Review

Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design

Benjamin K. Sovacool[a,b,*], Jonn Axsen[c], Steve Sorrell[a]

[a] Science Policy Research Unit (SPRU), University of Sussex Businees School, University of Sussex, United Kingdom
[b] Center for Energy Technologies, Department of Business Development and Technology, Aarhus University, Denmark
[c] School of Resource and Environmental Management, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

ARTICLE INFO

ABSTRACT

A series of weaknesses in creativity, research design, and quality of writing continue to handicap energy social science. Many studies ask uninteresting research questions, make only marginal contributions, and lack innovative methods or application to theory. Many studies also have no explicit research design, lack rigor, or suffer from mangled structure and poor quality of writing. To help remedy these shortcomings, this Review offers suggestions for how to construct research questions; thoughtfully engage with concepts; state objectives; and appropriately select research methods. Then, the Review offers suggestions for enhancing theoretical, methodological, and empirical novelty. In terms of rigor, codes of practice are presented across seven method categories: experiments, literature reviews, data collection, data analysis, quantitative energy modeling, qualitative analysis, and case studies. We also recommend that researchers beware of hierarchies of evidence utilized in some disciplines, and that researchers place more emphasis on balance and appropriateness in research design. In terms of style, we offer tips regarding macro and microstructure and analysis, as well as coherent writing. Our hope is that this Review will inspire more interesting, robust, multi-method, comparative, interdisciplinary and impactful research that will accelerate the contribution that energy social science can make to both theory and practice.

Slippery, indistinct, elusive, complex, diffuse, messy, textured, vague, unspecific, confused, disordered, emotional, painful, pleasurable, hopeful, horrific, lost, redeemed, visionary, angelic, demonic, mundane, intuitive, sliding and unpredictable.

- Professor John Law, describing the practice of social science research methods [1].

1. Introduction

It is surely a "fool's errand" to try to define quality research in academia, especially in a field as diverse as energy social science—a term which we use to describe the broad set of literatures that apply social science disciplines, perspectives and approaches to the study of energy, including production, distribution, conversion and consumption. Studies in this area draw upon concepts, methods and theories from a range of specializations and aim to produce insights that are relevant to many social problems. For energy social science is not only a collection of disciplines, but also a social or epistemic community of scholars, a compendium of methods or ways of doing research, a collection of related concepts or theories, and a wide set of interrelated topics.

Clearly, with such diversity and complexity, there is no one-size-fits-all approach, no "ten easy steps to quality". However, there are practices and guidelines that can improve the quality of research, and increase the probability of positive impact. And the applied and socially-relevant nature of the field is all the more reason to be sure that published research answers useful research questions, is rigorous, and is effectively communicated. In an effort to encourage improvements in research practice, this Review aims to review and provide guidelines for enhancing quality under the headings of novelty, rigor, and style.

The field of energy social science aims to address some of our most urgent and threatening global problems. For example, the International Energy Agency (IEA) estimates that, if society is to have a reasonable (> 66%) chance of avoiding dangerous climate change, global energy-related carbon emissions must peak by 2020 and fall by more than 70% over the next 35 years, despite growing populations and increasing affluence around the world [2]. Such deep decarbonisation will require transformational changes in most of the systems on which industrial

---

society depends [3–5]. At the same time, society must address other challenges such as air and water pollution [6,7], fuel poverty [8–11], energy insecurity [12–14] and energy injustice [15–17].

With so much on the line, it is worthwhile to pause and reflect on the state of research—are we producing high-quality studies and are they contributing to the solution of these real-world problems? A number of recent papers across fields as diverse as energy, buildings, transportation, sustainability, the life sciences and geography have asked similar questions, arguing that while social sciences must play a larger role in research on these issues [18,19], this research also needs to improve in terms of rigor (depth), interdisciplinary reach (breadth), policy-relevance, and the communication of results [20–34].

Unfortunately, evidence suggests that energy social science research is falling short of the social goal of promoting effective decarbonisation and frequently falling short of the professional goal of excellence. For a start, many published studies do not make novel contributions to the literature, have uninteresting (or poorly written) research questions, and do not rigorously apply a research design or method. In their survey of sustainability science, Brandt et al. noted that methods were often chosen based on the researcher's familiarity or specialization, rather than the method's suitability for a chosen research question [35]. Schmidt and Weight further observe that, within energy studies more broadly, interdisciplinary work remains rare: "despite the predominately socio-economic nature of energy demand, such interdisciplinary viewpoints – albeit on the rise – are still the minority within energy-related research" [36]. More generally, an independent review of the Research Excellence Framework in the United Kingdom noted that the academic community needed to deliver far more "game-changing" research that was both policy relevant and high quality [37]. Other more severe critics have attacked academia for publishing "nonsense" or "utterly redundant, mere quantitative 'productivity'" - owing in part to the "publish or perish" incentives created by the research funding system and the criteria for professional promotion [38]. These conditions risk creating "vast amounts of commodified but disposable knowledge," a sort of "fast food research" void of quality and nutrition [39].

Aside from lack of relevance or excellence, criticisms have also been levied at the lack of rigor in academic research. By this, we mean a mix of carefulness and thoroughness. The simple Oxford definition of rigor is "the quality of being extremely thorough and careful." This definition does not favor a particular research design, objective, discipline or method. Rather, this definition represents the practice of taking great care in establishing and articulating research objectives, selecting and implementing appropriate research methods and interpreting research results - while at the same time acknowledging omissions and limitations. Donnelly et al. thus define rigor in research as "identifying all relevant evidence" within the available resources or timeframe [40].

A critique of lacking rigor seems particularly justified in energy social science, given that an examination of 15 years of peer-reviewed publications in this field found that almost one-third (29%) of the 4,444 studies examined had no description of an explicit research design—or method—whatsoever [41]. In the related field of global environmental governance and politics, a review of 298 articles published over 12 years noted that only 35% included a discussion of, or a justification for, the research methods employed [42]. Even articles with explicit research designs can still suffer from flaws. Hamilton et al. note that in the domain of energy efficiency and buildings: "analysis is often limited to small datasets and results are not applicable more broadly due to an absence of context or baselines" [43].

Finally, drawing from our own experience as editors, peer-reviewers and readers of energy social science, we observe that many articles are stymied by bad "style" – that is, poor structure, unclear analysis and difficulties in expression. Even when they make a novel contribution and employ a rigorous research design, many authors struggle to communicate clearly due to a lack of care in writing or a lack of fluency in language. Their papers often lack persuasive or cohesive elements

such as signposts, roadmaps, figures and tables; have many grammatical mistakes and typos; and exhibit a poor standard of written English. Put another way: many submitted articles are poorly written, and if they are published they seem destined to have a low impact—even if the research itself is novel and/or rigorous.

To remedy these tripartite limitations of novelty, rigor, and style, this Review offers a guide for researchers so they can improve the quality of their research. We have four objectives:

1 Bring attention to the importance of clearly articulating research questions, objectives, and designs.
2 Provide a framework for conceptualizing novelty.
3 Suggest codes of practice to improve the quality and rigor of research.
4 Provide guidelines for improving the style and communication of results.

Our hope is that this Review will contribute to more coherent, creative, rigorous and effectively communicated research that will enhance the contribution that energy social scientists make to both theory and practice. Our primary audience is researchers in energy social science, as well as readers who want to evaluate such research. Using our collective experience, we focus our suggestions on how social science research has been applied (and misapplied) to energy-related research questions—though much of this content is relevant to other social science applications, especially to societal issues such as transport and mobility, or environmental and resource management. Further, while this Review is intended to be useful for early career researchers, we believe that researchers of all levels can benefit from an ongoing dialogue about what makes high quality, novel, rigorous and effective research in our field.

2. Getting started: research questions, frameworks, objectives and designs

Although the later parts of this Review will explore how to improve aspects of novelty, rigor, and style, a useful starting point is to consider four core elements: 1) asking concise, interesting, socially relevant, and answerable research questions; 2) applying and testing theoretical constructs or conceptual frameworks; 3) clearly stating research objectives and intended contributions; and 4) developing an appropriate research design. Although it is not always a linear process, our flow has a researcher starting with their research question (demarcating their topic), moving to discuss how they will approach it or filter data (theoretical or conceptual lens), identifying specific aims (research objectives), and explicating a research design (selecting and operationalizing a particular research method or methods).

Although there is a large element of subjectivity in the sections to come, our contention is that all good papers should include clear research questions, a clear conceptual or theoretical basis, precise objectives and an explicit research design. We start with these steps because, in our experience, their absence is often a fatal flaw.

2.1. Asking socially relevant (and impactful) research questions

With some overstatement, getting the research question(s) right could be half the work of writing a good paper. The research question guides a literature review or collection of data, suggests the type of answers a study can give and provides a strong disciplining device when writing. Bellemare [44] proposes that good papers contain interesting ideas when they do one of three things: ask a question that has not been asked before; ask a "Big Question" that affects the welfare of many people; or ask a question that has been asked before but can be answered in a better way. For more detailed suggestions for how to craft research questions, we suggest Hancke's Intelligent Research Design [45].
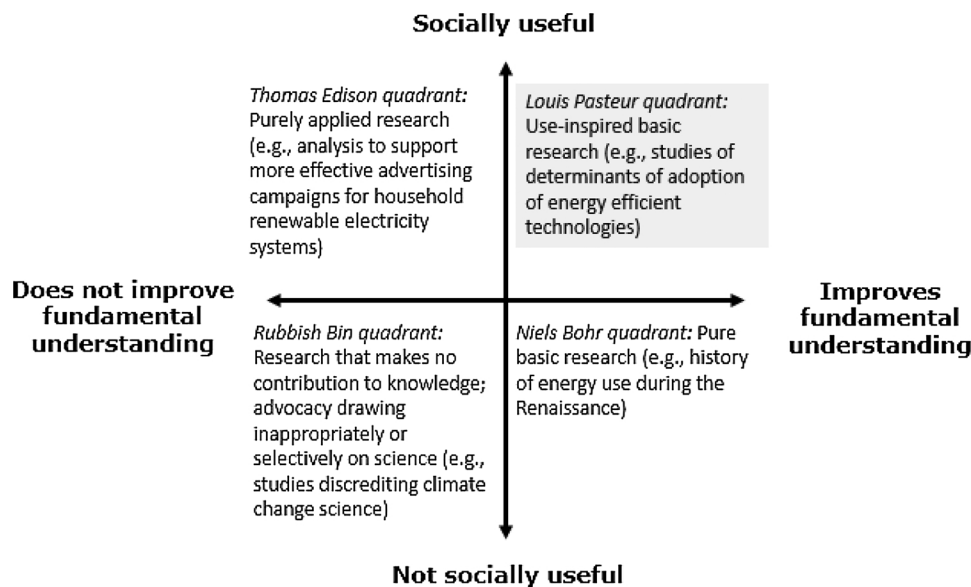
Socially useful

Thomas Edison quadrant:
Purely applied research
(e.g., analysis to support
more effective advertising
campaigns for household
renewable electricity
systems)

Louis Pasteur quadrant:
Use-inspired basic
research (e.g., studies of
determinants of adoption
of energy efficient
technologies)

Does not improve fundamental understanding ←————————————→ **Improves fundamental understanding**

Rubbish Bin quadrant:
Research that makes no
contribution to knowledge;
advocacy drawing
inappropriately or
selectively on science (e.g.,
studies discrediting climate
change science)

Niels Bohr quadrant: Pure
basic research (e.g., history
of energy use during the
Renaissance)

Not socially useful

Fig. 1. A typology of energy social science research contributions.
Note: Research in the top right-hand quadrant (shaded) is the most impactful, but also rare.
Source: Authors, modified from Stern et al. [49].

Here, we summarize three tips. First, build your question(s) from empirical or conceptual material—do "pre-search." No research question can be constructed without reading. All good research questions are the product of prior engagement with empirical and/or theoretical material. Second, ensure that your research question(s) are researchable. Is there reliable and accessible evidence that you can use to answer your question, or is there scope for producing such evidence? Will this evidence be available to others? Is your question limited in time or space? Does it have clear enough boundaries and a logical "end" that you work towards and explain or answer? Or are you chasing a moving target? Third, ensure that your research question is answerable. A research question needs to be asked in such a way that your expectations can be wrong (and that you know when they are wrong) and that you can be surprised. When confronted with reliable evidence, the answer to the question should be apparent.

Even better is a question that both advances theory and addresses a relevant social problem, meaning that your question matters to academia, practitioners and other stakeholders. The typology in Fig. 1 depicts four broad categories of research contribution. Stern et al. warn that too little research in energy social science falls into "Pasteur's Quadrant" of both advancing scientific or theoretical understanding and being immediately useful at addressing a pressing energy- or climate-related problem [46]. As Mourik put it recently, "We need scientists that are allowed to work in this in-between space, a boundary space between research and practice" [47]. Similarly, O'Neil and her colleagues write that more problem-driven research is needed that confronts social or environmental issues, rather than merely describing them [48]. Thus, asking socially relevant questions can facilitate broader social impact, something elaborated more in Box 1.

Crafting research questions in this way can make a study socially- and policy-relevant by design, helping to ensure relevant insights for policymakers, practitioners, managers and/or other stakeholder groups. Under this logic, research is not only an art or craft, but a civic duty. We argue that more applied research is needed in the field of energy social science, that researchers should think about policy/practitioner applications when developing their research objectives, and that, where appropriate, researchers should seek to integrate practitioners directly into the research process.

2.2. Engaging with theory and conceptual frameworks

Separate from an abundance of possible research questions, there is no shortage of conceptual frameworks, analytical frameworks and theories available to the scholar (terms that we use interchangeably, while acknowledging the literature on the differences between). The selection of theory can also flow from a "paradigm," a worldview or way of interpreting reality [50].

There are many excellent reviews of these theories available. For example, reviews relevant to energy social science include:

- Edomah et al.'s comparison of the theoretical perspectives related to energy infrastructure [51];
- Jackson's review of theories for consumer behavior and behavioral change [52];
- Kern and Rogge's survey of theories of the policy process and their relevance to sustainability transitions [53];
- Peattie's catalogue of theories relating to values, norms, and habits associated with "green consumption" [54];
- Scheller and Urry's survey of sociotechnical transitions, social practice theory and complexity theory for transport and mobility researchers [55];
- Sovacool and Hess's survey of 96 theories, concepts, and analytical frameworks for sociotechnical change (including energy transitions) [56];
- Wilson and Dowlatabadi's analysis of decision-making frameworks relevant to energy consumption [57].

As these theoretical reviews emphasize, different theories may be more or less suitable for different types of research question and may also have varying and sometimes incompatible foundational assumptions. Rather than dive into the many specific theories relevant to energy social science, we instead indicate some of the most important dimensions and features of those theories, and how these shape research questions, objectives and designs.

One way of classifying theories is to identify their underlying paradigm, that is, their assumptions about the nature of reality (ontology), the status of knowledge claims about that reality (epistemology) and the appropriate choice of research methods. For example, Table 1 highlights the assumptions associated with three broad

Box 1
Research impact.

Admittedly a prosaic concept, research impact can be roughly divided into academic and non-academic dimensions. Academic research impact is often (over)simplified to mere citations—an impactful article is well cited and utilized by others, with citation counts being offered by Google Scholar, Scopus, or ISI Web of Science. Other forms of academic impact include downloads (via the journal, an institutional or personal website, or online platform such as the Social Science Research Network or ResearchGate), requests for consultancies, and invitations to present work at academic conferences or media events. The Research Excellence Framework, or REF, in the United Kingdom categorizes academic impact according to five classifications: four star (quality that is world-leading in terms of its originality, significant and rigor), three star (quality that is internationally excellent but which falls short of the highest standards of excellence), two star (quality that is recognized internationally), one star (quality that is recognized nationally), and unclassified (quality that falls below the standard of nationally recognized work).

The REF defines non-academic impact as "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia." Typically, non-academic impacts arise when research directly informs national debates in parliament (or relevant political bodies), affects policy and regulation, influences how the private sector operates, connects with legal testimony and ongoing court decisions, and/or shapes national discourse via the broader mass media. Non-academic impacts are typically more difficult to both achieve and demonstrate.

Source: Authors, modified from REF documents.

---

paradigms or philosophies of science - positivism, interpretivism and critical realism. Theories in the positivist paradigm assume that reality is objective, focus upon generating and testing hypotheses and are well suited to quantitative research methods such as multivariate regression. In contrast, theories in the interpretive paradigm assume that reality is (at least partly) subjective, focus upon uncovering the meaning actors give to events and are well suited to qualitative research methods such as participant observation. Critical realism is a more recent philosophy of science that partly reconciles these different perspectives and is consistent with both quantitative and qualitative research methods. Some theories align closely and explicitly with one of the paradigms in Table 1, while others are more ambiguous, or combine elements from more than one perspective.

A second way of classifying theories is to identify their primary focus, namely: agency, structure or discourse - or a hybrid of these. As Table 2 indicates, agency-based theories prioritize the autonomy of the individual, and thus tend to emphasize individual behaviors and beliefs. For example, Karl Popper [59] famously recommended that "…all social phenomena, and especially the function of all social institutions, should always be understood as resulting from the decisions, actions, attitudes etc. of human individuals, and….we should never be satisfied by an explanation in terms of so-called 'collectives'…." In contrast, structural theories emphasize the opposite: macro-social relationships and technological infrastructures that constrain the autonomy of people and organizations. In contrast to both, discursive theories shift the focus away from individual choice and social structure and towards more cultural factors such as language and meaning. Again, different theories give differing degrees of emphasis to these factors, and many occupy a hybrid space, emphasizing the complex interactions among agency,

structure, and discourse (sometimes called "relational" or "processual" theories) [60,61]. Multilevel frameworks often sit within this hybrid category.

Whereas the first four types of theories in Table 2 are inherently descriptive or explanatory, a fifth type of theory is normative and attempts to assess whether a technology, practice, policy or other unit of analysis is a net positive or negative for society or individuals. To do so, normative theories often rely on criteria set by ethics, moral studies, social justice or political ecology. Put another way, the first four theories are about explanation, whereas normative theories are about evaluation.

A third way of classifying theories is to identify their particular assumptions about human behavior and decision-making (Table 2). These approaches range from those subscribing to a rational actor model that sees people as basing decisions on reasons, utility or logic to more complex theories incorporating broader dimensions such as attitudes, beliefs, morals, habits, and lifestyles. These dimensions are not mutually exclusive, but different theories vary in the relative emphasis given to each and hence may be more or less useful in explaining particular behaviours and decisions.

A final caveat to engaging with theory—especially within the positivist paradigm—is managing the tension between specificity and generalizability, as well as between parsimony and complexity [65]. Jackson notes that more complex theories can aid conceptual understanding but can be difficult to use in practice—for instance they are poorly structured for empirical quantification or surveys [66]. Less complex theories can be easier to test but may hinder comprehension by omitting key variables and relationships. Sartori found this to be the case in politics and international relations: as one moved up a ladder of

Table 1
Differing paradigms in energy social science research.
Source: Based on Mingers et al. [58].

	Positivism	Interpretivism	Critical realism
Ontology	Independent and objective reality	Socially constructed reality	Objective, stratified reality consisting of surface-level events
	Causality indicated by constant conjunctions of empirical events	Multiple realities possible	Real entities with particular structures and causal properties
Epistemology	Knowledge generated by discovering general laws and relationships that have predictive power	Knowledge generated by interpreting subjective meanings and actions of subjects according to their own frame of reference	Knowledge generated by process of "retroduction," used to create theories about the entities, structures and causal mechanisms that combine to generate observable events
	Emphasis on prediction	Emphasis on interpretation	Emphasis on explanation
Methodology	Specific, deductive hypothesis-testing	Exploratory, inductive, theory-generating	No preference for a particular method - choice depends upon the research question and the nature of the relevant entities and causal mechanisms Mixed methods encouraged
	Emphasis on generalizability	Emphasis on depth rather than generalizability	
	Quantitative methods, such as experiments, surveys and statistical analysis of secondary data	Qualitative methods, such as ethnographies and case studies	

Table 2
Differing focus and behavioral assumptions of theories in energy social science.
Source: Authors, modified from [62–64].

Focus	Assumptions relevant to energy social science	Examples of research themes
Theoretical emphasis		
Agency	Individuals have autonomy, behavior is at least partially driven by individual characteristics	Exploring the role of individual beliefs, attitudes and preferences in energy use
Structure	The macro-social, technological, or political environment determines (or prevents) energy systems stability and change	Exploring the positions of firms in an industry, the self-perpetuation of energy-using "practices", or relations between energy systems and the natural environment
Discourse	Energy decisions are mediated through language, symbolism, narratives, rhetorical visions, and discursive coalitions	Exploring the visions around hydrogen fuel cells or small modular reactors, revealing the narratives of fossil fuel incumbents
Hybrid	Energy decisions and policies occur across the categories of agency, structure, and discourse/meaning	Exploring the accelerated diffusion of new automobiles (involving drivers, policies, and marketing messages)
Normative	Energy systems can exert a distinct net positive or negative impact on society	Exploring the distribution of energy related externalities, or the fairness and due process of energy permitting decisions
Behavioural assumptions		
Rational actor (and variations)	Actors are deliberative, knowledgeable, and possess well-defined, stable preferences	Elicitation of preferences and valuation regarding goods and behavior
Attitudes-beliefs	Behavior is shaped by attitudes, beliefs and values	Exploring associations between attitudes or beliefs and behavior
Morals and norms	Behavior is driven by norms about what is right, or what others want	Exploring associations between norms and behavior
Habits	Many behaviors are routinized and engaged in without conscious deliberation	Exploring the role of context and structure in behavior
Lifestyle and identity	Behavior and beliefs are shaped by (and shape) worldview, identity, and social interactions	Exploring identity and behavior; understanding processes of social influences

abstraction, scope, purpose and concepts change to become more general but less robust [67]. Azjen adds that theories have scopes—some (such as his Theory of Planned Behavior) have to be adapted to each study or application, whereas others (such as Value Belief Norm theory) can use concepts and measures that apply across a large range of dependent variables [68].

The point here is that good studies not only employ a relevant theory or conceptual framework; they acknowledge its analytical emphasis, its underlying ontological and epistemological assumptions, its degree of complexity or abstraction and the strengths and limitations that result.

2.3. Stating research objectives

In addition to selecting research questions and theoretical framework(s), the rigorous researcher must also clearly articulate the research objectives. As concisely summarized by Babbie [69], researchers should aim to (1) specify as clearly as possible what they want to find out, and (2) determine the best way to do it. This entails providing a concise statement of exactly what the researcher aims to do in a particular study—what should prove to be the guiding statement for the eventual considerations and details of the specific research design. In our experience, one to four objectives are appropriate for a standard journal article, and we encourage researchers to clearly state these objectives at the end of their introduction section, and to continually reflect back on them throughout the article. We distinguish objectives from more general research questions (most often stated in qualitative or intepretivist research approaches), and more specific hypotheses (most often stated in quantitative, positivist research approaches).

Consider these oversimplified examples that draw from an application of value theory:

- *Research Question*: What consumer traits or motivations are associated with interest in electric vehicles?
- *Research Objective*: Determine which values are associated with interest in electric vehicles by estimating discrete choice models using choice data collected from a sample of UK car buyers.
- *Research Hypothesis*: Interest in electric vehicles is positively associated with higher levels of biospheric and altruistic values.

Well-articulated research objectives will communicate the type of analysis that is needed and the intended novelty of the contribution. As described by Babbie [70], the objective may be to: (1) "explore" new research categories or relationships; (2) "describe" or observe the state of something (e.g. reporting frequencies of citizen support for a given climate policy); or (3) "explain", typically meaning looking for causality through statistical analysis, experimental design or perhaps narrative analysis.

Similarly, the research objectives must also communicate the intended scholarly contribution of the research, which might be theoretical (developing or contributing to new theory or testing existing theory), methodological (developing or contributing to new methods) or empirical (new applications of existing methods or theories, or new types of evidence)—issues we explore in Section 3. A given study can be publishable if there is clear novelty in at least one of these categories, and sometimes in two. Only rare and exceptional papers make contributions across all three—and attempts to do so can lead to confusion or incoherence.

