* OR decreas* OR reduction)) OR "undernutrition" OR "undernourish*" OR "under-nutrition*" OR "under-nourish" OR ((sugar* OR fizzy OR carbonated OR cola) NEAR/2 (beverage* OR drink*)) OR (liquid* NEAR/2 carbohydrate*) OR cordial* OR "hygien*" OR "sanita*" OR "wash*" OR "open defecation" OR "water treatment" OR "water disinfect*" OR "water quality" OR "handwash*" OR "hand-washing" OR ((toilet OR latrine) NEAR/2 usage) OR "SODIS" OR "handwashing" OR "hand disinfection" OR "hand cleansing" OR "handscrubbing" OR "chlorine test" OR ((faeces" OR "feces" OR faecal" OR "defecat*" OR "excrement*" OR "human waste" OR "night soil" OR "excreta") NEAR/2 ("dispos*" OR "manag*")) OR "drinking water" OR "drinkwater" OR "soap" OR "water quality" OR "water supply" OR "water contamination" OR "water storage" OR "water source" OR "point of use" OR "environmental health" OR "water access" OR "potable water" OR (infection NEAR/2 (control OR prevention OR management)) OR "cross infection" OR "waste disposal" OR "adaptive behavior" OR "behaviour change" OR "adaption" OR "adoption" OR "adopting" OR "rejecting" OR "rejection" OR "observance" OR "conformity" OR "compliance" OR "diffusion" OR "substance abuse" OR "substance*" OR "drug abuse" OR "drug*" OR "smok*" OR "alcohol*" OR "drunk" OR "cigar*" OR "narcotic*" OR "stimulant*" OR "bidis" OR "drink*" OR (alcohol OR drink* OR cigarette OR tobacco) NEAR/2 (consumption OR use OR abuse OR misuse* OR intoxicat* OR harmful OR excess* OR binge* OR heavy OR temperance OR abstinence OR exposure OR quit OR quitting OR reduction OR stop OR stopping)) OR "smoke free" OR "smoke-free" OR "abstinence" OR (smoking NEAR/2 (cessation OR quit OR quitting OR reduction OR stop OR stopping)) OR "tobacco smoke" OR "second hand smoke" OR "nicotine replacement" OR "demand cigarette" OR "waterpipe smoking" OR "addiction" OR "addictive behavior" OR "tobacco" OR "nicotine" OR "temperan*" OR "drinking behavior" OR "drink*" OR "beer" OR "wine" OR "ethanol" OR addict*



Search	Results #	Search expression
		<p>OR recidivism OR "ANC" OR "ante-natal care" OR "antenatal care" OR "postnatal care" OR "post-natal care" OR "PNC" OR "institutional deliver*" OR "deliver*" OR "hospital delivery" OR "childbirth*" OR "child birth*" OR "bed net*" OR "bednet*" OR "itn use" OR "oral rehydration solution" OR (hiv NEAR/2 (test* OR learn*)) OR "check-up*" OR "checkup*" OR "check up*" OR "vaccin*" OR "immuni*" OR "deworm*" OR ((medicine* OR "medical device*" OR condom* OR "therapy session*") NEAR/2 (use OR usage OR purchas*)) OR "visitation" OR "counsel*" OR (("health care" OR healthcare OR "health service*") NEAR/2 (utilization OR utilisation OR entry)) OR inocul* OR innocul* OR "knowledge practice" OR (patient NEAR/2 (complan* OR adheren* OR dropout* OR attrition OR capacitance)) OR "treatment refusal" OR ((drug OR medication) NEAR/2 adherence) OR ltfu OR "loss* to follow-up" OR "self-directed exercise*" OR "self directed exercise*" OR "uptake" OR "take up" OR "satisfaction" OR "compliance" OR "comply" OR "comple*" OR "refus*" OR "modif*" OR "adaptation" OR "adaptive behavior" OR "behaviour change" OR "adaption" OR "adoption" OR "adopting" OR "rejecting" OR "rejection" OR "observance" OR "conformity" OR "diffusion" OR (knowledge NEAR/2 attitudes NEAR/2 practice)))</p> <p>AND LANGUAGE: (English)</p> <p>Indexes = SSCI Timespan = 2000–2020</p>
#6	12,448	<p>#5 AND #4 AND #3 AND #2 AND #1</p> <p>Indexes = SSCI Timespan = 2000–2020</p>