Further, in an interdisciplinary field, rigorous researchers know that their objectives must somehow communicate the paradigm that is guiding their inquiry, that is, the underlying assumptions (often discipline-specific) about the nature of reality, how the researcher interacts with reality and the appropriate methods to use [71]. While numerous paradigms exist, we focus here on the very broad dichotomy between the positivist paradigm, which emphasizes quantitative research methods, and the intepretivist paradigm, which emphasizes qualitative research methods. As noted above, quantitative methods are not just about numbers, but rather stem from a paradigm that emphasizes hypothesis testing, large and representative sample sizes, statistical analyses, prediction, generalization and the objectivity of the researcher—notions dominant in disciplines such as social psychology, economics, and American political science. In contrast, qualitative approaches could be characterized as theory or hypothesis generating, rather than hypothesis testing, and focus more upon understanding, meanings, interpretation, social construction and the subjectivity of the researcher [72]. These notions are dominant in disciplines such as anthropology, sociology, and European political science. These two broad paradigms (and associated disciplines) are associated with different rules, standards and guidelines so it is important for researchers to communicate the nature of their research objectives—at a minimum

Table 3
Dominant research methods within energy social science.
Source: Authors, modified from Tranfield et al. [73] and Sorrell [74].

	Experiments and quasi-experiments	Literature reviews	Surveys and quantitative data collection	Data analysis and statistics	Quantitative energy modeling	Qualitative research	Case studies
Core Disciplines	Behavioral science, social psychology, behavioral economics, medical and life sciences (when applied to energy social science topics)	All disciplines, though meta-analysis is more common in quantitative disciplines (e.g. psychology and economics)	Various, but especially, sociology and marketing	Various, but especially, economics, psychology and some traditions within political science	Economics, engineering, environmental science (for Integrated Assessment Models)	Anthropology, sociology, history, geography, policy studies, science and technology studies	Various, but similar to qualitative research
Description	Exemplified by randomized controlled trials, but also includes controlled before-and-after studies and various types of matched comparisons. Potentially provides reliable evidence of the causal effect of different mechanisms by explicitly controlling for the effect of different variables.	Reviews generally do not present new or original data. Instead, they scour existing peer-reviewed or grey literature, with the aim of identifying the current state of knowledge. Reviews occasionally use content or discourse analysis.	Survey data can provide valuable information about a given sample and population (e.g. consumers, citizens, or stakeholders), allowing the use of descriptive statistics and test of association among variables among variables	Technique for exploring quantitative hypotheses, such as comparing means across samples or testing associations of variables; can be performed on either new data collected by the researcher or analysis of existing (secondary) data.	Covers a variety of approaches to analyzing the operation and consequences of different mechanisms using simplified mathematical models. These abstract from real-world complexities and focus on key mechanisms, either conceptually or by combining theoretical assumptions with empirical data	A variety of techniques for obtaining information regarding the opinions, understandings, attitudes and perceptions of individuals and groups in different contexts. Examples include semi-structured interviews, participant observation and focus groups	In-depth, examination of one or more subjects of study (cases) and associated contextual conditions. Can rely upon multiple sources of both quantitative and qualitative evidence.
Research culture	Convergent, subject to rigorous scientific evaluation	Convergent for meta-analysis and systematic reviews, but largely divergent for other forms	Somewhat convergent, practices vary by discipline and nature of research question (e.g. descriptive or causal)	Somewhat convergent, general principles hold across disciplines, but some disciplines have developed more specific practices (e.g. econometrics)	Divergent, research questions and model assumptions differ greatly across disciplines and approaches	Divergent, split among different subcategories of qualitative/ interpretivist research, e.g. post-positivism, relativism, and constructivism	Divergent, split between different objectives, types of case (e.g. typical, extreme, influential) and types of evidence
Codes of practice for methodological rigor	Can be based upon a hierarchy of evidence, studies assessed against predetermined criteria, standardized reporting structures	Some standardized assessment criteria exist, particularly for systematic reviews and meta-analysis	Can be based upon increasingly accepted assessment structures	Based upon statistical principles, but preferred techniques and practices vary between disciplines (e.g. economics versus psychology)	Some codes have been proposed, but these vary with the model type	Data collection not always guided by explicit criteria	Depends on case study types, whether single or comparative cases are needed, and spatial or temporal variation

17

whether they intend to generate theory or hypotheses (or explore new categories), or to test theories and hypotheses (to quantify the size, nature or relationships of existing or new categories).

In short, the nature of the objectives will determine what types of methods, analysis and interpretations are appropriate—which leads into research design which we discuss next.

2.4. Explicating a research design

Our final suggestion is that every article ought to have a clearly articulated research design—this ensures the conceptual frameworks are operationalized, research questions are answered, objectives are met and/or hypotheses are tested. In very simple terms: a research method refers to a technique for gathering or analyzing data (e.g. the categories we outline next), while a research design is how exactly such a method, or methods, become executed in a particular study. The goal of a research design should be to provide enough detail to make the study transparent, helping readers to assess the study in light of the stated research objectives, while facilitating replicability.

In energy social science, most research designs use one of the seven categories summarized in Table 3 – or some combination thereof. Note that any taxonomy of research methods will give inadequate attention to some methods while missing others altogether—our taxonomy is merely an attempt to summarize the dominant categories within energy social science. Table 3 identifies the disciplines most associated with each of the seven research methods, describes their key elements, summarizes their research cultures and sketches out codes of practice for rigor. The table omits research designs using multiple or mixed methods, which we discuss further in Section 3.2. However, even where multiple methods are used in a single study, each individual method ought to follow the codes of practice summarized Section 4.

As Table 3 implies (and as noted in the previous section) the two general "classes" of research method—quantitative (surveys, statistics, modeling) and qualitative (interviews, focus groups, observation, case studies)—have different strengths and weaknesses. Quantitative methods are best for testing hypotheses or quantifying relationships (e.g. correlation), while qualitative methods are best for exploratory studies or accessing more in-depth information, such as how social actors construct meaning. Different methods may in turn be associated with different degrees of consensus (convergence) or debate (divergence) about what constitutes rigor. These tradeoffs and tensions will become more apparent as we examine codes of practice in Section 4, but here we offer a brief summary of each method category.

2.4.1. Experiments and quasi-experiments

Experiments involve human participants and seek to test for causal relationships between variables, while isolating the study or relationship from (or controlling for) other potentially influential variables [75,76]. "True experimental designs" are distinguished by: a) random selection and/or assignment of participants; and b) researchers having control over extraneous variables [77]. In contrast, quasi-experimental designs seek to identify the causal effect of some treatment or effect, but lack random assignment to treatment groups [78,79]. In some cases ("natural experiments") the experimental conditions are outside the control of the investigators, but nevertheless provide sufficient degree of control to permit causal inference. Experimental and quasi-experimental designs can be implemented in "lab" or "field" settings, as well as via trials, games, and simulations.

2.4.2. Literature reviews

A literature review is a compilation and integration of existing research, typically with the aim of identifying the current state of knowledge and specific research gaps. The relevant evidence may include both peer-reviewed and grey-literature. Reviews typically involve repeated searches of databases using specific keywords in order to identify large bodies of evidence. Depending upon the research

question(s), the search may impose relatively narrow criteria for inclusion, or much wider criteria that allows consideration of different research designs and types of evidence [80,81]. As discussed later, we distinguish between three broad types of literature review: meta-analysis, systematic review, and narrative review.

2.4.3. Surveys and data collection

Survey methods involve data collection using a survey instrument or structured questionnaire with a sample of respondents from a relevant target population. Surveys are used extensively within many social science disciplines, but both the practices and norms associated with implementing surveys and the interpretation of results can differ between those disciplines.

2.4.4. Data analysis and statistics

Quantitative data analysis typically utilizes statistical techniques, though norms of implementation can again vary between social science disciplines, as can the relative use of specific techniques (e.g. MANOVA versus multivariate regression). This divergence results in part from variations in the type of data that is commonly used. For example, social psychology relies heavily upon primary data collected via experiments or surveys, which provide good controls for confounding variables. In contrast, economics makes greater use of secondary data sources such as government statistics, which can be incomplete or nonexistent for some variables, and can be prone to measurement and other errors.

2.4.5. Quantitative energy models

Energy modeling includes techniques that quantitatively represent and analyze the technical, economic and (to a lesser degree) social aspects of energy systems, typically in a forward-looking manner [82]. These models may focus upon energy demand (e.g. vehicle stock model), energy supply (e.g. linear programming model of electricity generation) or whole energy systems; their scope may range from the very narrow (e.g. electricity distribution within a single city) to the very wide (e.g. the global energy system); they may utilize a range of behavioral assumptions (e.g. full or bounded rationality) and mathematical techniques (e.g. systems dynamics, agent based); and they may be integrated to a greater or lesser degree with broader economic models. Energy models are widely used to explore socially-relevant questions, such as how changes in income, technology or policy may shape energy consumption and carbon emissions over time, and what future energy systems may look like [83–85]. For the most part, all modeling exercises boil down to translating a series of assumptions into mathematical form (equations, algorithms, parameters) and then testing the logical consequences of those assumptions.

2.4.6. Qualitative research

Qualitative research designs cover a range of techniques for collecting and analyzing data about the opinions, attitudes, perceptions and understandings of people and groups in different contexts. Qualitative research methods differ according to the nature of data collection, as well as the means of analyzing that data. In energy social science, the most popular approaches to qualitative data collection tend to be semi-structured interviews, focus groups, direct observation, participant observation and document analysis [86–88]. What each of these methods has in common is that they are inductive and exploratory by nature, seeking to access a particular perspective in depth, rather than to test a specific hypothesis.

2.4.7. Case studies and cross-case comparisons

A final common research design is a case study, which is an in-depth examination of one or more subjects of study (cases) and associated contextual conditions. Case studies can use both quantitative and qualitative research techniques. George and Bennet define a case study as a "detailed examination of an aspect of a historical episode to

develop or test historical explanations that may be generalizable to other events" [89], while Yin defines it as "an investigation of a contemporary phenomenon within its real-life context when the boundaries between phenomenon and context are not clearly evident" [90]. Rather than using statistical analysis of data from a large sample, case study methods often involve detailed, longitudinal assessments of single or multiple cases - which may be individuals, groups, organizations, policies or even countries [91,92].

3. Promoting theoretical, methodological, and empirical novelty

This section of the review focuses on novelty: how to produce research that is original, fresh, or even exciting and unexpected. Studies can typically be classified by their primary form of novelty or contribution to the literature. Although this will vary, studies generally fall into one of three types:

- Theoretically-novel articles contribute to creating, testing, critiquing, or revising some type of academic concept, framework or theory;
- Methodologically-novel articles focus on the research process itself and include testing, revising or developing new research methods;
- Empirically-novel articles reveal new insights through new applications of existing methods and theories (e.g. to different regions, contexts or research questions), as well as through analysis of new types of evidence or data.

For the most part, articles that fit into the third category are more numerous—there tend to be far more applications of existing theories and methods than developments of new ones. Further, there is clearly overlap in these categories; e.g., a theoretically-novel article will frequently include some empirical novelty as well. The following sections describe each of categories in turn. In our experience, an article that does one of these three things well is sufficient. Seeking objectives that cross two can be better, but doing all three is overambitious and likely to lead to confusion rather than clarity.

3.1. Theoretical novelty: create, synthesize, or test theories

Theoretically novel studies can create, apply, advance, test, compare or critique concepts or theories. Here we briefly demarcate three types of theoretical novelty: inventing theories, synthesizing theories, and triangulating theories.

3.1.1. Theoretical invention

Perhaps the most rare (and difficult) is theoretical invention or innovation. Scholars can sometimes develop new frameworks (invention) or further elaborate and advance existing theories (innovation). Prominent examples relevant to energy social science would be the initial papers that presented "technological innovation systems" (with its emphasis on the functions of innovation systems) [93–95] and "social practice theory" (with its emphasis on materials, competencies, meanings, and connections) [96,97]. In both cases, the motivation for doing so was the perceived limitations of existing theories for explaining the phenomena in question.

3.1.2. Theoretical synthesis

Theoretical synthesis attempts to integrate existing theories or concepts into a new conceptual framework. For example, the Unified Theory of Acceptance and Use of Technology (UTAUT) model integrates concepts from psychology, technology studies, economics, and innovation studies [98]. Similarly, the "Multi-Level Perspective" (MLP) on sociotechnical transitions (with its emphasis on niches, regimes, and landscapes) integrates ideas from evolutionary economics (e.g., variation and selection, path dependence, lock-in), science and technology studies (e.g. actor-networks, social constructivism) and various

traditions within sociology (e.g. structuration, social practices, social expectations) [99–101]. At a more conceptually focused level, Axsen and Kurani integrate aspects of Rogers' Diffusion of Innovations with theories of social networks, comformity, and translation (such as Actor Network Theory and Social Construction of Technology) to create a "reflexive layers of influence" herustic to assess low-carbon consumer behaviour and social networks [102,103]. One must take care when syntheizing, however, to ensure that the theories being integrated are complementary and have commensurate underlying assumptions [104,105], and that the resulting framework is not overly complex.

3.1.3. Theoretical triangulation

Theoretical triangulation refers to the comparison, evaluation and/or testing of multiple theories or concepts [106]. This involves comparing a number of theories to see which best explain a particular set of empirical observations. One classic example from political science explained a single event, the Cuban Missile Crisis, through three different theories: Realism or Rationalism; Organizational or Institutional Theory; and Bureaucratic Politics and Negotiation [107]. A more recent study in the domain of energy and social science sought to explain the consumer adoption of residential solar PV systems in the United States by testing the validity of concepts from Rogers' Diffusion of Innovation theory, Azjen's Theory of Planned Behavior, and Dietz and Stern's Value-Belief-Norm Theory [108]. Similarly, Ryghaug and Toftaker triangulate Social Practice Theory with Domestication Theory to explain electric vehicle adoption in Norway [109]; while Sovacool et al. compare the MLP, the Dialectical Issue Lifecycle Theory, and Design-Driven Innovation to explain the obstacles to electric vehicle diffusion in Denmark and Israel [110]. Such theoretical triangulation can reduce bias in theory selection and improve theoretical constructions through critical reflection [111]. It can also help researchers select the most appropriate analytical tools for their research question, properly credit those who contributed towards the development of theory, and avoid dogmatic adherence to particular ideas that can stifle both conceptual advancement and communication between disciplines [112].

3.2. Methodological novelty: develop novel or cutting edge methods

Another category of novelty applies to papers where the primary contribution is to develop a research method that is new, or modified from a conventional version or combined with other methods in a new way. Given the size and diversity of the energy social science research community (spanning many disciplines and research designs), together with the dynamic nature of research methodology (new approaches and techniques are continually emerging), it is impossible to present an exhaustive or even representative list of state-of-the-art methods. In some cases, this type of novelty can involve taking methods from one discipline or area and attempting to make it "better," such as mixing it with other methods. In other cases, novelty can involve utilizing methods that are "new" and only beginning to emerge among academics more generally. To illustrate, we summarize three examples of novel methods in our field: multiple methods, longitudinal research and behavioral realism.

3.2.1. Multiple or mixed methods

A first example of methodological novelty is the use of "multiple methods" or "mixed-methods". The first term is more general and refers simply to any research design that uses or blends several different methods (e.g. semi-structured interviews and media analysis). The second term is more specific and refers to the integration of quantitative and qualitative research methods in a single study [113,114]. There is much debate about how to best implement mixed-methods [115], though in practice the most popular approach has been to combine quantitative surveys with qualitative interviews [116]. Creswell provides a typology of mixed-methods approaches, which vary in the sequence and intention of integration, with the most suitable approach

being one that is best matched to the research objective [117]. The term "methodological triangulation" is used to describe the use of multiple methods to view a given social phenomenon through multiple perspectives [118], though the term triangulation has become controversial in some disciplines due to the potential implication that there is a single reality to "see" rather than multiple valid, and potentially very different, perspectives [119]. Effective implementation of multiple methods can lead to more sophisticated answers to research questions [120] and can help overcome the limitations of individual research approaches [121].

3.2.2. Behavioral realism

A second example of methodological novelty is the addition of behavioral realism to quantitative energy models. A broad range of such models have been criticized for lacking realistic assumptions about behavior, including optimization models that assume that actors are hyper-rational and fully informed [122]; and agent-based models that lack an empirical foundation for their assumptions [123]. Behavioral realism broadly refers to improvements in the representation of agents or decision-makers in these models, especially consumers, to better match real-world behavior in the target population—which of course can vary by region and culture and over time. This realism can come from the use of empirical data, representation of both financial and non-financial motivations, and representation of diversity or heterogeneity in behaviors and motives. Improving the behavioral realism of energy models typically involves the combining of methods in some form, for example via translation of insights from an empirical method to the model in question [124]. As examples, some recent studies have sought to improve optimization models by using meta-analysis of behavioral studies to estimate parameters representing processes of social influence [125]; by representing heterogeneity in consumer valuation of product attributes [126]; and by incorporating "decision-making heuristics" such as present bias, habit formation and loss-aversion [127]. For agent-based models, innovative research is exploring how to use results from surveys, laboratory experiments, case studies and other sources to inform the selection of model parameters [128–130].

3.2.3. Repeated data collection and longitudinal research

A third example of methodological novelty is approaches to repeated data collection and longitudinal research design. While most surveys and interviews are cross-sectional (accessing respondents at a single point in time), longitudinal approaches offer the opportunity to improve the depth and reliability of collected data, as they aim to study changes in a sample of participants over time. Here, one can distinguish between "panel" studies that repeatedly survey the same participants, and "pooled cross-sectional" studies that repeatedly sample the same population but analyze different cross-sections over time (sometimes called a "time series cross sectional" study) [131]. Such approaches can allow more accurate inference of relevant parameters; provide greater control of confounding variables; facilitate the testing of more complicated behavioural hypotheses; and permit more reliable investigation of dynamic relationships [132,133]. For example, studies have shown that interview participants and survey respondents are able to express more stable preferences for electric vehicles if they have been given a multi-day trial of that vehicle [134–136]. Panel-type survey research is benefiting from improvements in information and communication technologies that make it easier to follow a given respondent over time [137]. The panel approach in particular comes with many challenges, including how to minimize and address attrition over time, and how to mitigate the behavioural effects of repeated surveying, such as conditioning—which can be costly and time consuming to overcome [138]. Another novel future direction is the meshing of qualitative narrative analysis with quantitative longitudinal data [139].

3.3. Empirical novelty: new applications, new data, and new types of evidence

The final type of research novelty is empirical—where we distinguish between new applications, new data, and new types of evidence.

3.3.1. New applications

This category represents the majority of studies in our field: those that apply existing theories and methods to new applications, such as new regions, case studies, contexts, or research questions. While such studies can provide incremental contributions to the testing of theories or the development of methods, their primary contribution is empirical, in improving understanding of the relevant topic or application. Such studies frequently score high on practicality, or the "immediate usefulness" dimension of Fig. 1 (above), but trend towards the "Thomas Edison" rather than "Louis Pasteur" quadrant.

Examples are highly diverse, including: using surveys to apply identity theory to different types of pro-environmental behaviors [140]; applying an existing technology adoption models to simulate compliance with US fuel economy standards [141]; using transaction cost economics to understand the conditions for success of energy service contracts [142]; and applying the MLP to the case of Norwegian electric vehicle policy [143]. Some empirically-novel studies have no strong theoretical framework, being primarily descriptive, exploratory, or grounded in data. For example, such studies may ask: how many English citizens would support a carbon tax? Or how have financial incentives influenced the uptake of household solar panels and electric vehicles? Many empirically-novel studies also tend to be socially-relevant by design, seeking to generate immediate insights for policymakers, practitioners, managers and other stakeholders.

3.3.2. New data from exceptional groups or populations

Empirical novelty also includes collecting and/or analysing new types of data; typically such data are either difficult to collect or access (e.g. lack of sampling frame, high costs, or needs for computing power), challenging to analyse, or neglected for some other reason (e.g. the paradigm or common practice in a discipline tends to ignore such viewpoints). To illustrate, we identify four types of "exceptional" stakeholder groups that often prove difficult to access: elites, experts, small populations (early adopters, venture capitalists), and vulnerable populations (minorities, indigenous people, the chronically poor). In some cases, collecting data from such populations can be a novelty itself.

Perhaps the most common example of this approach is data collection from elites: people in a position of power, influence or expertise regarding energy decision-making (as opposed to laypersons, consumers or voters) [144]. Examples of elites include business executives, heads of state, senior ministers, or senior directors and managers of energy programs [145]. Elite interviews are especially useful for revealing the motivations and actions behind policy formation and adoption, although access to the highest levels of politics or policy-making is often restricted and confidentiality concerns abound [146].

A second category is experts in a particular topic area, which may include inventors, entrepreneurs, researchers or intellectuals. Sampling or accessing such experts can be challenging, in particular because it may not be clear who makes up the target population (where to draw the boundaries), how to draw a sample, and how to best engage the sample. The perspective of experts can be accessed using "Delphi" techniques that can facilitate convergence towards a consensus view on a topic (e.g. future energy prices or the capabilities of energy storage technology) [147].

Small populations include, for example, pioneer adopters of low-carbon technologies [148,149] or venture capitalists [150]. These can be difficult to access due to small or non-existent sampling frames, yet their viewpoints can provide an important, often missing contribution to a given literature.

Finally, sensitive or vulnerable populations can include the survivors of energy accidents such as those at Chernobyl [151] or Fukushima [152], indigenous peoples [153–155], children [156], the elderly or ill [157,158], and the chronically poor [159–161]. Understandably, strategies for accessing these groups will be completely different from those for elites and experts, and will require cultural sensitivity and careful attention to ethics. Nevertheless, despite these added steps and challenges, it is often critically important for the perspectives of these groups to be considered in broader theory, research and decision-making.

3.3.3. New types of evidence

A third category of empirical novelty is new forms of evidence. Here we use the example of big data - interpreted as "extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions" [162]. These datasets may cover large populations, or achieve high temporal resolution (e.g. second-by-second observations), or both [163,164]. The data may be generated by people themselves (e.g. decades of information collected through time-use diaries [165]), but is more commonly measured automatically by digital technologies such as smart meters (in homes) [166], load-monitoring devices [167], and GPS devices (attached to people, phones, vehicles, or vessels) [168,169]. Although not (yet) widely used in energy social science, other sources of big data that could yield empirical insights are telematics in automobiles [170], online shopping profiles [171], and social media content such as Facebook and Twitter [172]. Some applications combine data sources: for example, Chatterton et al. aggregate data from 70 million domestic energy meters and vehicle odometers, with the aim of identifying areas in the United Kingdom with high household and vehicle energy consumption [173].

Hamilton et al use the term "energy epidemiology" to describe the use of such data to measure and explain energy demand patterns, and to predict future changes in energy demand from policy and other interventions [174]. As they write:

> Energy epidemiology is the study of energy demand to improve the understanding of variation and causes of difference among the energy-consuming population. It considers the complex interactions between the physical and engineered systems, socio-economic and environmental conditions, and individual interactions and practices of occupants. Energy epidemiology provides an over-arching approach for all the disciplines involved, where findings from large-scale studies both inform energy policy while providing a context for conventional small-scale studies and information input for predictive models [175].

Big data and energy epidemiology therefore open up new opportunities for exploring the relationships between consumer behavior and energy use. For example, automatically collected data can avoid the errors (or cognitive burden) of self-reported behavior, while data on consumer purchases can provide insights into consumer preferences while avoiding the limitations of hypothetical, stated choice experiments. But such applications raise complex and important questions about data privacy, transparency, security and accountability as well as third-party verification of data quality [176–179].

4. Promoting rigor via codes of practice, hierarchies of evidence, and appropriate balance

In this section, we focus on rigor: how to strive for careful and thorough research designs that ensure the research objectives are achieved. This definition relates to concepts of validity, which are defined in Box 2. We focus our discussion on three lessons:

- The usefulness of codes of practice for our seven research designs, where we advocate a "fit for purpose" approach.
- The limitation of hierarchies of evidence, where some disciplines

emphasize a ladder of approaches.
- The need for appropriateness and balance, where studies need not excel across all criteria.

4.1. Towards codes of practice

Here, we propose some basic "codes of practice" for different research designs—recognizing that the strength of a particular approach will depend on the context, objectives and research questions. Rather than offering a definitive checklist, this is more of a "toolbox," "horses for courses," or "fit for purpose" approach to rigor. More detailed guidelines for each of the research designs can be found in the cited sources. To be clear, these codes of practice are intended to emphasize which research designs or methods might be appropriate in particular settings, but the choice is dictated not only by the codes of practice, but also by the logic of inquiry and the research objectives.