Appendix 5: Distribution of Main Roles

Project lead, conception, funding and quality control

Professor Dr. Jyotsna Puri	Independent Evaluation Unit, Green Climate Fund
Dr. Martin Prowse	Independent Evaluation Unit, Green Climate Fund
Joseph Dickmann	Climate Investment Funds
Neha Sharma	Climate Investment Funds
Eirini Maria Pitta	Climate Investment Funds

Project execution

Center for Evaluation and Development (C4ED)	
Professor Dr. Markus Frölich	Quality assurance and technical support
Dr. Arne R. Weiss	Lead on the project execution and lead on the energy sector
Dr. Atika Pasha	Lead on the public health sector
Zharas Aitmambet	Project management
Consultants of the Africa Centre for Systematic Reviews and Knowledge Translation, Makerere University, Uganda	
Dr. Ekwaro A. Obuku	Africa Centre, Lead Consultant
Moses Ocan	Africa Centre, Supporting Consultant



Robert Apunyo	Africa Centre, Project Manager
Thomas Katairo	Africa Centre, Research Associate
External support	
Dr. Frank Renkewitz	University of Erfurt, technical support on systematic review, meta-analysis and detection of publication bias
John Eyers	Information specialist, support in designing search strategies for multiple databases
Sarah Khan	University of Göttingen, technical support on systematic review, specifically health sector
Internal GCF IEU review	
Emma De Roy	Independent Evaluation Unit, Green Climate Fund
Third-party review	
Howard White	Campbell Collaboration
Ashrita Saran	Campbell Collaboration

Acknowledgements The authors, who are listed alphabetically, are grateful for the comments received on this approach paper from Howard White, Ashrita Saran, Eirini Pitta, Emma De Roy, Frank Renkewitz, John Eyers, Sarah Khan and Ekwaro A. Obuku. Any errors or inconsistencies are entirely our responsibility.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arnott, B., L. Rehackova, L. Errington, F.F. Sniehotta, J. Roberts, and V. Araujo-Soares. 2014. Efficacy of behavioural interventions for transport behaviour change: Systematic review, meta-analysis and intervention coding. *International Journal of Behavioral Nutrition and Physical Activity* 11: 133.
- Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukerji, M. Shotland, and M. Walton. 2017. From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives* 31: 73–102. <https://doi.org/10.1257/jep.31.4.73>.
- Bassett, T.J., and C. Fogelman. 2013. Déjà vu or something new? The adaptation concept in the climate change literature. *Geoforum* 48: 42–53. <https://doi.org/10.1016/j.geoforum.2013.04.010>.
- Borenstein, M., L.V. Hedges, J.P. Higgins, and H.R. Rothstein. 2011. *Introduction to meta-analysis*. Hoboken: Wiley.
- Bouye, M., S. Hammeling, and N.-S. Schuld. 2018. *Connecting the dots: Elements for a joined-up implementation of the 2030 Agenda and Paris Agreement*. UNI Global Compact.
- Cartwright, N. 2020. *Using middle-level theory to improve programme and evaluation design*. CEDIL Methods Brief. Oxford: CEDIL. Available at: <https://doi.org/10.51744/CMB1>
- Cartwright, N., L. Charlton, M. Juden, T. Munslow, and R.B. Williams. 2020. Making predictions of programme success more reliable. CEDIL Methods Working Paper 1. Oxford: CEDIL. Available at: <https://doi.org/10.51744/CMWPI>



- Climate Investment Funds. 2019. *Evaluation of transformational change in the climate investment funds*. Climate Investment Funds.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates.
- CPI, A. Clark, A. Falconer, R. Macquarie, C. Meattle, R. Tolentino, and C. Wetherbee. 2019. *Global landscape of climate finance 2019*. San Francisco: CPI.
- Deaton, A. 2010. Instruments, randomization, and learning about development. *Journal of Economic Literature* 48: 424–455.
- Duvendack, M., J.G. Hombrados, R. Palmer-Jones, and H. Waddington. 2012. Assessing ‘what works’ in international development: Meta-analysis for sophisticated dummies. *Journal of Developmental Effectiveness* 4: 456–471. <https://doi.org/10.1080/19439342.2012.710642>.
- Ebinger, J., and W. Vergara. 2011a. *Climate impacts on energy systems: Key issues for energy sector adaptation* (No. 60051). Washington, DC: The World Bank.
- Ebinger, J., and W. Vergara. 2011b. *Climate impacts on energy systems: Key issues for energy sector adaptation* (No. 60051). Washington, DC: The World Bank.
- Falconí, F., R. Burbano, J. Ramos-Martin, and P. Cango. 2019. Toxic income as a trigger of climate change. *Sustainability* 11: 2448. <https://doi.org/10.3390/su11082448>.
- Feola, G. 2015. Societal transformation in response to global environmental change: A review of emerging concepts. *Ambio* 44: 376–390. <https://doi.org/10.1007/s13280-014-0582-z>.
- Ferraro, P.J. 2009. Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation* 122: 75–84.
- Few, R., D. Morchain, D. Spear, A. Mensah, and R. Bendapudi. 2017. Transformation, adaptation and development: Relating concepts to practice. *Palgrave Communications* 3: 1–9. <https://doi.org/10.1057/palcomms.2017.92>.
- Global Commission on Adaptation. 2019. *Adapt now: A global call for leadership on climate resilience*. Groningen: Global Commission on Adaptation.
- Guyatt, G.H., A.D. Oxman, S. Sultan, P. Glasziou, E.A. Akl, P. Alonso-Coello, D. Atkins, R. Kunz, J. Brozek, V. Montori, R. Jaeschke, D. Rind, P. Dahm, J. Meerpohl, G. Vist, E. Berliner, S. Norris, Y. Falck-Ytter, M.H. Murad, and H.J. Schünemann. 2011. GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology* 64: 1311–1316. <https://doi.org/10.1016/j.jclinepi.2011.06.004>.
- Intergovernmental Panel on Climate Change. 2018. *IPCC Sixth Assessment Report*. Geneva: IPCC.
- International Energy Agency. 2019a. *Securing investments in low-carbon power generation sources*. Paris: OECD. <https://doi.org/10.1787/53bce64c-en>.
- International Energy Agency. 2019b. *World Energy Outlook 2019*. Paris: IEA.
- International Finance Corporation. 2019. *AIMM Sector Framework Brief: Power (Brief), AIMM Sector Frameworks Consultation*. Washington, DC: International Finance Corporation.
- Kates, R.W., W.R. Travis, and T.J. Wilbanks. 2012. Transformational adaptation when incremental adaptations to climate change are insufficient. *Proceedings of the National Academy of Sciences of USA* 109: 7156–7161. <https://doi.org/10.1073/pnas.1115521109>.
- Lally, P., C.H.M. van Jaarsveld, H.W.W. Potts, and J. Wardle. 2010. How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology* 40: 998–1009. <https://doi.org/10.1002/ejsp.674>.
- Lee, J.-H., W. Fulp, K.J. Wells, C.D. Meade, E. Calcano, and R. Roetzheim. 2013. Patient navigation and time to diagnostic resolution: Results for a cluster randomized trial evaluating the efficacy of patient navigation among patients with breast cancer screening abnormalities, Tampa, FL. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0074542>.
- List, J.A. 2022. *The voltage effect*. London: Penguin.
- Loveman, E., G.K. Frampton, J. Shepherd, J. Picot, K. Cooper, J. Bryant, K. Welch, and A. Clegg. 2011. The clinical effectiveness and cost-effectiveness of long-term weight management schemes for adults: A systematic review. In *NIHR Health Technology Assessment programme: Executive Summaries*. NIHR Journals Library. <https://www.ncbi.nlm.nih.gov/books/NBK56859/>.
- Madrian, B.C. 2014. Applying insights from behavioral economics to policy design. *Annual Review of Economics* 6: 663–688. <https://doi.org/10.1146/annurev-economics-080213-041033>.
- Markard, J., R. Raven, and B. Truffer. 2012. Sustainability transitions: An emerging field of research and its prospects. *Research Policy* 41 (6): 955–967. <https://doi.org/10.1016/j.respol.2012.02.013>.