4.1.1. Experiments and quasi-experiments

Experiments have a long history in disciplines such as social psychology, but have been adopted more slowly in other areas of social science [184]. In short, they aim to isolate and establish evidence for the causes of particular effects of interest. "True experiments" and "randomised controlled trials" (RCTs) in particular are defined by the randomized assignment of subjects to treatment conditions. Such designs are appropriate for research questions that seek to establish causal relationships between variables, such as: "do time-of-use electricity tariffs lead to reductions in electricity consumption?" [185]; or "does the format and color of energy efficiency labels affect the adoption of efficient appliances?" [186]. While such relationships are frequently inferred from non-experimental or "associational" studies, those inferences may be invalid [187,188]. For example, survey data may indicate a positive correlation between reported happiness and reported engagement in pro-environmental behavior, but the causality may be in the opposite direction (happier people may engage in more pro-environmental behavior) or the correlation may result from a third variable that is not observed (e.g. people with more free time may be happier and more inclined towards pro-environmental behavior). In order to provide stronger evidence of causation, the defining characteristic of true experiments is that the subjects or participants are randomly assigned to treatment or control (non-treatment) groups. This minimizes the risk of selection bias and isolates both the magnitude and direction of the treatment effect. Experiments are most easily conducted in laboratory conditions, but extension to the field can allow for exploration of a broader range of research questions and may provide greater realism.

True experiments are becoming increasingly popular in social science [189], and are commonly seen as the "gold standard" for determining causality [190]. They also benefit from broad consensus on what constitutes best practice. For example, Bloom [191] provides a useful overview of experimental designs for different contexts, including differing research questions. However, true experiments are not widely used within energy social science, even in areas where they appear feasible - such as the evaluation of energy efficiency programs [192]. This is partly because energy social science asks a wide range of research questions, only a portion of which can be answered through experimental designs. But it is also because experiments can be time-consuming and expensive to conduct (compared to desk-based research, or a study using a small sample of interview respondents) and can raise practical and ethical difficulties. For instance, it may not be possible to randomly withhold subsidies for energy efficiency improvements from qualifying applicants. True experiments can also have limitations, such as usage of small or unrepresentative samples, vulnerability to the Hawthorne effect (where participants behave differently because they are being observed), difficulties incentivizing replication studies, and a lack of guidelines for how to increase the reproducibility of results [193]. Indeed, some argue that experiments must move beyond the bias

Box 2
Defining validity.

> Researchers will inevitably be concerned with validity when they design, implement and interpret their study. Broadly speaking, and more in line with the positivist paradigm, validity relates to whether the result or interpretation is correct. Although the concept is most clearly applicable to experiments and quasi-experiments [180]—that is, studies of causation or explanation—it is also relevant to other quantitative and qualitative methods [181]. Shadish et al. [182] present four types of validity, the two most commonly discussed of which are internal and external validity. Internal validity relates whether the observed effects are due to the identified variable(s) and not some other factors, whereas external validity refers to the generalizability of the study's results to other groups, contexts or time periods. Researchers will want to consider both forms of validity within their research design—through considering alternative explanations for what they observe (internal validity), and assessing how current observations may or may not apply to other contexts (external validity). Hammersley argues that while concepts of validity are useful, they must be applied differently for different research questions, methods and intentions for the produced knowledge [183].

towards Western, educated, industrialized, rich and democratic (WEIRD) societies [194]. Furthermore, the laboratory setting of most true experiments is rather artificial, and the results may be difficult to transpose to real-world settings. Those defending experiments counter that many of these limitations can be mitigated through either careful research design or the integration of experiments with other, complementary research methods [195].

Where true experiments are impractical, it may be feasible to employ a "natural" or "quasi-experimental" research design, that includes treatment and control groups, but where allocation to those groups is determined by factors beyond the researchers' control [196–198]. The key to success in a quasi-experimental design is to ensure that the assignment to treatment or control group is not related to other determinants of the relevant outcome. If successful, this can obviate the need to specify and control for all confounding variables. Some of the most common approaches to quasi-experiments are summarized in Box 3.

Quasi-experiments encompass a range of research designs, of varying degrees of robustness and sophistication. A recent variant utilizes "living laboratories" to provide user-centered social experiments with the aim of testing a particular technology, solution, idea or policy in a real-world environment [200–203]. Distantly related examples include "transition experiments" and "governance experiments" [204–206]. Still other designs utilize more complex simulations, games, or competitions [207] to understand bargaining strategies, including those using the labels of "serious games" [208] (games with a purpose other than entertainment), "adaptable simulations" (games for learning) [209], and "gamification" (games for an educational purpose) [210].

The codes of practice we recommend for experiments and quasi-experiments include:

1) Clearly specify the experiment's objectives, type ("true" or "quasi") and predicted result or effect [211];
2) Follow best practice for experimental design that aligns with the research objectives, including selection of sample size, choice of setting (field versus laboratory) and management of control groups [212];
3) Ensure recruitment of participants to be as representative as possible for the purpose at hand [213] (e.g. first year psychology students may not always be representative);
4) Utilize random assignment where feasible and appropriate, and where not, follow best practice for quasi-experimental design [214];
5) Acknowledge limitations in external validity, and, where possible, use a multi-method approach to mitigate those limitations (e.g. combining with a large, representative survey);
6) Where possible, consider replication or repeated experiments to gain stronger evidence of causality [215].

4.1.2. Literature reviews

A literature review is a study or compilation of other research—typically of peer-reviewed literature, though non-academic studies can also be included. We consider three types of review here, flowing from most to least structured: meta-analysis, systematic reviews, and narrative reviews (summarized in Table 4). A meta-analysis combines quantitative results across a set of studies to draw conclusions about a specific topic of interest. A systematic review aims to provide a comprehensive, unbiased and replicable summary of the state of knowledge on a well-defined issue. A narrative review provides an exploratory evaluation of the literature or a subset of literature in a particular area. Meta-analyses and systematic reviews can each be further distinguished between a priori reviews that start with fixed criteria or search strings that do not change once the search begins, and

Box 3
Quasi-experimental approaches.

> *Instrumental variables*: This involves identifying a variable that is correlated with the treatment variable but otherwise independent of the outcome. This variable can then be employed within a two-stage least squared (2SLS) regression. The difficulty lies in finding a suitable variable, since "weak instruments" can lead to bias [199].
>
> *Difference in differences*: This exploits the availability of "panel data," where repeated observations are made on two or more groups at two or more points in time. Neither group receives the treatment in the first period, and only one group receives it in the second period. For example, the groups could be two (or more) US states, one of which introduces a carbon tax. The approach relies on the assumption that, after controlling for relevant variables, the outcomes in the two groups would have changed identically in the absence of the treatment. However, this assumption is not always valid.
>
> *Regression discontinuity*: This assigns subjects to treatment and control groups according to whether an observed, continuous variable lies above or below a threshold. For example, the threshold may be that used by regulators to define eligibility for a subsidy scheme. The variable may or may not be associated with the relevant outcome, but provided the association is smooth, no bias should result.
>
> *Non-equivalent control group*: This involves matching the subjects in the treatment and control groups as far as possible on the basis of relevant, observed variables (e.g. energy consumption, building type, location). Given limitations on data, this tends to be the weakest of the four approaches.

Table 4
An illustrative summary of three approaches to literature reviews.
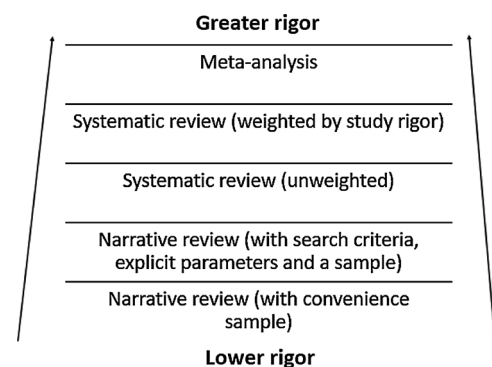Source: Authors.

Approach	Appropriate for	Limitations
Meta-analysis	Statistically aggregating quantitative results from a number of similar studies to increase the statistical power of tests and the precision of parameter estimates	Confined to quantitative evidence; does not bring insight into under-studied topics, or topics with more qualitative focus
Systematic review	Explicit and transparent methodology for synthesizing research results, including: clear specification of research question(s); systematic searching of the available literature; and applying explicit criteria for the inclusion or exclusion of studies. May also appraise the quality of included studies using transparent and standardized criteria.	Time-consuming and resource intensive (compared to a narrative review); focuses upon a narrow range of questions; biased towards quantitative research methodologies; unsuitable for addressing complex problems and policies; uses an "additive" approach to synthesis that neglects the complementary nature of different studies; narrow scope may prevent more in-depth insights.
Narrative review	Exploratory investigation of literature, involving less precise research objectives, a less systematic approach to article inclusion and allowing more in-depth qualitative insights to be obtained.	Prone to researcher bias; can selectively miss research; tends to place excessive reliance on individual studies and pays insufficient attention to methodological quality.

iterative reviews that modify search strings based on ongoing results, leading to repeated searches.

Meta-analysis is usually quantitative in nature, involving statistical analysis of the quantitative results from a series of comparable studies [216,217]. Aggregate results (or more rarely, individual data) can be pooled and analyzed with a meta-regression technique that estimates an overall effect size, while also explaining variations across studies (e.g. different samples or methods). There are several comprehensive guides to meta-analysis, which is now an established technique in many fields [218,219]. While the method is powerful, it is only appropriate for clear and precise research questions that have previously been addressed by a large pool of comparable quantitative studies. Put another way, meta-analyses may not be possible for some study types, and they do not always yield more useful results (for example if the included studies are too heterogeneous). Meta-analyses are common in fields such as medicine, but much less common within energy social science. There are exceptions, however, such as estimates of energy price elasticities [220], social influence effects for alternative fuel vehicle purchases [221], and the success of demand response programs [222].

Systematic reviews are also very structured, but are more descriptive and can include both quantitative and qualitative evidence. Such a review usually works in phases, such as: (1) crafting of explicit research questions; (2) systematically searching the available literature using defined search terms; (3) using explicit criteria for including or excluding studies; (4) determining and then executing a coding strategy or analytical protocol; and (5) analyzing or synthesizing the collected evidence. Compared to a typical narrative review, a systematic review aims to use an explicit and replicable research design, ensure comprehensiveness in the literature search, and reduce bias in the selection of studies [223]. Further, most systematic reviews give greater weight to methodologically rigorous studies, although not all meet this criteria. Some researchers even suggest that systematic reviews belong at the top of a list of most rigorous methods. For instance, when discussing reviews, Khalid et al. state that "reviews should never be done in any other way" (clearly placing systematic reviews as the method of choice) [224]. Further, Huebner et al. suggest that there may even be a continuum of "systematic-ness" in literature reviews, moving up from purely narrative reviews to systematic reviews, and finally meta-analysis at the top [225]. Fig. 2 is our own conceptualization of how such a continuum may look.

Systematic reviews can be applied to topics where both quantitative and qualitative evidence is relevant, experiments (true or quasi) may or may not be feasible, researchers are concerned with "what works" in what context, and multiple and competing factors are at play [228]. Examples of systematic reviews in energy social science include: an assessment of the cost impacts of intermittent generation on the UK electricity system [229]; a review of the evidence for a near-term peak

Fig. 2. A scale of rigorous methods for literature reviews.
Source: Authors, modified from [226,227].

in global oil production [230]; an analysis of the social acceptance of wind energy in North America [231]; and an analysis of the barriers to and opportunities of smart meter deployment in the UK [232]. The main drawback of systematic reviews is that they are resource intensive and time consuming. Systematic reviews are therefore not optimal in circumstances when resources are limitied or for fields where evidence is sparse or patchy[233]. Also, they are more suited to relatively narrow research questions rather than multidimensional problems; and they tend to employ an "additive" approach to synthesizing research results that can neglect the complementary nature of different studies and perspectives [234]. Further, a systematic review is not guaranteed to be comprehensive or unbiased—the inclusion and coding of articles is still sensitive to the researcher's selection of criteria and concepts.

Narrative reviews are the least structured and most common type of review, appearing in both review papers and the literature review sections of research papers. A narrative review synthesizes evidence familiar to an author on a given topic or theme, and is typified by the reviews published in *Annual Reviews of Environment and Resources*. Good narrative reviews will be comprehensive, and typically require an experienced author to uncover the nuances and themes of the relevant literature. The narrative review approach can be particularly useful for exploratory reviews that seek to synthesize insights from a variety of perspectives and disciplines, or areas where insufficient data exists to conduct a systematic review or meta-analysis. Further, a good narrative review will be organized in a way that is useful and easy to read: for example, by concept, theme, theory or discipline; or, if appropriate, by publication date [235–237]. However, narrative reviews typically lack transparency and replicability, especially if the author uses a "convenience" sample with no explicit criteria for inclusion [238]. Thus, narrative reviews can be more subject to bias compared to other

methods, mainly in the inclusion and exclusion of research and in the weighting of research evidence - or at least, that bias might be better hidden.

A final point relevant to all literature reviews is the need for careful use of citations. Many authors have had the experience of seeing their work cited, only to discover that their study has been misinterpreted (slightly or sometimes completely), or mixed-up with another study. Researchers thus need to be careful with the documentation and organization of papers and citations, treating these as carefully as their own data or analyses [239].

The codes of practice we recommend for literature reviews include:

1) Be as explicit as possible about the process of the review you use, explaining your rationale and approach;
2) Employ meta-analysis when there is a large number of comparable quantitative studies of the topic and the research questions are specific, clear and consistent;
3) Utilize systematic reviews to comprehensively summarize and interpret large bodies of quantitative or qualitative evidence on well-defined research questions, and when sufficient time and resources are available;
4) Undertake narrative reviews for exploratory and/or multi-dimensional research questions or when resources are more limited;
5) In all three approaches, be transparent: if applicable, report the sources/databases covered, the dates and time period examined, the search term(s) used, the languages searched, and whether any sampling of results was done (if the population of articles was too large);
6) Know your citations and references and ensure that you accurately utilize them.

4.1.3. Surveys and data collection

Surveys are a cornerstone of research in a range of disciplines, some of which have established criteria for best practice—though these are not always consistent with each other. Dillman's "tailored design method" provides one of the most accepted guides to survey research and is now in its fourth edition [240]. To set up this discussion, we first distinguish between the target population (the entire set of "elements" - such as individuals, households or organisations - that the researcher wants to learn about), the sampling frame (the list of elements that will be sampled from, e.g. a phone book or list of motor vehicle registrations), the invited sample (the subset of those elements selected from the sampling frame), and the realized sample (those that actually complete the survey and provide usable data). For example, a researcher might want to study a city of one million people (the population), and have a list of 100,000 motor vehicle owners (the sampling frame). They randomly select and invite 5000 of these vehicle owners (invited sample), and of those, 1000 end up completing the survey (realized sample). In this example, the response rate is 20% (1000 completes out of 5000 invites)—though researchers can vary in how they define and calculate response rate, so this should always be explained.

One key consideration for survey design is the mode employed to conduct the survey, which can include phone, internet, mail or in-person, or some blend of these. A number of publications outline the relative strengths and weaknesses of each, which vary for different research questions and target populations [241,242]. Internet surveys have become increasingly popular owing to their low-cost and versatility. Regardless of the survey mode, for many target populations it is difficult to find an appropriate sampling frame, and to recruit a realized sample of sufficient size and representativeness to achieve one's research objectives.

Dillman argues that researchers need to consider and minimize four types of error that threaten validity, namely: sampling error, coverage error, non-response error and measurement error. Unfortunately, many researchers focus almost exclusively on sampling error, which only

Table 5

The required sample size for obtaining an estimate of specified precision from different population sizes*.

Source: [243].

Population size	Sample sizes for the 95% confidence intervals					
	+/- 10%		+/- 5%		+/- 3%	
	50/50 split	80/20 split	50/50 split	80/20 split	50/50 split	80/20 split
100	49	38	80	71	92	87
200	65	47	132	111	169	155
400	78	53	196	153	291	253
600	83	56	234	175	384	320
800	86	57	260	188	458	369
1000	88	58	278	198	517	406
2000	92	60	322	219	696	509
4,000	94	61	351	232	843	584
6,000	95	61	361	236	906	613
8,000	95	61	367	239	942	629
10,000	95	61	370	240	965	640
20,000	96	61	377	243	1,013	661
40,000	96	61	381	244	1,040	672
100,000	96	61	383	245	1,056	679
1,000,000	96	61	384	246	1,066	683
1,000,000,000	96	61	384	246	1067	683

Note: * by 50/50 or 80/20 expected response split, for a binary question. Estimated sample sizes for simple random samples that include the finite population correction.

describes the lack of precision resulting from selecting a sample rather than surveying the entire population—often leading to the erroneous perception that large sample size is the primary or only indication of a rigorous survey method. Table 5 illustrates the relationship between population size, sample size and sampling error. For example, consider a researcher that aims to draw a random sample from a population of one million, and desires the result of a binary question (e.g. yes/no). If the researcher expects to have a 50/50 split in responses among respondents (50% yes, 50% no), and wants to know these observed proportions within a precision level of +/- 3% (at a 95% confidence level), the study would need a minimum random sample of 1067 respondents. It is this calculation that often leads to 1000 being considered the "magic number" for desired sample size among survey researchers.

However, the choice of appropriate sample size depends upon the research question (not to mention researcher resources, and accessibility of the population). Studies with descriptive research questions (e.g. the percentage of a population that holds a certain belief) may use Table 5 (or the calculations behind it) to anticipate the degree of precision a given sample size will attain regarding survey responses. Studies focusing upon tests of association or causality may employ more complex calculations, where the appropriate sample size depends upon the anticipated effect size, the desired significance level, the desired statistical power of the test and the expected variance of the explained variable [244,245]. For some causal or experimental studies, a very small sample size (e.g. n < 20) may be sufficient. Modest sample sizes (e.g., n < 100) may also be acceptable for studies trying to access a small population (e.g. a city or region) or the exceptional groups mentioned in Section 3.3.2. For example, if you want to assess the percentage of Russian citizens that support nuclear power, you will need a large, nationally representative sample of respondents. If, however, you want to undertake an exploratory study of how early adopters of smart homes in Wales feel about those technologies, a much smaller sample could be appropriate (e.g. 10–30 households). In all cases, the sample size needs to be considered in the context of the research objectives and the intended method of statistical analysis.

Despite the importance of sample considerations, we urge survey researchers to consider and balance efforts to mitigate sampling error

Table 6
Illustrative summary of methods for data analysis.
Source: Authors.

Approach	Typical applications	Limitations
Descriptive statistics or univariate analysis	Descriptive or exploratory	Involves a single variable or set of variables, validity largely depends on sample quality
Bivariate analysis (Pearson r correlation, ANOVA, chi-square)	Exploratory, justifiably small samples or limited data, or analysis of results from true experiments	Correlation only, vulnerable to omitted variable bias unless data derives from a true experiment.
Multiple regression (linear, logistic, MANOVA, ANCOVA, MANCOVA, etc.)	Explanatory, with clear hypotheses	Assumptions required for valid estimation (e.g. variables uncorrelated with the error term) are frequently violated and can lead to bias; problems with "data-mining" approaches
Structural Equation Modeling (SEM)	Explanatory or exploratory; relating to theories with multiple levels of causation	Subject to the problem of omitted variables, the importance of lower-order model components, and potential limitations of models judged to be well fitting
Factor analysis	Explanatory or exploratory; when calling for the collapsing or combining of variables	Simple confirmatory tests (e.g. Cronbach's alpha) are vulnerable to assumptions; variety of views about "best practices" for exploratory factor analysis
Cluster analysis (K-means)	Exploratory and explanatory; when needing segmentation of agents or cases	Not appropriate for tests of statistical significance, no clear consensus on how to select the number of clusters
Discrete choice analysis (multinomial, nested, mixed, probit, etc.)	Explanatory, with clear hypotheses; used for dependent variables that take a discrete number of values	Assumptions required for valid estimation (e.g. variables uncorrelated with the error term) are frequently violated and can lead to bias; can be too strongly embedded in rational actor theory
Latent-class modeling	Explanatory; also allows exploration of heterogeneity via segmentation	Similar limitations to discrete choice modeling; typically requires larger sample size due to model complexity

with efforts to minimize the other three categories of error identified by Dillman [246]. The second category is coverage error, where the sampling frame (e.g. a mailing list) is not fully aligned with the target population, i.e. it misses certain types of people and/or oversamples others. For example, a sampling frame of household telephone numbers would miss households without a telephone, and a traditional phone book could miss households that only use a cell phone. The third category is non-response error, where those that respond to the invitation (to become the realized sample) are systematically biased relative to the target population—say being higher income, older, or having a higher level of education. For example, a market survey of car buyers interested in electric vehicles could be more attractive to electric vehicle enthusiasts—since these are more likely to respond to the survey invitation, the realized sample may be biased. Survey results would then overestimate consumer interest in electric vehicles. Related to this is item non-response error, where a particular survey question is neglected by some subset of the realized sample – such as higher income households being more likely to refuse to report their income (again biasing the observed distribution of results relative to the target population). The final category is measurement error, where the survey instrument does not record the information that the researcher thinks it is recording, typically as a result of poor or confusing wording of questions or response categories. This final category moves beyond the sample to highlight the importance of careful design and pre-testing of the survey instrument itself.

In short, a rigorous survey research design should have an appropriate sample size, be representative of the target population (for descriptive research) and be effective in communicating questions and eliciting responses. The complexity of real-world research questions usually means that all four errors will be present in a survey project to some degree. However, rigorous survey researchers must address and manage such risks in their research design (within the limits of available resources), and report how they have done so in their article.

Thus, we propose the following codes of practice for survey data collection:

1) Consider and acknowledge the strengths and weaknesses of different survey implementation and sample recruitment modes (internet, phone, mail, convenience);
2) Aim to collect an appropriate sample size for the research objectives and context (achieving an acceptable degree of sampling error);
3) Examine and report how well the sample represents the target population (threatened by coverage error or non-response error)—especially for descriptive research objectives;
4) Carefully design and pre-test the survey instrument to maximize the accuracy of responses (minimizing measurement error);
5) Carefully interpret results according to the limitations of the realized sample.

4.1.4. Data analysis and statistics

Many studies will require statistical analysis of collected data, so researchers must be able to select the most appropriate statistical methods, apply those methods effectively and interpret the results correctly. This requires a firm grounding in statistical methods. The appropriate choice of method will depend upon:

- The nature of the research objective, which can be exploratory, descriptive, or explanatory [247]. Exploratory research does not have clear hypotheses and rarely requires statistical methods. Descriptive research simply summarizes the characteristics of the data (e.g. sample means or proportions) and only requires basic statistics. Explanatory research searches for relationships among variables, typically starting with clear hypotheses about those relationships and often requiring sophisticated statistical analysis. Most analysts caution against "data-mining," "p-hacking," or "reverse-engineering" a paper, where the researcher tests a large number of models and variables and works backwards to focus on relationships they find significant. But some traditions – such as the general-to-specific methodology in econometrics – view such approaches more favorably [248–250].
- Whether a relationship is analyzed and which type: univariate analyses confine attention to single variables, including estimates of

means, standard deviations and confidence intervals; bivariate analyses estimate the relationship between two variables, through correlation, ANOVA or a chi-square test of association; and multivariate analyses estimate relationships among many variables via multiple regression and other techniques.

- The types of variables to be analyzed, be they continuous, ordinal or nominal—which in some cases can be transformed from one type to another.
- The type of data to be analyzed, which can be cross-sectional (sample taken from a population at a given point in time), time-series (observations on several variables at regular intervals), pooled cross-section (cross-sectional sample from the same population taken at two or more intervals in time) or panel (similar to a pooled cross-section, but with data from the same units in each period). Further distinctions include aggregate versus disaggregate data (e.g. US states versus households) and different periodicities of time-series data (e.g. monthly, quarterly, annual).

Table 6 lists some major data analysis methods by their typical application and main limitations. For some research objectives, particularly descriptive research, a simple procedure might be warranted. For example, a survey of citizen support for a given climate policy might only require the reporting of the proportion of respondents in favor, along with a confidence interval. However, most statistical studies in energy social science are interested in the relationships between two or more variables. Bivariate analysis explores relationships between two variables, but typically provides only limited insight due to the potential for the identified relationships to be spurious, owing to omitted variables. The exception is data from a true experiment (Section 4.1.1), where bivariate analysis of the relationship between treatment and outcome can be interpreted as causal, due to the process used to generate the data.

Some research texts present data analysis methods from least to most rigorous. Fig. 3, for example, proposes an arrangement of data techniques. For most studies, multivariate analysis will be required to produce meaningful insights, although the rigor of individual applications may vary widely depending upon both the nature of the data and the care taken by the analyst - for example, in conducting model specification tests.

Among multivariate analyses, the most common approach is multiple regression, which explores how a number of independent (or explanatory) variables are associated with a single dependent variable. Techniques such as MANOVA are a simply a subset of multiple regression, but are widely used in disciplines that employ true experiments, such as social psychology. In contrast, economics relies almost exclusively upon multiple regression. Linear or non-linear regression is used for continuous dependent variables, while logistic regression is used for categorical dependent variables. The primary advantage of multiple regression is that researchers can explore hypotheses about the relationship between two variables (e.g. how household income predicts support for climate policy), while controlling for (holding constant) other variables that might also matter, such as respondent age, gender and political affiliation. Although such analyses can be powerful, researchers frequently pay insufficient attention to the various assumptions that must hold for different methods to give unbiased results. Nearly any introductory statistics or econometrics textbook will explain these assumptions, together with the tests required and strategies available when those assumptions do not hold [254]. These issues are particularly important when using secondary data sources (such as government statistics on energy consumption and prices) since these have multiple limitations that are largely beyond the researchers' control – such as short time series, measurement error and missing or endogenous variables. Much of the sophistication within econometrics results from attempts to overcome such problems – for example, econometricians have developed "cointegration" techniques to extract the relationship between variables that share a time trend [255]. However, since no amount of analytical sophistication can adequately compensate for poor quality data, there is an increasing trend towards the use of panel data (which permits more robust inferences) and quasi-experimental techniques [256,257].

Table 6 also lists some more advanced techniques, along with their main limitations. We can't possibly mention all methods, so we only highlight a few that have proven popular in energy social science. For example, structural equation models can be used to explore complex relationships among variables, particularly when a theory or hypothesis proposes several layers of causation [258]. For example, it may be hypothesized that a person's values influence their beliefs about a particular energy technology, which in turn influences their likelihood of purchasing that technology. While this approach is powerful, rigorous analysts need to use theory carefully to guide their inquiry [259]. Factor analysis includes methods that collapse or group similar variables into a single measure [260] (e.g. constructing a composite measure of pro-environmental attitudes based on several survey questions), and is used extensively within social psychology [261,262]. Cluster analysis groups agents or cases in such a way that members of the group are more similar to each other than to those in other groups (e.g. identifying consumer segments) [263,264], but the most popular technique (K-means clustering) cannot be used for tests of statistical significance, and there is no universally accepted method to select the "best" number of clusters. Discrete choice modeling is a particular form of logistic regression that explains and predicts choices between two or more discrete alternatives, such as between an energy efficient and inefficient appliance, based upon the characteristics of the different choices, the characteristics of the relevant actors (e.g. households) and other relevant variables. This approach has proven particularly popular in economics and transportation studies [265]. Discrete choice models were originally informed by expected utility theory [266], but increasingly use other social theories as well [267,268]. Finally, latent-class models are a particular type of discrete choice model that explicitly represent heterogeneity among individuals, splitting respondents into a number of similar classes or segments, and estimating choice models for each segment [269,270].

Appropriate applications of each of these methods must consider many more issues than we can cover here, and the rigorous analyst will need to become familiar with textbooks and papers relating to their chosen method.

In summary, the practices of the rigorous data analyst include:

1) Effectively match the data analysis technique to the research question and type of data available;
2) Where multiple methods are appropriate, consider and acknowledge their individual strengths and weaknesses;
3) Where data are available, conduct more sophisticated and robust analysis of association (e.g. multivariate rather than bivariate);
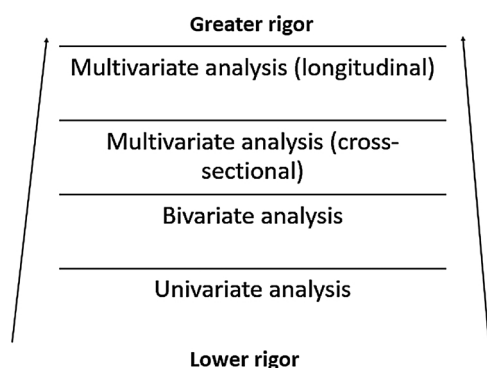4) For explanatory or comparative research questions, state hypotheses

Fig. 3. A scale of rigorous methods for data analysis and statistics.
Source: Authors, synthesized from [251–253].

Table 7
Strengths and limitations of different types of quantitative energy models.
Source: Authors, based partly on [275–280].

Broad class	Type	Examples	Claimed strengths	Critical limitations
Bottom-up	Optimization	TIMES, MARKAL, MEDEE, MESSAGE	Detailed and disaggregated representation of technologies; estimates "optimal path" for climate mitigation,	Generally lacking in behavioral realism
Top-down	Computable General Equilibrium (CGE)	EPPA, MSG-4, ELIAS, AMOS, AIM/CGE	Economy-wide, represents sector interlinkages and macroeconomic feedbacks (e.g. GDP, employment)	Not technologically explicit; often a lack of empirical basis for parameters
	Input/output	NAMEA, SDA	Economy-wide, represents sector interlinkages	Unrealistic for modeling large shocks or anything beyond short-term
Simulation	"Hybrid"	CIMS, NEMS, IMAGE, GCAM, MESSAGE-MACRO	Aim to have behavioral realism and to combine strengths of bottom-up and top-down models	Difficult to estimate behaviorally realistic parameters; model integration can be challenging
	System Dynamic	En-ROADS	Captures feedback loops and non-linearities	Tend to lack a firm empirical basis for behavioral assumptions
	Agent Based Models	EMCAS, N-ABLE, NEMSIM, MA3T	Represents heterogeneous agents (consumers, policymakers, companies, etc.)	Tend to lack a firm empirical basis for behavioral assumptions
Climate change integrated assessment models (IAMs)	Simplified cost-benefit IAMs	DICE, FUND, PAGE	Can simulate feedbacks between natural systems (e.g. climate) and social systems (energy systems and economy)	Social and natural components often oversimplified; integrated analysis face large uncertainty in monetized damage costs and discount rates
	Cost-effectiveness IAMs	IMAGE, GCAM, MESSAGE	Globally comprehensive versions of models noted above (e.g. optimization, macro-economic, hybrid)	Social/demand component has same drawbacks as model categories above; natural component often oversimplified

clearly up front, informed by theory and avoid re-working hypotheses to fit the results;

5) Balance the objectives of statistical performance with the interpretability and usefulness of the results;

6) Carefully distinguish between analyses of association versus causation;

7) Distinguish clearly between statistical significance and practical significance-where the latter relates to whether the difference is large enough to be of real-world importance.

4.1.5. Quantitative energy models

Quantitative energy models have held a central place in energy research for decades. Such models are computer-based, and are used for a variety of purposes, including exploring the range of possible futures under different assumptions and assessing the impact of particular policy interventions (e.g. carbon pricing or technology mandates). The different types of energy models can be classified in a variety of ways [271–274], including: geographical coverage (e.g. local, national, regional, global); sectoral coverage (e.g. single sector, multi-sector, whole economy); scope (e.g. energy demand, energy supply, whole economy); methodology (e.g. econometric, general equilibrium, simulation, optimization); and time horizon (e.g. single year, 5–15 years, decades, century).

For simplicity, Table 7 distinguishes four broad categories of model and highlights their main strengths and weaknesses. As with other research methods, the appropriate choice of model depends upon the research question, and therefore it is important to acknowledge the limitations of each model type – though model-based articles often neglect such acknowledgement and comparison. Given our focus on energy social science, we place particular weight on behavioral realism: that is, better energy models will have a strong empirical basis for their parameters, include some degree of heterogeneity between relevant groups, and/or represent the potential for a broad range of actor motivations (which, for many actors, will move far beyond financial

motivations).

We first distinguish "bottom-up" from "top-down" models, a distinction that represents the historical basis of many models. Although these categories have blurred in the last two decades (leaving some to discard their usage altogether) [281], we believe the broad distinctionis still a useful starting point. First is "bottom-up" models, a term that is often equated with optimization (typically "linear programming") models that have their origin in engineering and operations management. The term "bottom-up" is used because these models explicitly simulate the operation of individual energy-using technologies (the "bottom"), which are aggregated across individual sectors (e.g. electricity generation, households) or the energy system as a whole to give total energy use and emissions (the "up") [282]. These models simulate the ageing and replacement of technologies, with investment decisions being determined by capital costs, fuel prices, policy interventions and other factors. Bottom-up models usually include a large number of current and potential future technologies and simulate the "optimal" means of attaining some goal (typically minimizing discounted costs over the modeled time horizon) subject to constraints (usually including environmental goals). However, this optimization assumption is also the main weakness of conventional bottom-up models, as consumers, energy suppliers and other actors are frequently depicted as hyper-rational decision makers operating with perfect information and foresight and motivated purely by financial costs– assumptions contradicted by empirical research on human behavior [283]. However, significant efforts have been made to improve the behavioral realism of such models, including attempts to incorporate "myopic" decision-making [284], heterogeneity, intangible costs and benefits and social influences [285].

In contrast, "top-down" models are macroeconomic and aggregated in nature, and are commonly used to simulate how changes, or "shocks" in one sector (e.g. a carbon tax on electricity generation) impact the entire economy, including changes in prices, investment, employment and GDP [286]. Most common are computable general equilibrium

(CGE) models, which simulate regional or national economies by combining a social accounting matrix (showing transactions between different parts of the economy) with equations for the behavior of each sector, under the assumption that the economy tends towards an equilibrium. CGE models are calibrated to the economic transactions in a base year and make the assumption that firms maximize profits and consumers maximize utility [287]. System responses in a CGE model are strongly influenced by the assumed elasticities of substitution between factor inputs (e.g. capital and energy) and different types of consumption good [288]. Although the results are highly sensitive to these assumptions, their empirical basis is typically weak [289,290]. The aggregate nature of top-down models means that they do not represent specific technologies or actors, but instead use abstract relationships such as production functions [291,292]. This abstraction leads to the common perception of CGE models as "black boxes", lacking transparency regarding the assumptions and processes that lead to a given finding – though admittedly, most complex energy-economy models can suffer a similar problem. In most cases, the "black box" issue can be mitigated in part by comprehensive sensitivity tests and elaborations in the documentation of the economic mechanisms contributing to the observed results. This category also includes input-output models (I-O), which can be seen as simplified CGE models with a fixed production structure and no scope for substitution. I-O models benefit from simplicity and transparency, but are unable to model price changes, supply constraints and other market feedbacks and are only suitable for investigating the impact of relatively small system shocks over the short-term.

A third category may be called simulation models, grouping a variety of models that do not seek to optimize a system according to goals or macroeconomic assumptions—but instead seek to "simulate" real-world patterns of behavior. These models vary widely in structure and assumptions, making it particularly important for modelers to communicate those assumptions. In recent decades, so-called "hybrid" approaches have emerged, integrating aspects of top-down and bottom-up models, and attempting to balance the strengths of technological detail, behavioral realism and macroeconomic feedbacks [293,294]. Indeed, most widely used energy-economy models have either a bottom-up or top-down origin, but have since moved to some degree of hybridization. Methods have also been developed to improve the representation of consumer behavior and preference change in such models; for example the CIMS model draws from stated and revealed preference choice models to assign behavioral parameters representing car buyer preferences [295,296]. In turn, CIMS has been shown to produce more realistic estimates of the costs of emission reductions [297]. Similarly, the REPAC-IESD model pairs empirically-derived discrete choice models (one of vehicle purchase, and one of electric vehicle owner enrollment into a charging program) with an electricity-utility dispatch model, finding that the societal benefits of vehicle-grid-integration are lower than indicated by optimization models [298].

Another type of simulation model – systems dynamics - represent complex systems by means of stocks, flows, feedback loops, and time delays. It simulates the non-linear behavior of those systems over time – including phenomena such as increasing returns, path dependence and tipping points [299,300]. The systems modelled can range in scope from individual organizations to the global biosphere and can incorporate a wide range of assumptions about system behavior [301,302]. However, despite their long history, systems dynamics models have not been widely used in energy social science, in part due to their complexity and the lack of a firm empirical basis for the relevant assumptions. We also include agent-based models in this category, which are highly disaggregated models that simulate the behavior and interactions of multiple individual agents (e.g. firms, consumers, policymakers). Behavioral realism can vary widely in agent-based models, depending on how the modeler chooses to represent the determinants of decision-making, and whether there is an empirical basis for the parameters used. In contrast to system dynamics models, agent-

based models are becoming increasingly prominent in the energy field [303].

A final category is integrated assessment models (IAMs), a term that is sometimes applied loosely to any approach that combines more than one model—making it important to communicate what exactly is "integrated". Here we refer mainly to climate change IAMs, which can be further split between relatively simple cost-benefit IAMs (such as DICE, FUND and PAGE), and the more complex cost-effectiveness IAMs (including three already-noted models: IMAGE, GCAM and MESSAGE). The cost-benefit IAMs rely on very simplistic representations of both social and natural systems, and in some cases can be run with a single spreadsheet (e.g. DICE). Such IAMs have been widely used to estimate and monetize the damage caused by climate change and thereby to estimate the welfare impacts of different mitigation options. Specifically, they can explore the interlinkages and feedbacks between natural and social systems: for example, how economic activities lead to increased greenhouse gas emissions, which warms the climate and in turn create damages that impact the economy (e.g. sea level rise and increased prevalence of drought and storms). But this approach is controversial, owing to the highly simplified assumptions required, the enormous uncertainties about the magnitude of climate damages, the philosophical difficulties associated with monetizing those damages and the unresolved debates about the appropriate choice of discount rate [304–306].

In contrast, the complex cost-effectiveness IAM models integrate one of the previously mentioned categories of socio-economic model (optimization, macro-economic, simulation or hybrid) with one or more natural science models - usually a climate model, and sometimes other ecological or land-use models as well. Due to this integration, such IAMs tend to be highly complex, and are typically constructed and maintained by large groups that specialize in such models, such as the International Institute for Applied Systems Analysis or the researcher teams informing the Intergovernmental Panel on Climate Change. The unique strength of such IAMs is that they are globally comprehensive, accounting for all types of greenhouse gas emissions from all emitting sectors—which can then provide useful inputs into climate models of radiative forcing and temperature change. However, since the social science component of complex IAMs are equivalent to one of the modeling types noted above, they suffer the same drawbacks. Further, because integrating several sub-models require substantial computing power, the natural science models used in these IAMs tend be more simplistic than a dedicated climate model.

Based on the summary of energy models detailed above, we conclude that good practices of the rigorous modeler include:

1) Carefully select a model type based on its suitability for the research objectives (including data quality and availability), rather than prior familiarity;
2) Consider and acknowledge the strengths and weaknesses of different model types, even if only one is used;
3) Aim for a parsimonious and useful model that avoids excessive complexity (avoiding perceptions of a "black box");
4) Maximize transparency in the structure and operation of the model and in the selection of model parameters;
5) Seek a firm empirical basis for model assumptions and, where appropriate, strive towards behavioral realism;
6) Conduct sensitivity tests and investigate and acknowledge uncertainties in the results.

4.1.6. Qualitative research

Qualitative research methods are particularly suited to inductive and interpretive approaches. Inductive approaches begin with empirical observations and seek to identify new insights and categories, and to generate rather than test hypotheses [307]. Interpretive approaches aim to interpret the experience of individuals and to identify the meanings that those experiences hold, rather than looking only to

Table 8

An illustrative summary of qualitative methods of data collection and analysis.

Source: Authors.

	Category	Appropriate for :	Limitations
Qualitative methods of data collection	Interviews (semi-structured or unstructured)	Accessing individual stories, understandings, explanations and meanings	Vulnerable to interviewer bias and social desirability bias
	Focus groups	Accessing collectively formed ideas and meanings; testing theories at reduced cost	Vulnerable to interviewer bias; collective patterns of response might differ from individual patterns
	Direct (naturalistic) observation	Observing actual behavior (rather than relying on self-reports)	May misinterpret actor interpretations and meanings
	Participant observation (ethnography)	In-depth access to a culture or context (observation and interaction)	Very time intensive (months or years); immersion may not be possible
	Documents or media	Accessing social discourse as presented by particular stakeholders	Limited to the perspective, agenda and biases of those who produced the documents
Qualitative methods of data analysis	Content analysis	Well-structured research questions, with a priori analytical categories	Requires highly-structured framework; can lack depth of analysis
	Narrative or discourse analysis	Accessing greater depth, explanation and meaning	Methods vary widely; difficult to make transparent
	Grounded theory	Developing new theory; investigating empirical topics with very little literature or theory to draw from	Methods vary widely; difficult to make transparent

establish causal inferences [308]. However, qualitative methods can also support other forms of enquiry.

Qualitative methods are sometimes attacked for lacking the widely-accepted standards of rigor associated with some quantitative disciplines and methods. However, this need not make qualitative research less rigorous and there have been multiple efforts to establish more robust standards for qualitative rigor [309,310]. As with all research methods, qualitative research needs to be designed to suit the intended research objectives [311], and these objectives often differ in fundamental ways to those addressed by quantitative methods.

Table 8 summarizes four approaches to collecting qualitative data and three approaches to analyzing that data. The most common approach to data collection is qualitative interviews, which may be either semi-structured or unstructured; implemented with individuals or small groups (e.g. households); and targeted at either the general population or particular stakeholders (often in the form of "elite" or "expert" interviews, see Section 3.3.2). Interviews provide access to people's experience, motivations, beliefs, understandings and meanings – often providing a deeper understanding than surveys and allowing follow-up and more probing questions [312]. These attributes apply equally to stakeholder interviews, but these raise the additional challenge of determining how the interviewees' perspective relates to that of the organization they represent (e.g. government agency, company, NGO).

While interviews are generally effective at eliciting individual perspectives, focus groups allow the elicitation of perspectives from groups of individuals (typically five or more), leading to more socially negotiated responses. Perhaps due to their association with market research, focus groups are often seen primarily as a low-cost method (cheaper and quicker than a series of individual interviews) or an initial step in a larger study (e.g. leading to a larger quantitative survey) [313]. However, focus groups offer their own unique strengths, namely by constructing a social context in which participants can collectively generate, negotiate and express perceptions and meanings—though of course, a rigorous researcher must understand and acknowledge the limitations of that context [314].

The qualitative (that is, semi-structured or unstructured) nature of both interviews and focus groups makes it difficult to code answers, and responses will vary significantly between different persons and groups. As with any face-to-face data collection method there is also the risk of bias, including a tendency for participants to provide responses that they see as socially desirable, or desirable by the interviewer. Also, as

with surveys, interview participants may find it difficult to describe their behaviors, responses or motivations. More generally, effective implementation of qualitative interviews and focus groups requires the interviewer to develop a very different set of skills to those required for quantitative data collection methods [315].

The three remaining methods of qualitative data collection can avoid or mitigate the challenges of interviewer-participant interaction. The first two, direct observation and participant observation, involve the witnessing of relevant behaviors of individuals or groups [316]. Direct observation (or naturalistic observation) is unobtrusive by design, and might occur, as examples, in a study of environmental conditions at facilities, buildings, and other institutions [317]. In contrast, participant observation is more in-depth, describing studies where the researcher participates and becomes somewhat immersed in the relevant culture or practices over a long period of time. Researchers will interact directly with subjects, typically in day-to-day contexts, in a sense combining aspects of direct observation with unstructured or semi-structured interviews. However, such participant observation can be resource intensive, requiring months or even years of the researcher's time. The final category we consider is analysis of documents, such as reports, letters, websites and news media. Such data sources can provide insight into the information, frames and storylines presented by different actors, as well as the social interactions among them [318].

Qualitative data collection also raises questions of "sample" size—but sample is in quotations because the objective is rarely to draw a random sample from the population. Qualitative samples tend to be "purposive", that is, intending to access a variety of experiences to fit the purposes of the study [319]. Unfortunately, there are few guidelines on how many cases is "enough" and no equivalent to the calculations of sampling error used for quantitative survey research. Some qualitative researchers argue that "less is more" in terms of sample size, since depth is more important than breadth [320]. But there can also be value in larger samples, especially if that increases the breadth of perspectives, since this can strengthen both internal and external validity. Further, qualitative studies that compare samples from different cases, regions or settings can frequently produce more useful results (see Section 4.1.7). But that said, qualitative "sample" size needs to be examined and explained for each study's unique research objectives.

As with data collection, the analysis of qualitative data can take a range of forms – a feature that may have contributed to the perception that qualitative research lacks clear standards for analytical rigor. Here

we mention three broad types of data analysis that represent different degrees of structure—acknowledging that the diversity is greater than we can demonstrate here, and that many qualitative studies use no formal methods of data analysis at all. The most structured approach is content analysis, which involves coding samples of interview or focus group transcripts, documents and communication records with the aim of systematically identifying categories, themes and patterns and reporting these numerically or graphically [321,322]. Content analysis is most useful for studies that start with a clear theoretical framework or set of expected categories. However, it is not always effective for richer, deeper analysis or narrative description [323].

Richer analysis can be achieved through narrative analyses which seek to analyze text or utterances with the aim of identifying "story-lines" that particular actors or groups use to frame (i.e. perceive and/or communicate about) a topic or experience [324–326]. The objective here can be interpretive, or explanatory in the sense of linking cause and effect. Narratives can be identified at an individual level (e.g. how consumers explain their purchasing behavior) [327], or more broadly for formal or informal social groups (e.g. how oil companies respond to "attacks" from environmental groups) [328]. Discourse analysis can be even more sophisticated, attempting to capture how narratives and rhetoric coalesce into stable meaning systems, institutional practices, and power structures that can constrain or shape agency [329].

Finally, an example of the least structured analytical approach is grounded theory, which seeks to integrate the formulation of theory with the analysis of data, typically iteratively [330]. This research is called "grounded" because researchers seek to avoid wedding themselves to a particular theory before they begin their investigation, instead "grounding" their analysis inductively in the data itself [331,332]. One particular challenge for grounded approaches is that they appear in a number of forms, each with different descriptions and guidelines, across several sub-disciplines [333,334].

In summary, the practices of the rigorous qualitative researcher include:

1) Effectively match research objectives to the appropriate means of data collection;
2) Also match research objectives to the type of analysis (such as content analysis, narrative analysis, discourse analysis, or grounded theory);
3) Provide detail about the methods used - such as sample size, questions asked, interview duration, demographic details of respondents, whether results were transcribed, whether data is anonymized or attributed, etc.;
4) Clearly explain and justify the strengths (and weaknesses) of the chosen methods;
5) Include more data (more "sample") when interviews or focus groups are meant to access a wide range of experiences in a diverse and/or large population (e.g. a nation);
6) Use the qualitative data in an effective way within the manuscript - for example, by providing illustrative quotations or explaining example observations.

4.1.7. Case studies and cross-case comparisons

Case studies involve in-depth examination of particular subjects or phenomena (e.g. individuals, firms, cities, policies, adjustment to a new technology) as well as related contextual conditions, often using multiple sources of evidence (e.g. documents, interviews, direct observation) [335]. The most cited guide to case study research is by Yin [336], who recommends the use of case studies for "how or why" questions about contemporary phenomena where the researcher has little control over events. However, case studies are equally appropriate for historical investigations. Case studies are commonly employed within energy social science, but the standards of rigor vary widely [337–340]. We start by considering several dimensions: type, single versus comparative, temporal variation and spatial variation.

Table 9 summarizes six broad types of case study [341]. Typical case studies investigate common, frequently observed, representative, and/or illustrative cases. Examples include case studies of the energy transition in Germany [342,343], renewable portfolio standards in the United States [344,345] and climate change adaptation in Bangladesh [346,347]. Diverse cases attempt to demonstrate maximum variance along a relevant dimension, so they illuminate the full range of important differences. These capture the full variation of the population, but do not mirror the distribution of that variation. Examples include the nuclear phase out in Germany contrasted with the rebuild of nuclear in the UK [348], or a comparison of energy transitions in Mexico, South Africa, and Thailand [349]. Extreme cases look for deviant, outlier, or unusual values of some explanatory or explained variable, or an example that illustrates a rare but important occurrence. Essentially, they look for "surprises." Examples include case studies of the Chernobyl nuclear accident in 1986 [350,351] or the Fukushima accident in 2011 [352], Iceland's adoption of geothermal energy [353]; Denmark's ambitious wind energy program [354]; Brazil's ethanol program [355]; and the Deepwater Horizon oil spill [356]. Influential cases seek to challenge or test the assumptions behind a popular or well-established case in the academic literature, say by challenging typical cases. Sticking with our examples, this would include critiques or alternative explanations for the energy transition in Germany [357,358], renewable portfolio standards in the United States [359] or climate change adaptation in Bangladesh [360]. The most similar method chooses a pair of cases that are similar on all measured explanatory variables, except the variable of interest. An example would be the progression of the Canadian and American nuclear power programs, which began around the same time in similar market economies but resulted in entirely different designs (light water reactors versus naturally fueled CANDU reactors) [361]. The most different approach is the inverse, and refers to cases where just one independent variable as well as the dependent variable co-vary, and other independent variables show different values. An example is contrasting the Chinese nuclear program with that of India (which began at an entirely different time and under a different economic system) [362].