- Michie, S., M.M. van Stralen, and R. West. 2011. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science* 6: 42. <https://doi.org/10.1186/1748-5908-6-42>.
- Muralidharan, K., and P. Niehaus. 2017. Experimentation at scale. *Journal of Economic Perspectives* 31: 103–124. <https://doi.org/10.1257/jep.31.4.103>.
- Neufeldt, H., G.S. Martinez, A. Olhoff, C.M.S. Knudsen, and K.E. Dorkenoo. 2018. *The Adaptation Gap Report 2018*. Nairobi: United Nations Environment Programme (UNEP).
- Nisa, C.F., J.J. Bélanger, B.M. Schumpe, and D.G. Faller. 2019. Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nature Communications* 10 (1): 4545. <https://doi.org/10.1038/s41467-019-12457-2>.
- Pachuari, R.K., and L.A. Meyer. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Petticrew, M., and H. Roberts, eds. 2006. *Systematic reviews in the social sciences*. Oxford: Blackwell Publishing Ltd.
- Prowse, M., and B. Snilstveit. 2010. Impact evaluation and interventions to address climate change: A scoping study. *Journal of Development Effectiveness* 2: 228–262. <https://doi.org/10.1080/19439341003786729>.
- Puri, J. 2019. Transformational change: The challenge of a brave new world. In *Sustainable global value chains*, ed. M. Schmidt, D. Giovannucci, D. Palekhov, and B. Hansmann, 305–326. Cham: Springer. https://doi.org/10.1007/978-3-319-14877-9_17.
- Rankin, M., J. Jarvis-Thiebault, N. Pfeifer, M. Engelbert, S. Yoon, and A. Heard. 2016. *Adolescent sexual and reproductive health: An evidence gap map*. International Initiative for Impact Evaluation.
- Reiner, M., C. Niermann, D. Jekauc, and A. Woll. 2013. Long-term health benefits of physical activity—A systematic review of longitudinal studies. *BMC Public Health* 13 (1): 813. <https://doi.org/10.1186/1471-2458-13-813>.
- Robalino, J., A. Abarca, C. Sandoval, and K. Downs. 2014. Effectiveness of adaptation measures to climate change and climate variability: A scoping review.
- Sawilowsky, S.S. 2009. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods* 8: 597–599. <https://doi.org/10.22237/jmasm/1257035100>.
- Snilstveit, B., M. Vojtkova, A. Bhavsar, and M. Gaarder. 2013. *Evidence Gap Maps—A tool for promoting evidence-informed policy and prioritizing future research*. Washington, DC: World Bank.
- Stanley, T.D., and H. Doucouliagos. 2014. Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods* 5: 60–78. <https://doi.org/10.1002/jrsm.1095>.
- Stuart, C., and S. Escudero. 2017. *Energy and climate change adaptation in developing countries*. European Union Energy Initiative Partnership Dialogue Facility (EUEI PDF).
- SURE. 2011. *SURE checklist for making judgements about how much confidence to place in a systematic review*. SURE.
- Termeer, C.J.A.M., A. Dewulf, and G.R. Biesbroek. 2017. Transformational change: Governance interventions for climate change adaptation from a continuous change perspective. *Journal of Environmental Planning and Management* 60 (4): 558–576. <https://doi.org/10.1080/09640568.2016.1168288>.
- White, H. 2023. The use of middle-level theory in CEDIL-funded research studies. CEDIL Lessons Learned Paper 1, CEDIL, Oxford. Available at <https://doi.org/10.51744/LLP1>
- Wienges, S., et al. 2017. *Guidance note on how to design transformational mitigation projects, a checklist for project developers and managers*. Copenhagen: UNEP DTU Partnership.
- Wolfram, C., O. Shelef, and P. Gertler. 2012. How will energy demand develop in the developing world? *Journal of Economic Perspectives* 26: 119–138. <https://doi.org/10.1257/jep.26.1.119>.
- World Energy Council. 2015. *2015 World Energy Issues Monitor: Energy price volatility, the new normal*. London: World Energy Council.
- World Resources Institute. 2018. *Climate access indicators tool (CAIT): Paris contributions map*. Washington, DC: World Resources Institute.
- Zubala, A., S. MacGillivray, H. Frost, T. Kroll, D.A. Skelton, A. Gavine, N.M. Gray, M. Toma, and J. Morris. 2017. Promotion of physical activity interventions for community dwelling older adults: A systematic review of reviews. *PLoS ONE* 12 (7): e0180902. <https://doi.org/10.1371/journal.pone.0180902>.



Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