The second dimension to consider is single versus comparative case studies. Single cases are useful for exploration and for generating hypotheses - for creating new conjectures in a sort of "light bulb" moment.. Single case studies tend to be evidence-rich, allowing a range of relevant factors to be measured and assessed and allowing a consistent and coherent narrative and argument. A good example would be Geels' historical analysis of the transition from sailing ships to steamships [366]. By contrast, comparative cases are confirmatory and good for testing a hypothesis, or for refuting some of the conjectures arising out of single cases. A good example would be Oteman et al.'s comparative study of the conditions for success in community energy [367]. External consistency is dominant, and comparative cases are useful for examining causal effects beyond a single instance. Empirically, comparative cases must be similar enough to permit meaningful analysis. Comparative case studies thus have greater variation but frequently also less depth since not all relevant factors can be examined.

The third dimension to consider is whether a cross-case comparison requires temporal or spatial variation [368,369]. Spatial variation (across different countries, regions, scales) can provide diversity but also challenge comparability of results. Temporal variation can permit more natural (less artificial) boundaries around analysis as researchers can include as many relevant temporal events as needed, but may require more complex analysis to capture the greater complexity of data. Combinations of spatial and temporal variation can only enhance these strengths and weaknesses.

These thoughts lead us to the following codes of practice for case study research:

1) Carefully consider whether to use a single case (deep, exploratory) or comparative cases (broad, confirmatory), as well as whether and

Table 9
An illustrative summary of case study types, strengths and limitations.
Source: Authors, modified from [363–365].

	Category	Appropriate for:	Limitations
Type	Typical	Commonalities or representative occurrences	Focus on "average" instances can obscure diversity as well as outliers
	Diverse	Maximum variance or a range of differences	Can be difficult to compare and synthesize findings
	Extreme, illustrative, or deviant	Unusual or unique events, outliers or surprises	Less probable nature can make it difficult to draw common insights or recommendations
	Influential	Challenging popular or well-established cases	Requires one to first understand the case they are refuting
	Most-similar	Comparative, isolating the role of one variable (variation in only one variable)	Can be hard to find and identify
	Most-different	Comparative, identifying range of potential scenarios, or "boundaries" of extremes (variation in all but one variable)	Can be hard to find and identify
Number of cases	Single	Exploratory and descriptive work, hypothesis testing, deep or thick descriptions	Lack of external validity, provides limited insights (needs strong justification to be publishable), lack of breadth
	Comparative	Explanatory, hypothesis generating, broader generalizability	Requires similar access to data, challenges in isolating variables of interest, limited depth
Spatial variation	Yes	Emphasizing difference can enhance understanding of complex spatial variables	Variation within cases or across cases can reduce generalizability
	No	Uniformity among countries or geographic scales	Homogeneity can enhance generalizability but may force artificial "fits"
Temporal variation	Yes	Emphasizing difference can enhance understanding of complex temporal variables	Requires more complex diachronic methods (such as time-series cross sectional analysis or hierarchical time-series analysis)
	No	Uniformity and homogeneity among time periods	May require one to artificially bracket or confine research

how the latter will vary spatially or temporally;

2) Have a well-defined unit of analysis (a well-defined case or cases), with clear boundaries, consistent propositions and measurable dependent and independent variables;
3) Specify and justify the type of case study chosen, and justify single case studies to warrant publication;
4) Acknowledge the uniqueness (or generalizability) of the chosen case or cases;
5) Carefully interpret results according to the limitations of the evidence and acknowledge rival hypotheses and explanations.
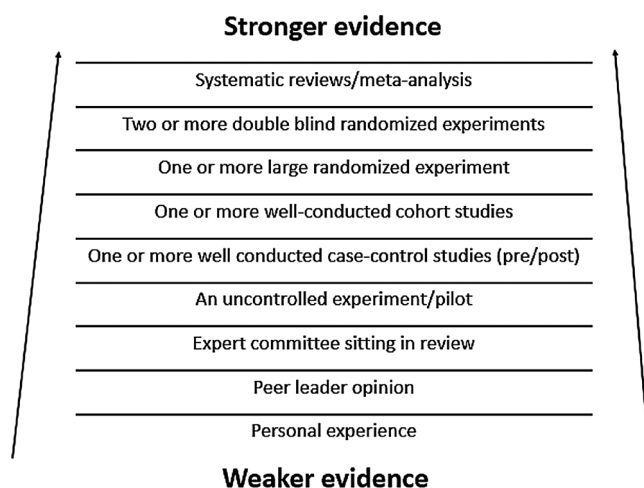
Fig. 4. Hierarchy of evidence for experimental or quantitative research.
Source: Authors, modified from [371,372].

4.2. Beware of hierarchies of evidence

Although we recommend a "codes of practice" approach to rigor, there are some disciplines, communities, and approaches where "hierarchies of evidence" are utilized to determine the strength of a particular study. The concept of hierarchies is most prominent in the health and medical literatures as part of developing concepts of "evidence-based research" or "evidence based policy and practice" and has since expanded to other fields such as social psychology and behavioral economics. The initial hierarchy (Fig. 4) is most relevant to research based on experimental designs, and it epitomizes a positivist view, placing personal experience at the bottom (the lowest level of the hierarchy) moving up through uncontrolled experiments to cohort studies and then multiple double blind experiments and randomized controlled trials, and with meta-analysis of randomized controlled trails as the "gold standard" [370].

Similarly, although less prominent, Daly et al. [373] have proposed another hierarchy for qualitative research and case studies with personal experience or a single qualitative case study at the bottom, descriptive studies in the middle, and conceptual or generalizable summaries or analyses of cases at the top. We have modified this hierarchy in Fig. 5 by adding more details about types and variation within case studies.

These hierarchies of evidence have at least two strengths. They are transparent about expectations in a given field, being exceptionally clear about what constitutes "good" or "better" research among peers in that discipline. Second, the implication that different methods can lead to cumulative impact, where studies can serve as the building blocks for others, can be useful and perhaps effective in moving towards a common understanding of certain, specific phenomena in a given field. For communities and disciplines that subscribe to such hierarchies, research methods at the lower levels—notably anecdotal experience,
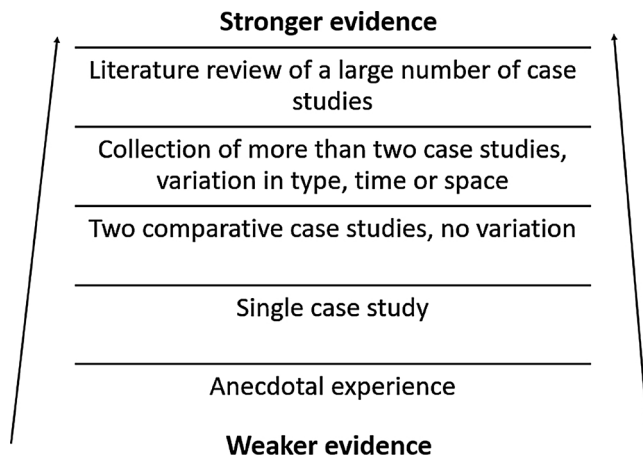
Stronger evidence

Literature review of a large number of case studies
Collection of more than two case studies, variation in type, time or space
Two comparative case studies, no variation
Single case study
Anecdotal experience

Weaker evidence

Fig. 5. Hierarchy of evidence for qualitative case studies.
Source: Authors, modified from [374–376].

uncontrolled experiments, pilots, or single case studies—are not necessarily seen as being inferior to "higher" methods or having no value. Indeed, moving up the hierarchy is not possible unless others lay the bricks at the base of the period; meta-analysis for instance depends on the single cases or cohort studies placed lower in a hierarchy.

However, these hierarchies are positivist by nature, and tend to reflect and propagate the narrow views of a particular discipline. Some disciplines have been known to rigidly subscribe to such hierarchies, systematically rejecting work that uses methods from a "lower level". On a related note, the hierarchical view may reinforce the unfortunate notion that quantitative research is necessarily more rigorous, valid, or just plain "better" than qualitative research. As we argue throughout this paper, we favor a more neutral perspective on rigor—identifying codes or principles that improve the quality of each type of social research method. Ultimately, researchers will have to decide which view better aligns with their perspective - taking into account their objectives and disciplinary affiliations. But in general, we advise caution with regard to hierarchies of evidence and recommend the broader codes of practice summarized above.

4.3. Appreciate balance, appropriateness and multi-method designs

Excellent or at least effective research requires a balance between the codes of practice we mention above. By balance, we mean that studies should not focus solely on maximizing one criteria of rigor, e.g. having an enormous sample size, using a particularly sophisticated simulation model, or providing a particularly "thick" description of a case study—at least not just for the sake of doing so. More generally, and perhaps contradictorily, academic research has been criticized for placing too much emphasis on rigor at the expense of impact or creativity—leading to careful but boring research [377] with little social relevance [378].

Instead, the effective use of each method requires tradeoffs. For example: large sample sizes can be costly, and are not necessarily representative; complex energy models can lack transparency, be difficult to parameterise and add uncertainty; and in-depth analysis of a case study might be too detailed to permit extraction of practical, generalizable insights. In short, there are always tensions in research design, which rigorous researchers will consider, and effectively (and humbly) communicate in their research.

Another theme that runs throughout our proposed codes of practice is appropriateness: the methods used must be well-suited to the research questions and research objectives. This consideration applies to the overall mode of inquiry (positivist or intepretivist), the research method applied (e.g. the seven we summarize, or some combination thereof), and the specific research design, including level of sophistication and depth of

analysis. It is not possible to produce a complete guide of how to work through this "matching" process—though we provide some guidelines here. Overall, we argue that no method itself is necessarily "best", or "good" or "bad" – rather it all depends on the context and goals of the project. That said, we have identified certain principles or codes that should lead to higher quality research.

In considering balance and appropriateness, we emphasize that some (perhaps even many) studies can involve more than one research method. A paper could start with a narrative review to determine a gap and justify or frame a research question before attempting to answer it with a case study that draws from data collected via qualitative interviews. Another study could begin by surveying a group of actors to solicit their perceptions and expectations, then conduct semi-structured interviews with a subset of that sample to elicit richer, in-depth narratives of how those actors connect those perceptions with their identity and lifestyles.

Mixed-method approaches hold particular promise, given that the two rough classes of inquiry—quantitative and qualitative—have particular advantages and disadvantages. Quantitative methods are very good at validating theories about how and why phenomena occur, testing hypotheses, eliminating variables and assessing correlations. However, weaknesses include the fact that a researcher's categories may not reflect local understanding or context, may miss phenomena because of the focus on testing rather than generating new ideas or insights (confirmation bias), and may focus inappropriately on measurable variables (for which data is available) rather than underlying causal mechanisms [379,380]. In contrast, qualitative methods enable data to be based on a participant's own categories of meaning, are useful for studying a limited number of cases in depth, can be effective in describing complex phenomena or cases, and can better reveal how social actors "construct" different viewpoints [381]. The drawbacks are that qualitative knowledge may not be generalizable to other people or settings, may be of no help in making quantitative predictions, may take more time to collect, and may be more easily influenced by the researchers' own bias. Thus, there is much to be gained by mixing quantitative and qualitative methods, to avoid the weaknesses and to capitalize on the strengths of each.

In this way, our definition of rigor is about being "careful and thorough" in one's research, but not necessarily using the most advanced, sophisticated or complicated method. All methods have their strengths and limitations, so an effective definition of rigor is more of a "good balance across multiple criteria." In fact, overly complex research designs can be counterproductive, due to limited resources (lack of time, funding, access), lack of transparency in the process or results, or diminishing marginal returns for the added effort (e.g. doubling a sample size from 1000 to 2000 may have little impact on the size of confidence intervals around survey responses). In short, temper ambition and do not become paralyzed by seeking perfection.

5. Promoting style via structure, clarity, and critical thinking

We now turn to perhaps the most prosaic of our three dimensions of what makes good research: style. Although novelty and rigorous research designs are incredibly important, it can be equally important to effectively package and present your ideas to journal editors, peer reviewers, and eventual readers [382]. In that vein, we have three suggestions:

- Seek a coherent and cohesive macrostructure to an article, including elements such as titles, sub-headings, placement of paragraphs and regular signposting;
- Pursue clarity of expression in microstructure (the content of paragraphs, sentences, choice of words, tables and figures);
- Aim for transparency, think critically and examine and communicate the limitations of the analysis, especially insofar as you can explicitly preempt objections, and bring humility to your research.

These components of style amount to conveying information in a meaningful, accessible and well-reasoned manner. They remind researchers that producing research—asking research questions, designing a study, collecting data, analyzing data—is still very distinct from reporting that research on paper [383].

5.1. Seek cohesive and coherent macrostructure

This first element of style emphasizes the "big picture" of how a manuscript looks and reads. An effective writing structure boxes your analysis (sets its boundaries, limitations, what is included and excluded) and funnels information (so that it flows like a funnel from more general statements down to specific statements, then an expansion of those statements). To assist researchers in developing better macrostructure, we offer a few tips.

First, although the "standard" IMRAD [384] structure of "Introduction," "Materials/Methods," and "Results and Discussion" can work well for many manuscripts, authors can deviate from parts of it. For instance, both the "Literature Review" and "Results and "Discussion" of the paper can be organized in numerous creative ways [385–387]. For example:

- **A chronological** structure portrays events, or presents cases, as they happened over time, aiming to provide an overview or history of the relevant topic [388,389].
- A **conceptual** structure adheres to the units of analysis, components, or sub-components of a particular academic theory [390,391].
- A **cross-disciplinary** structure presents data according to the specific disciplines or domains of knowledge it comes from, e.g. linguistics, sociology, history, mathematics, or anthropology [392,393].
- A **hypothesis-testing** structure first introduces various hypotheses or suppositions and then organizes the results around testing, validating, or challenging them [394–397].
- A **spatial** or country structure organizes results by the countries or geographic case studies being examined [398,399].
- A **technological** structure organizes results by the specific systems (transport, electricity), technologies (solar, wind), or energy services (heat, mobility) being analyzed [400,401].
- A **thematic** structure organizes results around the themes emerging from the analysis, from different dimensions (technical, economic, political, social) to recurring topics (climate change mitigation, climate change adaptation) [402,403].
- A **narrative** structure organizes the data and results around a compelling storyline [404–406].
- A **hybrid** structure combines some of the structures above, such as: laying out a theory (conceptual structure) alongside country case studies (spatial structure) [407], by summarizing country case study results (spatial structure) by theme (thematic structure) [408], or by presenting propositions (hypothesis-testing structure) from within the disciplines they originate (cross-disciplinary structure) [409].

Indeed, a compelling case has been made for greater use of narrative structures (involving physical settings, events, characters and protagonists, stories and plots) as an effective form of communication given that human beings are dramatic creatures at heart [410,411]. That said, many students and novice writers may want to start with a more conventional structure. In any case, papers should aim to tell a good story, and the structure needs to be decided before writing commences—and in most cases will be adjusted as the writing proceeds. We also recommend beginning a paper by generating a high-level outline (perhaps as brief as one page or less, or longer if using the topic-sentence outline mentioned below), to help plan the structure and to assess how it all fits together.

Once a structure has been chosen and a condensed outline generated, we have a few other tips for structuring a manuscript [412,413].

- Authors should **carefully select their title, headings and sub-headings**, as these will help signpost an article. Titles are especially important, and should mention not only the topic but also (potentially) findings and case studies (if applicable).
- **Provide roadmaps** and textual bridges that connect the different sections of a manuscript; at times, summative tables and figures that preview or synthesize an article's findings or structure can be useful. By leafing or scrolling through an article, a reader should be able to spot the main findings easily, as well as figure out how the research was conducted, and locate any crucial definitions needed to understand its results.
- **Aim for similarity of length** between the comparable sections of a manuscript—for example, cases or sub-sections should be roughly the same size. At the same time, do not force this, as in some instances there can be a good reason to have different sizes.
- **Maintain paragraph cohesion** and a clear flow of logic: paragraphs need to be tied together in a smooth manner, otherwise it appears as if an author is simply throwing facts at the reader. Some find particular success with the use of a "topic sentence outline" that specifies each section title, and a single, topic sentence to represent each paragraph of the manuscript. Such an exercise helps to initially map out the article, and can be adjusted iteratively with the eventual manuscript throughout the drafting process. Such outlines can be particularly effective for planning and organizing expectations among a set of co-authors.

Recognizing there is a strong subjective element to "good" structural writing, we nevertheless recommend the list in Table 10 as a starting point. It contrasts a generically "good" paper with a "bad" paper across the constituent (and formulaic) components of a typical manuscript (as we have previously noted, not all articles need or even should utilize such a structure).

5.2. Pursue clarity of expression in microstructure

If an article's overall macrostructure is the foundation on which a manuscript is built, then the microstructure—sentences, words, diagrams, tables, figures, references—are its mortar and bricks. Although there is no universal approach to the mechanics of microstructure, most (if not all) well-written manuscripts maintain the following [415–417]:

- **Paragraph unity**, or "one idea per paragraph." Each paragraph should have one topic sentence. That is, a sentence that contains a subject, verb and object that define what the paragraph is all about (i.e., "The price of oil is increasing"). In most cases, the topic sentence is the first sentence but it can appear elsewhere. All other sentences are support sentences - intended to support the claim made in the topic sentence. So in this case, one would expect to see evidence that demonstrate the price is increasing. The paragraphs should not have any other information. So, if an author wants to explain why the price of oil is increasing, it should be either done in a separate paragraph with a new topic sentence (i.e. "There are three reasons for such price increases") or the topic sentence for the original paragraph should be rewritten (i.e. "Three factors are causing increases in oil prices").
- **Paragraph parsimony**. Authors should keep most paragraphs to a reasonable length (e.g. typically not more than 4–7 sentences); avoid excessive support sentences or examples, and let a paragraph rest when the point has been made.
- **Subject or verb/object congruence**. Authors should ensure analysis or examples are coherent. For example, if one writes that "the price of oil is booming," this is incongruent as prices cannot boom,

Table 10
Key structural ingredients of good and bad papers.
Source: Authors, based partially on [414].

	Good papers	Bad papers
Title	Describes topic but also key findings, themes, and contributions, and/or cases	Describes only the topic or method
	Identifies the geographic location of the research (if relevant)	Does not mention location or case study (if relevant)
Abstract	Clearly states research objectives or questions, methods, findings, limitations, and future directions	Focuses only on one or two aspects of the manuscript
	Is closely copy edited, is not repeated later in the text	Is full of typos, or repeated in the text itself verbatim
Introduction	Is short and sharp, often with an attention getting device at the start	Has a messy introduction that is too long
	Presents the core argument or question within the first few paragraphs	Presents the core argument too late
	Is well linked with the rest of the paper	Is poorly-linked with the rest of the paper
	Is well linked with the conclusion and findings	Ignores the link between the introduction and conclusion
	Previews the structure of the paper to come	Does not give the structure of the argument
Research Questions, Frameworks, Methods and Designs	Has a clear, answerable, interesting research question or questions	Has an unclear research question or none at all
	If appropriate, engages with a conceptual framework or frameworks	Does not state an appropriate theoretical or conceptual framework
	Is explicit about research design	Does not clarify research design
	Follows or acknowledges codes of practice for its research design	Does not consider codes of practice
	Mentions and pre-empts methodological limitations	Ignores or hides methodological limitations
Results	Actively interprets data	Lets data speak for itself
	Is selective and judicious about data utilized	Presents data not directly linked to the core argument
	Tightly couples data and analysis	Decouples the presentation of data from the analysis
Discussion/ Conclusion	Aims to make the conclusion the best part of the article	Has a thin conclusion
	Does not start a new argument in the conclusion	Starts a new argument in the conclusion
	Does not present new data in the conclusion	Presents new data in the conclusion
	Uses the conclusion to discuss findings as well as future research directions	Lets the conclusion be a summary and nothing else
	Cautiously discusses limitations and generalizability of findings (or lack thereof)	Ignores limitations and/or inappropriately presents findings as fully universal or generalizable
General structure	Tells a compelling story for the reader	Lets the reader wonder what the results mean
	Has coherent, logical structure with clear headings and subheadings	Has jumbled structure and no headings or subheadings
	Strong paragraph unity	Lacks paragraph unity
	Is well signposted	Forgets signposts

however often reported as such in the media. Idioms and colloquialisms work only when compatible.

- **Comprehensive referencing**. Authors should properly reference every factual claim, statistic, direct quote, or study/finding that influenced your argument. Always err on the side of referencing (to avoid unintentional plagiarism), and always go to the original source. Further, authors should strive to put others' work into their own words—and be sure to use quotation marks in those rare instances where it is appropriate to use the source's original words.
- **Appropriate length**. As a general rule, authors should aim for

brevity (present article excluded!). If a researcher can say it in fewer words, or with fewer examples, do so. As the saying goes, "I would have written you a shorter letter but ran out of time." Conveying information via a condensed number of words is often more difficult than lengthy exposition—yet the condensed version can be much more readable and useful to a target audience.

- **Minimal jargon and acronyms**. Arguably, any piece of writing should seek to be accessible to a wide audience, and this is especially true for the interdisciplinary and applied work in our field. Authors should thus take the time to identify and carefully define

any pieces of "jargon" used in the paper (terminology that is unique to a particular field, discipline or sub-discipline), and to minimize the use of such jargon where possible. Similarly, acronyms should be used sparingly, and when used should be carefully spelled out when first introduced (and potentially reintroduced in later sections), or summarized in a list of abbreviations at the beginning of the manuscript.

Admittedly, the above tips are mostly about the mechanics of writing. What about the stylistic elements—adding vim, vigor, flair, and character to your writing so the words sparkle and the manuscript keeps readers riveted? Here, although it is even more difficult to distil lessons, we advocate a few. Aristotle believed that effective communication rested not only on logic (logos) but also emotional connection (pathos) and credibility (ethos)—good manuscripts often possess all three. Writing more than a half century ago, George Orwell [418] critiqued writing for being prone to dying metaphors that have worn out and lost all power (e.g., "two sides of the same coin," "chicken or the egg," "a tale of two cities," "Achilles heel"); for using phrases instead of verbs (such as "render inoperative" instead of "break," or "mitigate against" instead of "stop"); and for dressing up simple statements with big or foreign words. To counter these trends, Orwell offered six general rules that we find helpful:

- Never use a metaphor, simile, or other figure of speech which you are used to seeing;
- Never use a long word where a short one will do;
- If it is possible to cut a word out, always cut it out;
- Never use the passive voice where you can use the active;
- Never use a foreign phrase, a scientific word, or a jargon word if you can think of an everyday English equivalent;
- Break any of these rules sooner than say anything outright barbarous.

And, because Orwell was talking about writing in general (from fiction and poetry to non-fiction), we have a few more tips tailored especially for academic articles:

- Effectively utilize visual aids (figures and tables, photographs, maps, infographics and other "visualization tools" [419]) to enhance the impact of your writing;
- Use rhetorical devices (examples, analogies, anecdotes, epitaphs, poems, even jokes) to (selectively) enhance the appeal of your writing;
- Have fun, be creative, and don't be afraid to experiment [420]. Writing is too important a part of the academic career to not enjoy at least part of it.

5.3. Aim for transparency, test and critically examine your analysis

Our last suggestion is to be transparent about assumptions, to think critically and to actively acknowledge and explain limitations. Although such an exercise could fall partly under rigor (a part of being careful is considering contrary viewpoints), we have put it in style because it is an important stylistic technique that we wish every manuscript employed.

One way of systematically being critical is to always consider the five "tests" for a manuscript [421]. Do the assumptions of a model or a theory fit? Do the conclusions follow from the premises? Do the implications of the argument find confirmation in the data? How much better is the argument than other, competing explanations? How useful is the explanation for understanding or explaining other cases? Considering these tests may mean explicitly adding text to your manuscript that acknowledges the key limitations in method, theory, generalizability of findings and so on.

Furthermore, part of aiming for transparency, reflection, and humility is to appreciate the necessity of the process of revising and

editing. Experienced writers commonly report that only 20% of their writing time is on the first draft, with the remaining 80% on revisions, edits and re-writes. Kazuo Ishiguro, who won the 2017 Nobel Prize in Literature, remarks that good writing requires "a willingness to be terrible" the first time around, before people see it [422]. Feedback from others—colleagues, peers, editors, even expected critics—is always good before submission. Actively seek comments and criticism on a manuscript (it's best to know potential weaknesses as early as possible), since these are far more helpful than praise.

6. Conclusion

To conclude, we've thrown a capacious amount of recommendations at readers. As such, it is difficult (and admittedly contestable) to offer any type of definitive guidance or checklist for how to design, implement and write more novel, rigorous, and stylistic studies. After all, in many ways research itself is a "method of discovery" [423] or a "craft of inquiry" [424] with no predetermined answers or fully agreed upon processes. Albert Einstein is reputed to have said that "if we knew what we were looking for, it wouldn't be called 're-search'." In particular, the codes of practice and hierarchies of evidence that we identify reveal a diversity of research designs and very different approaches, goals, and aims.

All too often, when one moves away from the limits of a single disciplinary idea of novelty, rigor, or style, then the guidelines disappear, so we end up with an abundance of low quality work, and in some cases a lack of appreciation for high quality work. Thus, given the clear importance of interdisciplinarity in energy social science, we argue that guidelines are strongly needed. This is not to say that a rigorous researcher needs to be completely interdisciplinary, fully trained in all relevant research methods—but at a minimum they need to have a basic awareness and appreciation of alternative paradigms, viewpoints, and methods. Such appreciation will inject an appropriate level of humility into their work and will improve their ability to conduct and comprehend literature reviews, identify research gaps and effectively build collaborative, interdisciplinary research teams.

In this admittedly lengthy but hopefully holistic review, we have sought to establish a comprehensive and clear set of guidelines for the interdisciplinary field of energy social science. These are not dogmatic, but instead highlight general principles that are often missing or implied. We therefore posit that stronger research tends to:

- **Clearly state objectives**. Good papers explicitly ask a research question (or questions) and/or set out to achieve particular aims and objectives.
- **Be empirically grounded in evidence**. Good research is data-driven, based on a foundation of empirical data rather than opinion (or worse, bias).
- **Have and communicate a research design**. Good papers are as explicit as possible about the research design and methods employed, cognizant of codes of practice, and appropriate and balanced in their execution.
- **Appreciate multiple methods**. Rigorous researchers will explain how their method compares to alternative methods and approaches. Even better, novel and rigorous research designs can combine at least two complementary methods.
- **Theorize**. Many good papers connect themselves to social science concepts or theories. They test concepts, engage in debates, and elaborate on conceptual findings about the relationship between energy and society.
- **Address generalizability**. Comparative research (e.g. across technologies, policies, regions) can have broader impact. Research in one region, such as a survey conducted in one country, or a single case study, needs to make a strong argument for how the results contribute to theoretical development or are applicable beyond that case.

- **Be stylistically strong**. Good papers utilize a coherent macro-structure and microstructure, and are written in a way that is crisp, clear and (at times) creative and fun.
- **Emphasize strengths and weaknesses**. Rigorous researchers fully acknowledge, explain, and (when possible) preempt limitations in design, case study selection, methods or analysis.

These principles suggest that energy social science research is enhanced by the principles of diversity (intellectual, theoretical, methodological, empirical), inclusion (professional, geographic, disciplinary), creativity (experimentation, curiosity, ambition) and reflection (appreciation or even omniscience of other work, transparency, critical thinking, and modesty). Such research is clearly conveyed so assumptions are apparent as well as strengths and weaknesses. It may require teams of researchers and years of hard work to make a significant contribution, thus requiring both persistence and patience.

There is value to smaller-scale, incremental contributions, where the guidelines we provide above apply just as well. Each new published insight can contribute to the broader body of knowledge, in particular through eventual literature reviews on the subject. Similarly, in more positivist, quantitative disciplines, individual experiments and statistical analyses are the building blocks for a later systematic review or meta-analysis.

That said, as much as we want to offer tips and guidance, we must also remember that energy social science is both a science and an art [425]. It must be not only logical but emotionally impactful and credible. It is not only dialectic but rhetoric. It is not only analysis but argument – the effective presentation of ideas to an audience. While energy social science remains a collective endeavor, outstanding research shines when it excels across the three dimensions of novelty, rigor, and style.

Acknowledgments

References

[1] J. Law, After Method: Mess in Social Science Research, Routledge, London, 2004.
[2] International Renewable Energy Agency & International Energy Agency, Perspectives for the Energy Transition: Investment Needs for a Low-carbon Energy System, available at: (2017) http://www.irena.org/menu/index.aspx?mnu = Subcat&PriMenuID = 36&CatID = 141&SubcatID = 3828.
[3] C. Figueres, et al., Three years to safeguard our climate, Nature 546 (2017) 593–595.
[4] F.W. Geels, B.K. Sovacool, T. Schwanen, S. Sorrell, Sociotechnical transitions for deep decarbonisation, Science 357 (6357) (2017) 1242–1244 September 22.
[5] C. Felix, et al., Towards demand-side solutions for mitigating climate change, Nat. Clim. Chang. 8 (April) (2018) 260–271.
[6] J. Hodbod, W. Neil Adger, Integrating social-ecological dynamics and resilience into energy systems research, Energy Res. Soc. Sci. 1 (March) (2014) 226–231.
[7] A. Bridges, F.A. Felder, K. McKelvey, I. Niyogi, Uncertainty in energy planning: estimating the health impacts of air pollution from fossil fuel electricity generation, Energy Res. Soc. Sci. 6 (March) (2015) 74–77.
[8] R. Walker, C. Liddell, P. McKenzie, C. Morris, S. Lagdon, Fuel poverty in Northern Ireland: humanizing the plight of vulnerable households, Energy Res. Soc. Sci. 4 (December) (2014) 89–99.
[9] L. Middlemiss, R. Gillard, Fuel poverty from the bottom-up: characterising household energy vulnerability through the lived experience of the fuel poor, Energy Res. Soc. Sci. 6 (March) (2015) 146–154.
[10] C. Snell, M. Bevan, H. Thomson, Justice, fuel poverty and disabled people in England, Energy Res. Soc. Sci. 10 (November) (2015) 123.
[11] R. Gillard, C. Snell, M. Bevan, Advancing an energy justice perspective of fuel poverty: household vulnerability and domestic retrofit policy in the United Kingdom, Energy Res. Soc. Sci. 29 (July) (2017) 53–61.
[12] C. Emily, Opening the black box of energy security: a study of conceptions of electricity security in the United Kingdom, Energy Res. Soc. Sci. 21 (November) (2016) 1–11.
[13] M.A. Brown, Y. Wang, B.K. Sovacool, A.L. D'Agostino, Forty years of energy security trends: a comparative assessment of 22 industrialized countries, Energy Res. Soc. Sci. 4 (December) (2014) 64–77.
[14] C.R. Jones, D. Kaklamanou, L. Lazuras, Public perceptions of energy security in Greece and Turkey: exploring the relevance of pro-environmental and pro-cultural orientations, Energy Res. Soc. Sci. 28 (June) (2017) 17–28.
[15] K. Jenkins, D. McCauley, R. Heffron, H. Stephan, R. Rehner, Energy justice: a conceptual review, Energy Res. Soc. Sci. 11 (January) (2016) 174–182.
[16] N. Simcock, C. Mullen, Energy demand for everyday mobility and domestic life: exploring the justice implications, Energy Res. Soc. Sci. 18 (August) (2016) 1–6.
[17] K. Bickerstaff, G. Walker, H. Bulkeley (Eds.), Energy Justice in a Changing Climate: Social Equity and Low-Carbon Energy, Zed Books, New York, 2013.
[18] G. Bosch, Train PhD students to be thinkers not just specialists, Nature 554 (2018) 277.
[19] G. Schuitema, N.D. Sintov, Should we quit our jobs? Challenges, barriers and recommendations for interdisciplinary energy research, Energy Policy 101 (2017) 246–250.
[20] B. Turnheim, F. Berkout, F. Geels, A. Hof, A. McMeekin, B. Nykvist, D. van Vuuren, Evaluating sustainability transitions pathways: bridging analytical approaches to address governance challenges, Glob. Environ. Change 35 (2015) 239–253.
[21] N. Castree, Geography and the new social contract for global change research, Trans. Inst. Br. Geogr. 41 (3) (2016) 328–347.
[22] N. Castree, W.M. Adams, J. Barry, et al., Changing the intellectual climate, Nat. Clim. Change 4 (2014) 763–768.
[23] A.C.G. Cooper, Building physics into the social: enhancing the policy impact of energy studies and energy social science research, Energy Res. Soc. Sci. 26 (April) (2017) 80–86.
[24] U. Felt, Within, across and beyond: reconsidering the role of social sciences and humanities in Europe, Sci. Cult. 23 (3) (2014) 384–396.
[25] F. Gralla, D.J. Abson, D. Lang, A. Møller, et al., Nuclear accidents call for transdisciplinary nuclear energy research, Sustain. Sci. 10 (1) (2015) 179–183 January.
[26] F. Harris, F. Lyon, Transdisciplinary environmental research: building trust across professional cultures, Environ. Sci. Policy 31 (2013) 109–119.
[27] B. Mallaband, G. Wood, K. Buchanan, S. Staddon, N.M. Mogles, E. Gabe-Thomas, The reality of cross-disciplinary energy research in the United Kingdom: a social science perspective, Energy Res. Soc. Sci. 25 (March) (2017) 9–18.
[28] B.K. Sovacool, Energy studies need social science, Nature 511 (7511) (2014) 529–530 July 31.
[29] B.K. Sovacool, S.E. Ryan, P.C. Stern, K. Janda, G. Rochlin, D. Spreng, M.J. Pasqualetti, H. Wilhite, Lutzenhiser L integrating social science in energy research, Energy Res. Soc. Sci. 6 (2018) 95–99.
[30] D. Spreng, Transdisciplinary energy research – reflecting the context, Energy Res. Soc. Sci. 1 (March) (2014) 65–73.
[31] M. Winskel, Embedding social sciences in interdisciplinary research: recent experiences from interdisciplinary energy research, Sci. Cult. 23 (3) (2014) 413–418.
[32] G. Marsden, L. Reardon, Questions of governance: rethinking the study of transportation policy, Transportation Research Part A: Policy and Practice 101 (2017), pp. 238–251.
[33] H.R. Hughes, M. Paterson, Narrowing the climate field: the symbolic power of authors in the IPCC's assessment of mitigation, Rev. Policy Res. 34 (2017) 744–766.
[34] E. Corbera, L. Calvet-Mir, H. Hughes, M. Paterson, Patterns of authorship in the IPCC working group III report, Nat. Clim. Chang. 6 (2016) 94–99.
[35] P. Brandt, et al., A review of transdisciplinary research in sustainability science, Ecol. Econ. 92 (2013) 1–15.
[36] S. Schmidt, H. Weigt, Interdisciplinary energy research and energy consumption: What, why, and how? Energy Res. Soc. Sci. 10 (2015) 206–219.
[37] C. Havergal, Stern aims for lower REF workload, more 'game-changing' research, Times High. Educ. 28 (July) (2016).
[38] M. Binswanger, Excellence by Nonsense: The Competition for Publications in

Modern Science. Opening Science, Springer, London, 2013, pp. 49–72.

[39] M. Marinetto, Fast food research in the era of unplanned obsolescence, J. Manage. Stud. (2018), https://doi.org/10.1111/joms.12338 Accepted Author Manuscript https://onlinelibrary.wiley.com/doi/abs/10.1111/joms.12338.

[40] C.A. Donnelly, et al., Four principles for synthesizing evidence, Nature 558 (June (21)) (2018) 361–364.

[41] B.K. Sovacool, What are we doing here? Analyzing fifteen years of energy scholarship and proposing a social science research agenda, Energy Res. Soc. Sci. 1 (March) (2014) 1–29.

[42] K. O.' Neil, et al., Methods and global environmental governance, Annu. Rev. Environ. Resour. 38 (2013) 441–471.

[43] I.G. Hamilton, A.J. Summerfield, R. Lowe, P. Ruyssevelt, C.A. Elwell, T. Oreszczyn, Energy epidemiology: a new approach to end-use energy demand research, Build. Res. Inf. 41 (4) (2013) 482–497.

[44] M.F. Bellemare, How to Publish in Academic Journals. AAEA Early Career Professionals Workshop. Vail, CO, June 2 (2017).

[45] B. Hancke, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[46] P.C. Stern, B.K. Sovacool, T. Dietz, Towards a science of climate and energy choices, Nat. Clim. Change 6 (June) (2016) 547–555.

[47] R. Mourik, Free Our Minds!, June 13 (2017) http://www.eceee.org/all-news/columns/free-our-minds/.

[48] K.O.' Neil, et al., Methods and global environmental governance, Annu. Rev. Environ. Resour. 38 (2013) 441–471.

[49] P.C. Stern, B.K. Sovacool, T. Dietz, Towards a science of climate and energy choices, Nat. Clim. Change 6 (June) (2016) 547–555.

[50] T.S. Kuhn, The Structure of Scientific Revolutions, University of Chicago, 1962.

[51] N. Edomah, C. Foulds, A. Jones, Influences on energy supply infrastructure: a comparison of different theoretical perspectives, Renew. Sustain. Energy Rev. 79 (2017) 765–778.

[52] T. Jackson, Motivating Sustainable Consumption: A Review of Evidence on Consumer Behaviour and Behavioural Change, Sustainable Development Research Network, University of Surrey, Guildford, Surrey, 2005.

[53] F. Kern, K.S. Rogge, Harnessing theories of the policy process for analysing the politics of sustainability transitions: a critical survey, Environ. Innovation Societal Transitions 27 (2018) 102–117.

[54] K. Peattie, Green consumption: behavior and norms, Annu. Rev. Environ. Resour. 35 (2010) 195–228.

[55] M. Scheller, J. Urry, Mobilizing the new mobilities paradigm, Appl. Mobilities 1 (1) (2016) 10–25.

[56] B.K. Sovacool, D.J. Hess, Ordering theories: Typologies and conceptual frameworks for sociotechnical change, Social Stud. Sci. 47 (5) (2018) 703–750.

[57] C. Wilson, H. Dowlatabadi, Models of decision making and residential energy use, Annu. Rev. Environ. Resour. 32 (1) (2007) 169–203.

[58] J. Mingers, Realising Systems Thinking: Knowledge and Action in Management Science, Springer Science & Business Media, 2006.

[59] K.S. Popper, The Open Society and Its Enemies, Routledge, 2012.

[60] F.W. Geels, Foundational ontologies and multi-paradigm analysis, applied to the socio-technical transition from mixed farming to intensive pig husbandry (1930–1980), Technol. Anal. Strateg. Manag. 21 (7) (2009) 805–832.

[61] J. Rutherford, O. Coutard, Urban energy transitions: places, processes and politics of socio-technical change, Urban Stud. 51 (7) (2014) 1353–1377.

[63] F.W. Geels, F. Berkhout, D.P. van Vuuren, Bridging analytical approaches for low-carbon transitions, Nat. Clim. Change 6 (2016) 576–583 2016.

[64] T. Jackson, Motivating Sustainable Consumption: A Review of Evidence on Consumer Behaviour and Behavioural Change, Sustainable Development Research Network, University of Surrey, Guildford, Surrey, UK, 2005.

[65] D. Collier, J.E. Mahon, Conceptual "Stretching" revisited: adapting categories in comparative analysis, Am. Polit. Sci. Rev. 87 (December (4)) (1993) 845–855.

[66] T. Jackson, Motivating Sustainable Consumption: a Review of Evidence on Consumer Behaviour and Behavioural Change. A Report to the Sustainable Development Research Network January, (2005).

[67] G. Sartori, Concept misformation in comparative politics, Am. Polit. Sci. Rev. 64 (4) (1970) 1033–1053 Dec..

[68] I. Ajzen, Values, attitudes, and behavior, in: S. Salzborn, E. Davidov, J. Reinecke (Eds.), Methods, Theories, and Empirical Applications in the Social Sciences: Festschrift for Peter Schmidt, Springer VS, Berlin, 2012, pp. 33–38.

[69] E. Babbie, The Practice of Social Research, 12th edition, Wadsworth, 2010.

[70] E. Babbie, The Practice of Social Research, 12th edition, Wadsworth, 2010.

[71] E. Guba, Y. Lincoln, Competing paradigms in qualitative research, in: N. Denzin, Y. Lincoln (Eds.), Handbook of Qualitative Research, Sage, Thousand Oaks, 1994.

[72] G. McCracken, The Long Interview, Sage Publications, Newbury Park, California, 1988.

[73] D. Tranfield, et al., Towards a methodology for developing evidence-informed management knowledge by means of systematic review, Br. J. Manag. 14 (2003) 207–222.

[74] S. Sorrell, Improving the evidence base for energy policy: the role of systematic reviews, Energy Policy 35 (2007) 1858–1871.

[75] S. Sorrell, Improving the evidence base for energy policy: the role of systematic reviews, Energy Policy 35 (2007) 1858–1871.

[76] H. Bloom, The core analytics of randomized experiments for social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[77] D.T. Campbell, J.C. Stanley, N.L. Gage, Experimental and Quasi-experimental Designs for Research (No. 04; Q175, C3.), Houghton Mifflin, Boston, 1963.

[78] T. Cook, V. Wong, Better quasi-experimental practice, in: P. Alasuutari,

[79] L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

Frans Sengers, Anna J. Wieczorek, Rob Raven, Experimenting for sustainability transitions: A systematic literature review Technological Forecasting & Social Change (in press 2017) https://www.sciencedirect.com/science/article/pii/S0040162516302530.

[80] T. Greenhalgh, et al., Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review, Soc. Sci. Med. 61 (2005) 417–430.

[81] R. Whittemore, K. Knafl, The integrative review: updated methodology, J. Adv. Nurs. 52 (5) (2005) 546–553.

[83] H. Collins, T. Pinch, The Golem at Large: What You Should Know About Technology, Cambridge University Press, Cambridge, 1998.

[84] M. Jefferson, Closing the gap between energy research and modelling, the social sciences, and modern realities, Energy Res. Soc. Sci. 4 (December) (2014) 42–52.

[85] M. Jefferson, Energy realities or modelling: which is more useful in a world of internal contradictions? Energy Res. Soc. Sci. 22 (December) (2016) 1–6.

[86] E. O'Sullivan, G.R. Rassel, M. Berner, Research Methods for Public Administrators, Pearson Longman, New York, 2010.

[87] G. King, R.O. Keohane, S. Verba, Designing Social Inquiry: Scientific Inference in Qualitative Research, Princeton University Press, Princeton, NJ, 1994.

[88] B. Hancké, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[89] A.L. George, A. Bennett, Case Studies and Theory Development in the Social Sciences, Harvard University Press, Cambridge, MA, 2004.

[91] B. Flyvbjerg, Five misunderstandings about case study research, Qual. Inq. 12 (April (2)) (2006) 219–245 2006.

[92] B. Flyvbjerg, Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again, Cambridge University Press, Cambridge, MA, 2001 2001.

[93] S. Jacobsson, A. Johnson, The diffusion of renewable energy technology: an analytical framework and key issues for research, Energy Policy 28 (2000) 625–640.

[94] M.P. Hekkert, R.A.A. Suurs, S.O. Negro, S. Kuhlmann, R.E.H.M. Smits, Functions of innovation systems: a new approach for analysing technological change, Technol. Forecast. Soc. Change 74 (4) (2007) 413–432.

[95] A. Bergek, S. Jacobsson, B. Carlsson, S. Lindmark, A. Rickne, Analyzing the functional dynamics of technological innovation systems: a scheme of analysis, Res. Policy 37 (3) (2018) 407–429.

[96] I. Ropke, Theories of practice: new inspiration for ecological economic studies of consumption, Ecol. Econ. 68 (10) (2009) 2490–2497.

[97] E. Shove, M. Pantzar, M. Watson, The Dynamics of Social Practice: Everyday Life and How It Changes, Sage, London, 2012.

[98] V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, User acceptance of information technology: toward a unified view, MIS Q. 27 (3) (2003) 425–478.

[99] F.W. Geels, Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study, Res. Policy 31 (8) (2002) 1257–1274.

[100] F.W. Geels, From sectoral systems of innovation to socio-technical systems: insights about dynamics and change from sociology and institutional theory, Res. Policy 33 (6–7) (2004) 897–920.

[101] F.W. Geels, J.W. Schot, Typology of sociotechnical transition pathways, Res. Policy 36 (3) (2007) 399–417.

[102] J. Axsen, K.S. Kurani, Social influence, consumer behavior, and low-carbon energy transtioins, Annu. Rev. Environ. Resour. 37 (2012) 311–340.

[103] J. Axsen, K.S. Kurani, Social influence and proenvironmental behavior: the reflexive layers of influence framework, Environ. Plann. B Plann. Des. 41 (5) (2014) 847–862.

[104] R. Galvin, M. Sunikka-Blank, Schatzkian practice theory and energy consumption research: time for some philosophical spring cleaning? Energy Res. Soc. Sci. 22 (December (1)) (2016) 63–68.

[105] F.F. Sniehotta, J. Presseau, V. Araújo-Soares, Time to retire the theory of planned behaviour, Health Psychol. Rev. 8 (November (1)) (2013) 1–7 1.

[106] U. Flick, Triangulation, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018, p. 444.

[107] G.T. Allison, Conceptual models and the cuban missile crisis, Am. Polit. Sci. Rev. 63 (September (3)) (1969) 689–718 1969.

[108] K.S. Wolske, P.C. Stern, T. Dietz, Explaining interest in adopting residential solar photovoltaic systems in the United States: toward an integration of behavioral theories, Energy Res. Soc. Sci. 25 (March) (2017) 134–151.

[109] M. Ryghaug, M. Toftaker, A transformative practice? Meaning, competence, and material aspects of driving electric cars in Norway, Nat. Cult. 9 (2) (2014) 146–163 2004.

[110] B.K. Sovacool, L.D. Noel, R.J. Orsato, Stretching, embeddedness, and scripts in a sociotechnical transition: explaining the failure of electric mobility at better place (2007-2013), Technol. Forecast. Soc. Change 123 (October) (2017) 24–34.

[112] D. Youngblood, Multidisciplinarity, interdisciplinarity, and bridging disciplines: a matter of process, J. Res. Pract. 3 (2) (2007) M18.

[113] J.W. Creswell, Controversies in mixed methods research, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research Methods, 4th edition, SAGE Publications, London, UK, 2011.

[114] C. Teddlie, A. Tashakkori, Mixed methods research: contemporary issues in an emerging Field, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, 4th edition, SAGE Publications, London, UK, 2011.

[115] N. Denzin, Y. Lincoln, Strategies of inquiry, in: N. Denzin, Y. Lincoln (Eds.), The Sage Handbook of Qualitative Research, 5th edition, SAGE Publications, London, UK, 2018.

[116] A. Bryman, Integrating qualitative and quantitative research: how is it done? Qual. Res. 6 (2006) 97–113.

[117] J.W. Creswell, Research Design: Qualitative, Quantitative and Mixed Methods Approaches, 2nd edition, Sage Publications, London, UK, 2003.

[118] N. Denzin, Sociological Methods: A Sourcebook, 5th edition, Aldine Transaction, 2006 ISBN 978-0-202-30840-1.

[119] U. Flick, Triangulation, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018.

[120] U. Flick, Triangulation, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018.

[121] U. Flick, Triangulation, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018.

[122] R. Murphy, M. Jaccard, Energy efficiency and the cost of GHG abatement: a comparison of bottom-up and hybrid models for the US, Energy Policy 39 (11) (2011) 7146–7155.

[123] B. Al-Alawi, T. Bradley, Review of hybrid, plug-in hybrid, and electric vehicle market modeling studies, Renew. Sustain. Energy Rev. 21 (2013) 190–203.

[124] D.L. McCollum, C. Wilson, H. Pettifor, K. Ramea, V. Krey, K. Riahi, C. Bertram, Z. Lin, O.Y. Edelenbosch, S. Fujisawa, Improving the behavioral realism of global integrated assessment models: an application to consumers' vehicle choices, Transp. Res. D Transp. Environ. 55 (2017) 322–342.

[125] H. Pettifor, C. Wilson, et al., Modelling social influence and cultural variation in global low-carbon vehicle transitions, Glob. Environ. Chang. Part A 47 (2017) 76–87.

[126] D.L. McCollum, C. Wilson, et al., Improving the behavioral realism of global integrated assessment models: an application to consumers' vehicle choices, Transp. Res. D Transp. Environ. 55 (Suppl. C) (2017) 322–342.

[127] K. Safarzyńska, Integrating behavioural economics into climate-economy models: some policy lessons, Clim. Policy (2017) 1–14.

[128] T. Gnann, P. Plötz, et al., Modelling market diffusion of electric vehicles with real world driving data – German market and policy options, Transp. Res. Part A Policy Pract. 77 (2015) 95–112.

[129] M. Wolinetz, J. Axsen, How policy can build the plug-in electric vehicle market: insights from the respondent-based preference and constraints (REPAC) model, Technol. Forecast. Soc. Change 117 (2017) 238–250.

[130] M.A. Janssen, E. Ostrom, Empirically based, agent-based models, Ecol. Soc. 11 (2) (2006) 643–655.

[131] N. Beck, J. Katz, What to do (and not to do) with time-series cross-section data, Am. Polit. Sci. Rev. 89 (3) (1995) 634–647.

[132] J. Elliot, J. Holland, et al., Longitudinal and panel studies, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[133] C. Hsiao, Panel data analysis—advantages and challenges, Test 16 (1) (2007) 1–22.

[134] A.F. Jensen, E. Cherchi, et al., On the stability of preferences and attitudes before and after experiencing an electric vehicle, Transp. Res. D Transp. Environ. 25 (Suppl. C) (2013) 24–32.

[135] E. Graham-Rowe, B. Gardner, et al., Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: a qualitative analysis of responses and evaluations, Transp. Res. Part A Policy Pract. 46 (1) (2012) 140–153.

[136] S.M. Skippon, N. Kinnear, et al., How experience of use influences mass-market drivers' willingness to consider a battery electric vehicle: a randomised controlled trial, Transp. Res. Part A Policy Pract. 92 (Suppl. C) (2016) 26–42.

[137] D. Dillman, J. Smyth, et al., Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons Inc, New Jersey, USA, 2014.

[138] D. Dillman, J. Smyth, et al., Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons Inc, New Jersey, USA, 2014.

[139] J. Elliot, J. Holland, et al., Longitudinal and panel studies, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[140] L. Whitmarsh, S. O'Neill, Green identity, green living? The role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours, J. Environ. Psychol. 30 (3) (2010) 305–314.

[141] F. Xie, Z. Lin, Market-driven automotive industry compliance with fuel economy and greenhouse gas standards: analysis based on consumer choice, Energy Policy 108 (2017) 299–311.

[142] S. Sorrell, The economics of energy service contracts, Energy Policy 35 (1) (2007) 507–521.

[143] E. Figenbaum, Perspectives on Norway's supercharged electric vehicle policy, Environ. Innov. Soc. Transit. 25 (2017) 14–34.

[144] L. Dexter, Antony, Elite and Specialized Interviewing, Northwestern University Press, Evanston, IL, 1970.

[145] M. Woods, Rethinking elites: networks, space, and local politics, Environ. Plan. A 30 (12) (1998) 2101–2119.

[146] M.E. Drumwright, P.E. Murphy, How advertising practitioners view ethics, J. Advert. 33 (2) (2004) 7–24.

[147] Rowe and Wright, The Delphi technique as a forecasting tool: issues and analysis, Int. J. Forecast. 15 (October (4)) (1999) 1999.

[148] J. Axsen, S. Goldberg, J. Bailey, How might potential future plug-in electric vehicle buyers differ from current "Pioneer" owners? Transp. Res. D Transp. Environ. 47 (2016) 357–370.

[149] J. Axsen, K.S. Kurani, Developing sustainability-oriented values: insights from households in a trial of plug-in hybrid electric vehicles, Glob. Environ. Chang. Part A 23 (February (1)) (2013) 70–80.

[150] M.J. Bürer, R. Wüstenhagen, Which renewable energy policy is a venture capitalist's best friend? Empirical evidence from a survey of international cleantech investors, Energy Policy 37 (12) (2009) 4997-4500.

[151] L.I. Remennick, Immigrants from Chernobyl-affected areas in Israel: the link between health and social adjustment, Soc. Sci. Med. 54 (January (2)) (2002) 309–317.

[152] M. Tone, T. Stone, What we can learn about recovery: lessons from the Fukushima survivors, Nursing and Helath Sciences 16 (March (1)) (2014) 52–55.

[153] L. Necefer, G. Wong-Parodi, P. Jaramillo, M.J. Small, Energy development and Native Americans: values and beliefs about energy from the Navajo Nation, Energy Res. Soc. Sci. 7 (May) (2015) 1–11.

[154] G. Graetz, Energy for whom? Uranium mining, Indigenous people, and navigating risk and rights in Australia, Energy Res. Soc. Sci. 8 (July) (2015) 113–126 ISSN 2214-6296.

[155] D.B. Carson, D.A. Carson, G. Nordin, P. Sköld, Lessons from the Arctic past: the resource cycle, hydro energy development, and the human geography of Jokkmokk, Sweden, Energy Res. Soc. Sci. 16 (June) (2016) 13–24.

[156] A. Bielschowsky, Kids and Kilowatts: Socialisation, Energy Efficiency, and Electricity Consumption in New Zealand, Energy Res. Soc. Sci. 44 (2018) 178–186.

[157] E.M. Mistur, Health and energy preferences: rethinking the social acceptance of energy systems in the United States, Energy Res. Soc. Sci. 34 (December) (2017) 184–190.

[158] C. Liddell, C. Morris, B. Gray, A. Czerwinska, B. Thomas, Excess winter mortality associated with Alzheimer's disease and related dementias in the UK: a case for energy justice, Energy Res. Soc. Sci. 11 (January) (2016) 256–262.

[159] B. Southwell, K. Ronneberg, K. Shen, E. Jorgens, J. Hazel, R. Alemu, J. Ross, L. Richman, D. Vermeer, Energy information engagement among the poor: predicting participation in a free workshop, Energy Res. Soc. Sci. 4 (December) (2014) 21–22.

[160] M. Bazilian, S. Nakhooda, Thijs Van De Graaf, energy governance and poverty, Energy Res. Soc. Sci. 1 (March) (2014) 217–222.

[161] D. Hernández, D. Phillips, Benefit or burden? Perceptions of energy efficiency efforts among low-income housing residents in New York City, Energy Res. Soc. Sci. 8 (July) (2015) 52–59.

[162] Oxford Dictionary (2017). https://en.oxforddictionaries.com/definition/big_data.

[163] L. Noel, The hidden economic benefits of large-scale renewable energy deployment: integrating heat, electricity and vehicle systems, Energy Res. Soc. Sci. 26 (2017) 54–59.

[164] L. Noel, J.F. Brodie, W. Kempton, C.L. Archer, C. Budischak, Cost minimization of generation, storage, and new loads, comparing costs with and without externalities, Appl. Energy 189 (March) (2017) 110–121.

[165] B. Anderson, Laundry, energy and time: insights from 20 years of time-use diary data in the United Kingdom, Energy Res. Soc. Sci. 22 (December) (2016) 125–136.

[166] S. D'Oca, S.P. Corgnati, T. Buso, Smart meters and energy savings in Italy: determining the effectiveness of persuasive communication in dwellings, Energy Res. Soc. Sci. 3 (September) (2014) 131–142.

[167] J.M. Burgett, Fixing the American energy leak: the effectiveness of a whole-house switch for reducing standby power loss in U.S. residences, Energy Res. Soc. Sci. 6 (March) (2015) 87–94.

[168] D.A. Kroodsma, et al., Tracking the global footprint of fisheries, Science 359 (6378) (2018) 904–908 23 Feb.

[169] L. Sikanen, et al., Transport control of forest fuels by fleet manager, mobile terminals and GPS, Biomass Bioenergy 28 (February (2)) (2005) 183–191.

[170] K. Koscher, et al., Experimental security analysis of a modern automobile, 2010 IEEE Symposium on Security and Privacy (2010) 447–462, https://doi.org/10.1109/SP.2010.34.

[171] S.C. Isley, et al., Online purchasing creates opportunities to lower the life cycle carbon footprints of consumer products, PNAS 113 (35) (2016) 9780–9785.

[172] M.J. Culnan, P.J. McHugh, J.I. Zubillaga, How large US companies can use Twitter and other social media to gain business value, MIS Q. Executive 9 (4) (2010).

[173] T.J. Chatterton, J. Anable, J. Barnes, G. Yeboah, Mapping household direct energy consumption in the United Kingdom to provide a new perspective on energy justice, Energy Res. Soc. Sci. 18 (2016) 71–87.

[174] I.G. Hamilton, A.J. Summerfield, R. Lowe, P. Ruyssevelt, C.A. Elwell, T. Oreszczyn, Energy epidemiology: a new approach to end-use energy demand research, Build. Res. Inf. 41 (4) (2013) 482–497.

[175] I. Hamilton, A. Summerfield, T. Oreszczyn, P. Ruyssevelt, Using epidemiological methods in energy and buildings research to achieve carbon emission targets, Energy Build. 154 (2017) 188–197.

[176] S. Pfenninger, et al., The importance of open data and software: is energy research lagging behind? Energy Policy 101 (February) (2017) 211–215.

[177] N.M. Taleb, Beware the big errors of big data, Wired (February (2)) (2013) Available at https://www.wired.com/2013/02/big-data-means-big-errors-people/.

[178] S. Kaisler, F. Armour, J.A. Espinosa, W. Money, Big data: issues and challenges moving forward, Proceedings of the IEEE 46th Annual Hawaii International Conference on System Sciences (HICSS' 13), (2013), pp. 995–1004 January.

[179] I. Hamilton, T. Oreszczyn, A. Summerfield, P. Steadman, S. Elam, A. Smith, Co-benefits of energy and buildings data: the case for supporting data access to achieve a sustainable built environment, Procedia Eng. 118 (2015) 958–968.

[180] D.T. Campbell, J.C. Stanley, Experimental and Quasi-Experimental Designs for Research, Rand McNally College Publishing Co, Chicago, 1963.

[181] M. Hammersley, Assessing validity in social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[182] W.R. Shadish, T.D. Cook, D.T. Campbell ("SCC"), Experimental and Quasi-experi¬mental Designs for Generalized Causal Inference, Houghton-Mifflin, Boston, 2002.

[183] M. Hammersley, P. Alasuutari, L. Bickman, J. Brannen (Eds.), Assessing Validity in Social Research. The SAGE Handbook of Social Research Methods, SAGE

Publications, London, UK, 2008.

[184] A. Falk, J.J. Heckman, Lab experiments are a major source of knowledge in the social sciences, Science 326 (5952) (2009) 535–538.

[185] M. Nicolson, G. Huebner, D. Shipworth, Are consumers willing to switch to smart time of use electricity tariffs? The importance of loss-aversion and electric vehicle ownership, Energy Res. Soc. Sci. 23 (January) (2017) 82–96.

[186] S. Waechter, B. Sütterlin, J. Borghoff, Michael Siegrist, Letters, signs, and colors: how the display of energy-efficiency information influences consumer assessments of products, Energy Res. Soc. Sci. 15 (May) (2016) 86–95.

[187] A.C.G. Cooper, Evaluating energy efficiency policy: understanding the' energy policy epistemology' may explain the lack of demand for randomized controlled trials, Energy Effic. 11 (4) (2018) 997–1008.

[188] A. Deaton, N. Cartwright, Understanding and misunderstanding randomized controlled trials, Soc. Sci. Med. 210 (2018) 2–21.

[189] J.A. List, Why economists should conduct field experiments and 14 tips for pulling one off, J. Econ. Perspect. 25 (3) (2011) 3–15.

[190] T. Cook, V. Wong, Better quasi-experimental practice, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[191] H. Bloom, The core analytics of randomized experiments for social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[192] E. Vine, M. Sullivan, L. Lutzenhiser, C. Blumstein, B. Miller, Experimentation and the evaluation of energy efficiency programs, Energy Effic. 7 (4) (2014) 627–640.

[193] G. Huebner, M. Nicolson, M. Fell, H. Kennard, S. Elam, C. Johnson, D. Shipworth, Are we heading towards a replicability crisis in energy efficiency research? A toolkit for improving the quality, transparency and replicability of energy efficiency impact evaluations, Proceedings of the European Council for an Energy Efficient Economy ECEEE 2017 Summer Study on Energy Efficiency: Consumption, Efficiency and Limits (2018).

[194] J. Henrich, S.J. Heine, et al., Most people are not WEIRD, Nature 466 (2010) 29.

[195] A. Falk, J.J. Heckman, Lab experiments are a major source of knowledge in the social sciences, Science 326 (5952) (2009) 535–538.

[196] M. Greenstone, T. Gayer, Quasi-experimental and experimental approaches to environmental economics, J. Environ. Econ. Manage. 57 (1) (2009) 21–44.

[197] B.D. Meyer, Natural and quasi-experiments in economics, J. Bus. Econ. Stat. 13 (2) (1995) 151–161.

[198] J.D. Angrist, J.S. Pischke, The credibility revolution in empirical economics: how better research design is taking the con out of econometrics, J. Econ. Perspect. 24 (2) (2010) 3–30.

[199] J. Bound, D.A. Jaeger, R.M. Baker, Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak, J. Am. Stat. Assoc. 90 (430) (1995) 443–450.

[200] M. Korsnes, T. Berker, Ruth Woods, domestication, acceptance and zero emission ambitions: insights from a mixed method, experimental research design in a Norwegian Living Lab, Energy Res. Soc. Sci. 39 (2018) 226–233.

[201] Y. Voytenko, K. McCormick, J. Evans, G. Schliwa, Urban living labs for sustainability and low carbon cities in Europe: towards a research agenda, J. Clean. Prod. 123 (2016) 45–54.

[202] S. Claude, et al., The Living Lab methodology for complex environments: Insights from the thermal refurbishment of a historical district in the city of Cahors, France, Energy Res. Soc. Sci. 32 (2017) 121–130.

[203] W. Canzler, et al., From "living lab" to strategic action field: bringing together energy, mobility, and Information Technology in Germany, Energy Res. Soc. Sci. 27 (2017) 25–35.

[204] C. Luederitz, et al., Learning through evaluation – a tentative evaluative scheme for sustainability transition experiments, J. Clean. Prod. 169 (December (15)) (2017) 61–76.

[206] P. Kivimaa, M. Hildén, D. Huitema, A. Jordan, J. Newig, Experiments in climate governance –a systematic review of research on energy and built environment transitions, J. Clean. Prod. 169 (2017) 17–29.

[207] E.L. Vine, C.M. Jones, Competition, carbon, and conservation: assessing the energy savings potential of energy efficiency competitions, Energy Res. Soc. Sci. 19 (September) (2016) 158–176.

[208] V. Rai, A.L. Beck, Play and learn: serious games in breaking informational barriers in residential solar energy adoption in the United States, Energy Res. Soc. Sci. 27 (May) (2017) 70–77.

[209] Joseph M. Brown, Efficient, Adaptable Simulations: A Case Study of a Climate Negotiation Game, Journal of Political Science Education, in press (2018) https://www.tandfonline.com/doi/abs/10.1080/15512169.2018.1431129?journalCode=upse20.

[210] L. Morganti, F.a Pallavicini, E. Cadel, A. Candelieri, F. Archetti, F. Mantovani, Gaming for Earth: serious games and gamification to engage consumers in pro-environmental behaviours for energy efficiency, Energy Res. Soc. Sci. 29 (July) (2017) 95–102.

[211] E.R. Frederiks, K. Stenner, E.V. Hobman, M. Fischle, Evaluating energy behavior change programs using randomized controlled trials: best practice guidelines for policymakers, Energy Res. Soc. Sci. 22 (December) (2016) 147–164.

[212] H. Bloom, The core analytics of randomized experiments for social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[213] E.R. Frederiks, K. Stenner, E.V. Hobman, M. Fischle, Evaluating energy behavior change programs using randomized controlled trials: best practice guidelines for policymakers, Energy Res. Soc. Sci. 22 (December) (2016) 147–164.

[214] T. Cook, V. Wong, Better quasi-experimental practice, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods,

SAGE Publications, London, UK, 2008.

[215] E.R. Frederiks, K. Stenner, E.V. Hobman, M. Fischle, Evaluating energy behavior change programs using randomized controlled trials: best practice guidelines for policymakers, Energy Res. Soc. Sci. 22 (December) (2016) 147–164.

[216] E. Patall, H. Cooper, Conducting a meta-analysis, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[217] M.J. Borenstein, L.V. Hedges, et al., Introduction to Meta-analysis, J Wiley & Sons Ltd, Chichester, UK, 2009.

[218] E. Patall, H. Cooper, Conducting a meta-analysis, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[219] M.J. Borenstein, L.V. Hedges, et al., Introduction to Meta-analysis, J Wiley & Sons Ltd, Chichester, UK, 2009.

[220] X. Labandeira, J.M. Labeaga, X. López-Otero, A meta-analysis on the price elasticity of energy demand, Energy Policy 102 (2017) 549–568.

[221] H. Pettifor, C. Wilson, et al., Social influence in the global diffusion of alternative fuel vehicles – a meta-analysis, J. Transp. Geogr. 62 (2017) 247–261.

[222] A. Srivastava, S. Van Passel, E. Laes, Assessing the success of electricity demand response programs: a meta-analysis, Energy Res. Soc. Sci. 40 (June) (2018) 110–117.

[223] M. Petticrew, H. Roberts, Systematic Reviews in the Social Sciences: A Practical Guide, Wiley-Blackwell, Malden, MA, 2006.

[224] S. Khalid, Khan, et al., Five steps to conducting a systematic review, J. R. Soc. Med. 96 (March) (2003) 118–121.

[225] G. Huebner, M. Nicolson, M. Fell, H. Kennard, S. Elam, C. Hanmer, C. Johnson, D. Shipworth, Are we heading towards a replicability crisis in energy efficiency research? A toolkit for improving the quality, transparency and replicability of energy efficiency impact evaluations, Proceedings of the European Council for an Energy Efficient Economy ECEEE 2017 Summer Study on Energy Efficiency: Consumption, Efficiency and Limits (2018).

[226] S. Khalid, Khan, et al., Five steps to conducting a systematic review, J. R. Soc. Med. 96 (March) (2003) 118–121.

[227] G. Huebner, M. Nicolson, M. Fell, H. Kennard, S. Elam, C. Hanmer, C. Johnson, D. Shipworth, Are we heading towards a replicability crisis in energy efficiency research? A toolkit for improving the quality, transparency and replicability of energy efficiency impact evaluations, Proceedings of the European Council for an Energy Efficient Economy ECEEE 2017 Summer Study on Energy Efficiency: Consumption, Efficiency and Limits (2018).

[228] D. Tranfield, et al., Towards a methodology for developing evidence-informed management knowledge by means of systematic review, Br. J. Manag. 14 (2003) 207–222.

[229] R. Gross, et al., The Costs and Impacts of Intermittency: an Assessment of the Evidence on the Costs and Impacts of Intermittent Generation on the British Electricity Network, (2006).

[230] S. Sorrell, et al., Global Oil Depletion: an Assessment of the Evidence for a Near-term Peak in Global Oil Production: UK Energy Research Centre, (2009).

[231] J. Rand, B. Hoen, Thirty years of North American wind energy acceptance research: What have we learned? Energy Res. Soc. Sci. 29 (July) (2017) 135–148.

[232] B.K. Sovacool, P. Kivimaa, S. Hielscher, K. Jenkins, Vulnerability and resistance in the United Kingdom's smart meter transition, Energy Policy 109 (October) (2017) 767–781.

[233] W.J. Sutherland, F.R. Wordley, A fresh approach to evidence synthesis, Nature 558 (June (21)) (2018) 364–366.

[234] M. Hammersley, On 'systematic'reviews of research literatures: a 'narrative'r-esponse to Evans & Benefield, Br. Educ. Res. J. 27 (5) (2001) 543–554.

[235] A. Fink, Conducting Research Literature Reviews: From Paper to the Internet, Sage, Thousand Oaks, 1998.

[236] H.M. Cooper, Synthesizing Research: a Guide for Literature Reviews, Sage, Thousand Oaks, 1998.

[237] R. Shaw, Conducting literature reviews, in: M.A. Forrester (Ed.), Doing Qualitative Research in Psychology: A Practical Guide. Sage, London, 2010.

[238] A. Oakley, D. Fullerton, A Systematic Review of Smoking Prevention Programmes for Young People, Centre for Evidence Informed Education Policy and Practice (EPPI), London, 1995.

[239] H.S. Becker, Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article, University of Chicago Press, Chicago, 2007.

[240] D. Dillman, J. Smyth, et al., Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons Inc, New Jersey, USA, 2014.

[241] D. Dillman, J. Smyth, et al., Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons Inc, New Jersey, USA, 2014.

[242] E. Leeuw, Self-administered questionnaires and standardized interviews, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[244] J. Cohen, Statistical power analysis, Curr. Dir. Psychol. Sci. 1 (3) (1992) 98–101.

[245] H.C. Kraemer, C. Blasey, How Many Subjects? Statistical Power Analysis in Research, Sage Publications, 2015.

[246] D. Dillman, J. Smyth, et al., Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons Inc, New Jersey, USA, 2014.

[247] E. Babbie, The Practice of Social Research, 12th edition, Wadsworth Publishing, London, UK, 2010.

[248] K.D. Hoover, S.J. Perez, Data mining reconsidered: encompassing and the general-to-specific approach to specification search, Econom. J. 2 (2) (1999) 167–191.

[249] J. Campos, N.R. Ericsson, D.F. Hendry, General-to-specific Modeling: an Overview and Selected Bibliography, (2005).

[250] K.D. Hoover, S.J. Perez, Three attitudes towards data mining, J. Econ. Methodol. 7

(2) (2000) 195–210.

[251] D.A. De Vaus, Surveys in Social Research, 3rd edition, Allen and Unwin Publishing, North Sydney, 1991.

[252] B.G. Tabachnick, L.S. Fidell, Using Multivariate Statistics, 5th ed., Allyn & Bacon/Pearson Education, Boston, MA, 2007.

[253] D.J. Denis, Applied Univariate, Bivariate, and Multivariate Statistics, Wiley, London, 2015.

[254] P. Kennedy, A Guide to Econometrics, 6th edition, Wiley-Blackwell, Toronto, Canada, 2008.

[255] D.A. Dickey, D.W. Jansen, D.L. Thornton, A primer on cointegration with an application to money and income, Cointegration, Palgrave Macmillan, London, 1994, pp. 9–45.

[256] J.D. Angrist, J.S. Pischke, The credibility revolution in empirical economics: how better research design is taking the con out of econometrics, J. Econ. Perspect. 24 (2) (2010) 3–30.

[257] J.M. Wooldridge, Econometric Analysis of Cross-sectional and Panel Data, MIT Press, 2010.

[258] N.M.A. Huijts, E.J.E. Molin, et al., Hydrogen fuel station acceptance: a structural equation model based on the technology acceptance framework, J. Environ. Psychol. 38 (0) (2014) 153–166.

[259] S. Mulaik, Objectivity in Science and Structural Equation Modeling, SAGE Publications, London, UK, 2004.

[260] R. Rummel, Applied Factor Analysis, Northwestern University Press, 1970.

[261] R. Hoyle, J. Duvall, Determining the number of factors in exploratory and confirmatory factor analysis, in: D. Kaplan (Ed.), The SAGE Handbook of Quantitative Methodology for the Social Sciences, SAGE Publications, London, UK, 2004.

[262] A. Costello, J. Osborne, Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis, Pract. Assess. Res. Eval. 10 (7) (2005) 1–9.

[263] J. Axsen, Citizen acceptance of new fossil fuel infrastructure: value theory and Canada's Northern Gateway Pipeline, Energy Policy 75 (2014) 255–265.

[264] J. Anable, Complacent Car Addicts' or' Aspiring Environmentalists'? Identifying travel behavior segments using attitude theory, Transp. Policy 12 (2005) 65–78.

[265] D. Hensher, J. Rose, et al., Applied Choice Analysis, Cambridge University Press, Cambridge, U.K, 2005.

[266] K. Train, Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand, MIT Press, Cambridge, MA, 1986.

[267] G. Ewing, E. Sarigollu, Assessing consumer preferences for clean-fuel vehicles: a discrete choice experiment, J. Public Policy Mark. 19 (1) (2000) 106–118.

[268] J. Axsen, J. Bailey, et al., Preference and lifestyle heterogeneity among potential plug-in electric vehicle buyers, Energy Econ. 50 (2015) 190–201.

[269] J. Magidson, J. Vermunt, D. Kaplan (Ed.), Latent Class Models. The SAGE Handbook of Quantitative Methodology for the Social Sciences, SAGE Publications, London, UK, 2004.

[270] J. Shen, Latent class model or mixed logit model? A comparison by transport mode choice data, Appl. Econ. 41 (22) (2009) 2915–2924.

[271] Paul P. Craig, Ashok Gadgil, Jonathan G. Koomey, What can history teach us? A retrospective examination of long-term energy forecasts for the United States, Ann. Rev. Energy Environ. 27 (2002) 83–118.

[272] L. Mundaca, L. Neij, et al., Evaluating energy efficiency policies with energy-economy models, Annu. Rev. Environ. Resour. 35 (1) (2010) 305–344.

[273] M. Jaccard, Combining top down and bottom up in energy economy models, in: J. Evans, L. Hunt (Eds.), International Handbook on the Economics of Energy, Edward Elgar Publishing, Cheltenham, UK, 2009.

[274] V. Krey, Global energy-climate scenarios and models: a review, Wiley Interdiscip. Rev. Energy Environ. 3 (4) (2014) 363–383.

[275] Van den Bergh, C.J.M. Jeroen (Eds.), Handbook of Environmental and Resource Economics, Edward Elgar, Cheltenham, UK, 1999.

[276] V. Bosetti, R. Gerlagh, S.P. Schleicher (Eds.), Modeling Sustainable Development: Transitions to a Sustainable Future, Edward Elgar, Cheltenham, UK, 2009.

[277] E. Trutnevyte, Does cost optimization approximate the real-world energy transition? Energy 106 (2016) 182–193.

[278] L. Bergman, Energy policy modelling: a survey of general equilibrium approaches, J. Policy Model. 10 (3) (1988) 377–399.

[279] A. Weidlich, D. Veit, A critical survey of agent-based wholesale electricity market models, Energy Econ. 30 (2008) 1728–1759.

[280] S.C. Bhattacharyya, Applied general equilibrium models for energy studies: a survey, Energy Econ. 18 (1996) 145–164.

[281] V. Krey, Global energy-climate scenarios and models: a review, Wiley Interdiscip. Rev. Energy Environ. 3 (4) (2014) 363–383.

[282] M. Jaccard, Combining top down and bottom up in energy economy models, in: J. Evans, L. Hunt (Eds.), International Handbook on the Economics of Energy, Edward Elgar Publishing, Cheltenham, UK, 2009.

[283] M. Jaccard, Combining top down and bottom up in energy economy models, in: J. Evans, L. Hunt (Eds.), International Handbook on the Economics of Energy, Edward Elgar Publishing, Cheltenham, UK, 2009.

[284] F.F. Nerini, I. Keppo, N. Strachan, Myopic decision making in energy system decarbonisation pathways. A UK case study, Energy Strategy Rev. 17 (2017) 19–26.

[285] D.L. McCollum, C. Wilson, et al., Improving the behavioral realism of global integrated assessment models: an application to consumers' vehicle choices, Transp. Res. D Transp. Environ. 55 (Suppl. C) (2017) 322–342.

[286] International Risk Governance Council (2015).

[287] S.J. DeCanio, Economic Models of Climate Change: A Critique, Palgrave MacMillan, New York, 2003.

[288] M. Jaccard, Energy Planning and Management: Methodologies and Tools. Encyclopedia of Life Support Systems. UNESCO, UNESCO, Paris, France, 2002.

[289] S. Sorrell, Energy substitution, technical change and rebound effects, Energies 7 (5) (2014) 2850–2873.

[290] A. Kaya, D. Csala, S. Sgouridis, Constant elasticity of substitution functions for energy modeling in general equilibrium integrated assessment models: a critical review and recommendations, Clim. Change 145 (1-2) (2017) 27–40.

[291] M. Jaccard, Energy Planning and Management: Methodologies and Tools. Encyclopedia of Life Support Systems. UNESCO, UNESCO, Paris, France, 2002.

[292] International Risk Governance Council, Assessment of Future Energy Demand: a Methodological Review Providing Guidance to Developers and Users of Energy Models and Scenarios, (2015).

[293] M. Jaccard, Combining top down and bottom up in energy economy models, in: J. Evans, L. Hunt (Eds.), International Handbook on the Economics of Energy, Edward Elgar Publishing, Cheltenham, UK, 2009.

[294] S. Messner, L. Schrattenholzer, MESSAGE–MACRO: linking an energy supply model with a macroeconomic module and solving it iteratively, Energy 25 (3) (2000) 267–282.

[295] J. Axsen, D.C. Mountain, et al., Combining stated and revealed choice research to simulate the neighbor effect: the case of hybrid-electric vehicles, Resour. Energy Econ. 31 (3) (2009) 221–238.

[296] M. Horne, M. Jaccard, et al., Improving behavioral realism in hybrid energy-economy models using discrete choice studies of personal transportation decisions, Energy Econ. 27 (1) (2005) 59–77.

[297] R. Murphy, M. Jaccard, Energy efficiency and the cost of GHG abatement: a comparison of bottom-up and hybrid models for the US, Energy Policy 39 (11) (2011) 7146–7155.

[298] M. Wolinetz, J. Axsen, et al., Simulating the value of electric-vehicle–grid integration using a behaviourally realistic model, Nat. Energy 3 (2) (2018) 132–139.

[299] J. Sterman, Business Dynamics: Systems Thinking and Modeling for a Complex World, McGraw Hill, New York, NY, 2000.

[300] J. Struben, J. Sterman, Transition challenges for alternative fuel vehicles and transportation systems, Environ. Plann. B Plann. Des. 35 (6) (2008) 1070–1097.

[301] M. Dale, S. Krumdieck, P. Bodger, Global energy modelling—a biophysical approach (GEMBA) part 2: methodology, Ecol. Econ. 73 (2012) 158–167.

[302] U. Bardi, The Limits to Growth Revisited, Springer Science & Business Media, 2011.

[303] V. Rai, A.D. Henry, Agent-based modelling of consumer energy choices, Nat. Clim. Chang. 6 (2015) 556–562.

[304] N. Stern, Current climate models are grossly misleading, Nature 530 (February (25)) (2016) 407–409 2016.

[305] G. Metcalf, J. Stock, The Role of Integrated Assessment Models in Climate Policy: A User's Guide and Assessment. Discussion Paper 2015-68, Harvard Project on Climate Agreements, Cambridge, Mass, 2015 March.

[306] L.D. Anadón, E. Baker, V. Bosetti, Integrating uncertainty into public energy research and development decisions, Nat. Energy 2 (2017) 1707.

[307] G. McCracken, The Long Interview, Sage Publications, Newbury Park, California, 1988.

[309] J. Morse, Reframing Rigor in Qualitative Inquiry, SAGE Publications, London, UK, 2018.

[310] M. Hammersley, Assessing validity in social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[311] M. Hammersley, Assessing validity in social research, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[312] G. McCracken, The Long Interview, Sage Publications, Newbury Park, California, 1988.

[313] J. Smithson, Focus Groups, SAGE Publications, London, UK, 2008.

[314] J. Smithson, Focus Groups, SAGE Publications, London, UK, 2008.

[315] G. McCracken, The Long Interview, Sage Publications, Newbury Park, California, 1988.

[316] J. Bratich, Observation in a surveilled world, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, 5th edition, SAGE Publications, London, UK, 2018.

[317] R.K. Yin, Case Study Research: Design and Methods, Sage, London, 2003.

[318] L. Prior, Documents and action, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[319] J. Maxwell, Qualitative Research Design: An Interactive Approach, 2nd edition, Sage Publications, Thousand Oaks, CA, 2005.

[320] G. McCracken, The Long Interview, Sage Publications, Newbury Park, California, 1988.

[321] L. Prior, Documents and action, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[322] E. Margolis, R. Zunjarward, Visual research, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018.

[323] C. Geertz, The Interpretation of Cultures, Basic Books, New York, Inc, 1973.

[324] R. Burnett, Accounts and narratives, in: B. Montgomery, S. Duck (Eds.), Studying Interpersonal Interaction, The Guildford Press, New York, 1991, pp. 121–140.

[325] C. Antaki, Discourse analysis and conversation analysis, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[326] M. Hyvarinen, Analyzing narratives and story-telling, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[327] J. Axsen, K.S. Kurani, Developing sustainability-oriented values: insights from households in a trial of plug-in hybrid electric vehicles, Glob. Environm. Changee 23 (1) (2013) 70–80.

[328] S. Gunster, P. Saurette, Storylines in the sands: news, narrative and ideology in the Calgary Herald, Can. J. Commun. 39 (3) (2014) 333–359.

[329] D. Hugh, K. Brown, Ten years to prevent catastrophe? Discourses of climate change and international development in the UK press, Glob. Environ. Change 19 (2009) 191–202.

[330] K. Charmaz, Reconstructing grounded theory, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[331] B.G. Glaser, A.L. Strauss, The Discovery of Grounded Theory, Aldine, Chicago, 1967.

[332] A.L. Strauss, Qualitative Analysis for Social Scientists, Cambridge University Press, Cambridge, 1987.

[333] S. Chase, Narrative inquiry: towards theoretical and methodological maturity, in: N. Denzin, Y. Lincoln (Eds.), The SAGE Handbook of Qualitative Research, SAGE Publications, London, UK, 2018.

[334] K. Charmaz, Reconstructing grounded theory, in: P. Alasuutari, L. Bickman, J. Brannen (Eds.), The SAGE Handbook of Social Research Methods, SAGE Publications, London, UK, 2008.

[335] R.K. Yin, Case Study Research: Design and Methods, Sage publications, 2013.

[336] R.K. Yin, Case Study Research: Design and Methods, Sage publications, 2013.

[337] A. Bennett, C. Elman, Qualitative research: recent developments in case study methods, Annu. Rev. Political Sci. 9 (1) (2006) 455–476.

[338] B. Flyvbjerg, Five misunderstandings about case-study research, Qual. Inq. 12 (2) (2006) 219–245.

[339] K.M. Eisenhardt, M.E. Graebner, Theory building from cases: opportunities and challenges, Acad. Manag. J. 50 (1) (2007) 25–32.

[340] S. Lieberson, Small N's and big conclusions: an examination of the reasoning in comparative studies based on a small number of cases, Soc. Forces 70 (2) (1991) 307–320.

[341] J. Seawright, J. Gerring, Options case selection techniques in case study research: a menu of qualitative and quantitative options, Polit. Res. Q. 61 (2008) 294–308.

[342] S. Strunz, The German energy transition as a regime shift, Ecol. Econ. 100 (2014) 150–158.

[343] M.A. Brown, B.K. Sovacool, Germany's feed-in tariff, 1990-2009, Climate Change and Global Energy Security: Technology and Policy Options, MIT Press, Cambridge, 2011, pp. 253–260.

[344] B.K. Sovacool, C. Cooper, Green means 'Go?' A colorful approach to a national renewable portfolio standard, Electr. J. 19 (August/September (7)) (2006) 19–32.

[345] R. Wiser, S. Pickle, C. Goldman, Renewable energy policy and electricity re-structuring: a California case study, Energy Policy 26 (6) (1998) 465–475.

[346] A. Rawlani, B.K. Sovacool, Building Responsiveness to Climate Change through Community Based Adaptation in Bangladesh, Mitig. Adapt. Strateg. Glob. Change 16 (December (8)) (2011) 845–863.

[347] R. Brouwer, S. Akter, L. Brander, E. Haque, Socioeconomic vulnerability and adaptation to environmental risk: a case study of climate change and flooding in Bangladesh, Risk Anal. 27 (April (2)) (2007) 313–326.

[348] P. Johnstone, A. Stirling, Comparing Nuclear Power Trajectories in Germany And the UK: From 'Regimes' to 'Democracies' in Sociotechnical Transitions and Discontinuities, SPRU Working Paper Series, 2015/8 (June), available at (2018) https://www.sussex.ac.uk/webteam/gateway/file.php?name=2015-18-swps-johnston-stirling.pdf&site=25.

[349] B. Rennkamp, S. Haunss, K. Wongsa, A. Ortega, E. Casamadrid, Competing coalitions: the politics of renewable energy and fossil fuels in Mexico, South Africa and Thailand, Energy Res. Soc. Sci. 34 (2017) 214–223.

[350] L.A. Malone, The chernobyl accident: a case study in international law regulating state responsibility for transboundary nuclear pollution, 12 colum, J. Environ. Law. 203 (1987).

[351] B. Wynne, Sheepfarming after chernobyl: a case study in communicating scientific information, Environment 31 (2) (1989) 10–39.

[352] V.H.M. Visschers, M. Siegrist, How a nuclear power plant accident influences acceptance of nuclear power: results of a longitudinal study before and after the Fukushima disaster, Risk Anal. 33 (2013) 333–347.

[353] K. Araujo, Low Carbon Energy Transitions: Turning Points in National Policy and Innovation, Oxford University Press, Oxford, 2018.

[354] B.K. Sovacool, H.H. Lindboe, O. Odgaard, Is the Danish wind energy model replicable for other countries? Electr. J. 21 (March (2)) (2008) 27–38.

[355] J. Goldemberg, S.T. Coelho, P.M.ário Nastari, O. Lucon, Ethanol learning curve—the Brazilian experience, Biomass Bioenergy 26 (3) (2004) 301–304.

[356] J. Espen Skogdalen, J. Erik Vinnem, Quantitative risk analysis of oil and gas drilling, using Deepwater Horizon as case study, Reliab. Eng. Syst. Saf. 100 (2012) 58–66.

[357] C. Kuzemko, C. Mitchell, M. Lockwood, R. Hoggett, Policies, politics and demand side innovations: the untold story of Germany's energy transition, Energy Res. Soc. Sci. 28 (June) (2017) 58–67.

[358] V. Lauber, S. Jacobsson, The politics and economics of constructing, contesting and restricting socio-political space for renewables—the German Renewable Energy Act, Environ. Innov. Soc. Trans. 18 (2016) 147–163.

[359] B.K. Sovacool, The best of both worlds: environmental federalism and the need for federal action on renewable energy and climate change, Stanf. Environ. Law J. 27 (June (2)) (2008) 397–476.

[360] B.K. Sovacool, Bamboo beating bandits: conflict, inequality, and vulnerability in the political ecology of climate change adaptation in Bangladesh, World Dev. 102 (February) (2018) 183–194.

[361] B.K. Sovacool, S.V. Valentine, The National Politics of Nuclear Power: Economics, Security, and Governance, Routledge Global Security Studies Series, London, 2012.

[362] B.K. Sovacool, S.V. Valentine, The socio-political economy of nuclear energy in China and India, Energy 35 (September (9)) (2010) 3803–3813.

[363] J. Gerring, Case Study Research, Cambridge University Press, New York, 2005.

[364] J. Gerring, What is a case study and what is it good for? Am. Polit. Sci. Rev. 98 (May (2)) (2004) 341–354.

[365] B. Flyvbjerg, Five misunderstandings about case study research, Qual. Inq. 12 (April (2)) (2006) 219–245 2006.

[366] F.W. Geels, Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study, Res. Policy 31 (8–9) (2002) 1257–1274.

[367] M. Oteman, M. Wiering, J.K. Helderman, The institutional space of community initiatives for renewable energy: a comparative case study of the Netherlands, Germany and Denmark, Energy Sustain. Soc. 4 (1) (2014) 11.

[368] J. Gerring, Case Study Research, Cambridge University Press, New York, 2005.

[369] J. Gerring, What is a case study and what is it good for? Am. Polit. Sci. Rev. 98 (May (2)) (2004) 341–354.

[370] J. Popay, et al., Rationale and standards for the systematic review of qualitative literature in health services research, Qual. Health Res. 8 (May (3)) (1998) 341–351.

[371] D. Tranfield, et al., Towards a methodology for developing evidence-informed management knowledge by means of systematic review, Br. J. Manag. 14 (2003) 207–222.

[372] M. Petticrew, H. Roberts, Theory and methods: evidence, hierarchies, and typologies: horses for courses, J. Epidemiol. Community Health 57 (2003) 527–529.

[373] J. Daly, et al., A hierarchy of evidence for assessing qualitative health research, J. Clin. Epidemiol. 60 (2007) 43–49.

[374] J. Daly, et al., A hierarchy of evidence for assessing qualitative health research, J. Clin. Epidemiol. 60 (2007) 43–49.

[375] J. Gerring, Case Study Research, Cambridge University Press, New York, 2005.

[376] J. Gerring, What is a case study and what is it good for? Am. Polit. Sci. Rev. 98 (May (2)) (2004) 341–354.

[377] M. Alvesson, Y. Gabriel, R. Paulsen, Return to Meaning: a Social Science With Something to Say, Oxford University Press, 2017.

[378] B. Flyvbjerg, Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again, Cambridge University Press, Cambridge, 2001.

[379] A. Sayer, Method in Social Science: Revised, 2nd edition, Routledge, 2010.

[380] R. Burke Johnson, A.J. Onwuegbuzie, Mixed methods research: a research paradigm whose time has come, Educ. Res. 33 (7) (2004) 14–26.

[381] R. Burke Johnson, A.J. Onwuegbuzie, Mixed methods research: a research paradigm whose time has come, Educ. Res. 33 (7) (2004) 14–26.

[382] S.R. Barley, When I write my masterpiece: thoughts on what makes a paper interesting, Acad. Manag. J. 49 (1) (2006) 16–20.

[383] B. Hancke, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[384] P. Belt, M. Mottonen, J. Harkonen, Tips for Writing Scientific Journal Articles, Industrial Engineering and Management Working Papers, (2011), p. 5.

[385] A. Fink, Conducting Research Literature Reviews: From Paper to the Internet, Sage, Thousand Oaks, 1998.

[386] H.M. Cooper, Synthesizing Research: A Guide for Literature Reviews, Sage, Thousand Oaks, 1998.

[387] R. Shaw, Conducting literature reviews, in: M.A. Forrester (Ed.), Doing Qualitative Research in Psychology: A Practical Guide. Sage, London, 2010.

[388] B. Turnheim, F. Geels, Regime destabilisation as the flipside of energy transitions: lessons from the history of the British coal industry (1913–1997), Energy Policy 50 (November) (2012) 35–49.

[389] J.C.D. Roberts, Discursive destabilisation of socio-technical regimes: negative storylines and the discursive vulnerability of historical American railroads, Energy Res. Soc. Sci. 31 (September) (2017) 86–99.

[390] F.W. Geels, J.W. Schot, Typology of sociotechnical transition pathways, Res. Policy 36 (3) (2007) 399–417.

[391] C.-fei Chen, X. Xiaojing, L. Arpan, Between the technology acceptance model and sustainable energy technology acceptance model: investigating smart meter acceptance in the United States, Energy Res. Soc. Sci. 25 (March) (2017) 93–104.

[392] A. Cherp, J. Jewell, The three perspectives on energy security: intellectual history, disciplinary roots and the potential for integration, Curr. Opin. Environ. Sustain. 3 (4) (2011) 202–212.

[393] T. Van de Graaf, B.K. Sovacool, A. Ghosh, F. Kern, M.T. Klare, States, markets, and institutions: integrating international political economy and global energy politics, in: T. Van de Graaf, B.K. Sovacool, F. Kern, A. Ghosh, M.T. Klare (Eds.), The Palgrave Handbook of the International Political Economy of Energy, Palgrave, Basingstoke UK/New York USA, 2016, pp. 1–44.

[394] B.K. Sovacool, S.V. Valentine, M.J. Bambawale, M.A. Brown, T.D.F. Cardoso, S. Nurbek, G. Suleimenova, L. Jinke, X. Yang, A. Jain, A.F. Alhajji, A. Zubiri, Exploring propositions about perceptions of energy security: an international survey, Environ. Sci. Policy 16 (January (1)) (2012) 44–64.

[395] B.K. Sovacool, A. Gilbert, D. Nugent, Risk, innovation, electricity infrastructure and construction cost overruns: testing six hypotheses, Energy 74 (September) (2014) 906–917.

[396] B.K. Sovacool, G. Walter, T. Van de Graaf, N. Andrews, Energy governance, transnational rules, and the resource curse: exploring the effectiveness of the extractive industries transparency initiative (EITI), World Dev. 86 (July) (2016) 179–192.

[397] M. Broman Toft, G. Schuitema, J. Thøgersen, The importance of framing for consumer acceptance of the Smart Grid: a comparative study of Denmark, Norway

and Switzerland, Energy Res. Soc. Sci. 3 (September) (2014) 113–123.

[398] M. Ratinen, P. Lund, Policy inclusiveness and niche development: examples from wind energy and photovoltaics in Denmark, Germany, Finland, and Spain, Energy Res. Soc. Sci. 6 (March) (2015) 136–145.

[399] F. Kern, P. Kivimaa, M. Martiskainen, Policy packaging or policy patching? The development of complex energy efficiency policy mixes, Energy Res. Soc. Sci. 23 (January) (2017) 11–25.

[400] S. Pacala, R. Socolow, Stabilization wedges: solving the climate problem for the next 50 years with current technologies, Science 305 (August (13)) (2004) 968–972.

[401] M.Z. Jacobson, Review of solutions to global warming, air pollution, and energy security, Energy Environ. Sci. 2 (2009) 148–173.

[402] B.K. Sovacool, Rejecting renewables: the socio-technical impediments to renewable electricity in the United States, Energy Policy 37 (November (11)) (2009) 4500–4513.

[403] V.V. Raman, S. Iniyan, R. Goic, A review of climate change, mitigation and adaptation, Renew. Sustain. Energy Rev. 16 (1) (2012) 878–897.

[404] S.J. Darby, Coal fires, steel houses and the man in the moon: local experiences of energy transition, Energy Res. Soc. Sci. 31 (September) (2017).

[405] V. Drummond, E. Grubert, Fault lines: seismicity and the fracturing of energy narratives in Oklahoma, Energy Res. Soc. Sci. 31 (September) (2017) 128–136.

[406] D. Pargman, E. Eriksson, M. Höök, J. Tanenbaum, M. Pufal, J. Wangel, What if there had only been half the oil? Rewriting history to envision the consequences of peak oil, Energy Res. Soc. Sci. 31 (September) (2017) 170–178.

[407] R. Garud, P. Karnoe, Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship, Res. Policy 32 (2003) 277–300 2003.

[408] N. Balta-Ozkan, B. Boteler, O. Amerighi, European smart home market development: public views on technical and economic aspects across the United Kingdom, Germany and Italy, Energy Res. Soc. Sci. 3 (September) (2014) 65–77.

[409] B.K. Sovacool, G. Walter, Major hydropower states, sustainable development, and energy security: insights from a preliminary cross-comparative assessment, Energy 142 (1) (2018) 1074–1082.

[410] M. Moezzi, K.B. Janda, S. Rotmann, Using stories, narratives, and storytelling in energy and climate change research, Energy Res. Soc. Sci. 31 (September) (2017) 1–10.

[411] S.B. Chatman, Story and Discourse: Narrative Structure in Fiction and Film, Cornell University Press, 1978.

[413] A. Oshima, A. Hogue, Writing Academic English, Longman, 2000.

[414] B. Hancke, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[415] W. Strunk, E.B. White, The Elements of Style, 4th ed., Pearson, 2000.

[416] B. Hancke, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[418] G. Orwell, Politics and the English language, The American Guide to Modern Literature, Norton & Company, New York, 1946.

[419] Kate O' Neil, et al., Seeing complexity: visualization tools in global environmental politics and governance, J. Environ. Stud. Sci. 7 (4) (2017) 490–506.

[420] A.C. Niehaus, Write fiction to discover something new in your research, Nature 557 (2018) 269 2018.

[421] B. Hancke, Intelligent Research Design: A Guide for Beginning Researchers in the Social Sciences, Oxford University Press, Oxford, 2009.

[422] M. Garber, Writing Advice from a (Newly Minted) Nobel Winner (October 5), available at, The Atlantic, 2017, https://www.theatlantic.com/entertainment/archive/2017/10/writing-advice-from-a-newly-minted-nobel-winner/542136/.

[423] A. Abbott, Methods of Discovery: Heuristics for the Social Sciences, W.W. Norton & Co., New York, 2004.

[424] R.R. Alford, The Craft of Inquiry: Theories, Methods and Evidence, Oxford University Press, 1998.

[425] M.F. Bellemare, How to Publish in Academic Journals. AAEA Early Career Professionals Workshop, Vail, CO, June 2 (2017).