



Strengthening research systems to produce Transparent, Reproducible, and Ethical Evidence (TREE) – A guidance note on 3ie’s TREE Review Framework

Abstract: As researchers generating evidence to inform decision-making, our goal is to produce credible, unbiased evidence that is meaningful for decision-makers and conducted in an ethical manner. With this goal in mind, this paper presents a framework for conducting timely risk monitoring of international social science research through the intersecting lens of transparency, reproducibility, and ethics. The Transparent, Reproducible, and Ethical Evidence (TREE) Review Framework aims to achieve three objectives: (i) establish ethical standards; (ii) better integrate best practices in TREE into research workflow; and (iii) establish a timely, independent process that documents, monitors, and mitigates risks to achieve our goal. The paper motivates this framework by reflecting on the credibility crisis in social science research and calls for best practices in transparency and reproducibility to address this crisis, as well as recent attention on the need for improved transparency in research ethical decision-making. The paper then discusses gaps in the existing international social science research system, including a lack of common ethical requirements beyond the foundational ethical principles of beneficence, respect for persons, and justice, as well as gaps in the existing Institutional Review Board (IRB) process. The TREE Review Framework is then presented as a response to these gaps. The framework is guided by ten ethical requirements, informed by a growing literature on application of clinical research ethical requirements to social science. The TREE Review Framework questionnaire and process are then discussed, and the paper concludes with a set of lessons learned from a pilot within the International Initiative for Impact Evaluation (3ie) portfolio. As both an evidence producer and consumer, the 3ie experience offers insights on how to resource TREE practices that produce credible, unbiased, meaningful evidence in an ethical manner.

Keywords: open science, transparency, reproducibility, ethics, scientific integrity, responsible conduct of research,

New Delhi

215-216, Rectangle One
D-4, Saket District Centre
New Delhi - 110017, India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

London

c/o LIDC, 20 Bloomsbury Square
London WC1A 2NS
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 207 958 8351/8350

Washington, DC

1111 19th Street, NW, Suite 700
Washington, DC 20036
United States of America

3ieus@3ieimpact.org
Tel: +1 202 629 3939

Acknowledgements

The development of this guidance note is possible thanks to the International Initiative for Impact Evaluation (3ie) and its investments in global public goods for Transparent, Reproducible, and Ethical Evidence (TREE). This work is led by Jennifer Sturdy and Sayak Khatua from 3ie and is informed by calls for more transparency and reproducibility in research design and implementation (Miguel, Camerer, et al 2014) and research ethics (Asiedu, Karlan, et al 2021) and Evans 2021), as well as calls to reimagine researchers' role in ethical oversight for human research protections (Devereaux 2014; Bloss, Nebeker, et al 2016). The ethical requirements proposed in this paper are informed by a growing literature base examining how clinical research ethics apply to social science research ([MacKay, forthcoming](#)). Special thanks to peer reviewers Douglas MacKay (UNC), Scott Desposato (UCSD), Sean Grant (IUPUI), Sebastian Martinez (3ie), Chandan Jain (3ie), Sixto Aquino, participants in the 2022 BITSS Annual Meeting, Leah Bridle (CEGA), Aleksander Bogdanaski (CEGA), Kristina Hallez (CEGA), Carson Christiano (CEGA), Charlotte Lane (3ie), Saumya Mathur (3ie), Tom Wein (IDInsight), Mikaela Cochran (C4ED) and other participants in the Network for Research Ethics (Nethics) group. The development of the proposed questionnaire and process was strengthened by the informative feedback from the Swashakt (led by Bidisha Barooah) and Collective Impact (led by Stuti Tripathi) 3ie research teams.

Contents

- Motivation.....4
- Objectives 12
- Ethical requirements..... 12
 - Use transparency as a tool 12
 - Maximize social value and meaningful use 13
 - Balance power and align incentives 14
 - Preserve standard of care..... 15
 - Value and prioritize community engagement 16
 - Use fair methods..... 16
 - Ensure fair treatment for participants 17
 - Ensure appropriate informed consent and protection of confidentiality..... 18
 - Ensure favorable risk-benefit ratio and accountability 19
 - Ensure favorable cost-benefit ratio and accountability 23
- TREE Review Framework 25
 - Questionnaire 25
 - Review Committee..... 25
 - Process..... 26
- Lessons learned from 3ie pilot..... **Error! Bookmark not defined.**
- Appendix 1: Understanding vulnerability..... 27
- Appendix 2: TREE Review Questionnaire 28
- References..... 29



Motivation

More evidence in decision-making is better than less evidence in decision-making. There is a moral argument¹ that motivates investments in rigorous social science research² – such as program evaluations that use randomized control trials (RCTs) and quasi-experimental methods. This argument is focused on the need to measure and understand the efficacy and effectiveness of interventions and policies designed to improve social and economic outcomes (CGD 2006). This argument is supported by data that shows a correlation between rigorous monitoring and evaluation systems and stronger programs and program outcomes³. However, three conditions must be satisfied for this argument to remain true and each of the three conditions face credible risks: (i) application of social science research methods produce credible, unbiased evidence; (ii) evidence produced is useful to decision-makers; and (iii) evidence generation and use improves people’s lives and does not introduce harm.

The application of rigorous research methods leads to credible and unbiased data, analysis, and interpretation by researchers, but there are risks. Evidence generated by the application of rigorous research methods can improve decision-making and policy. However, designing and implementing rigorous research, particularly in challenging settings, is hard. It is fraught with risks, including studying solutions that do not align with problems, survey design errors, low response rates, and poor data quality (Karlan and Appel 2016). In addition, a decade of evidence regarding failures to replicate study results across large bodies of evidence (Ioannidis 2005; Open Science Collaboration 2015) calls into question the credibility of published social science research findings. This has drawn attention to additional risks which contribute to research credibility and waste, including perverse incentives created by funders and publishers for researchers to selectively report findings and/or p-hack analysis toward a specific statistical significance threshold to increase the likelihood of publication or future funding. In addition, a lack of access to data and code prevents independent verification of findings (Wood, Müller, and Brown 2018; Gertler, Galiani, and Romero 2018). And it remains unknown how prevalent outright fraud and data manipulation is, though recent examples highlight this is a risk as well (Broockman, Kalla, and Aronow 2014; DataColada 2021).

The generation of evidence leads to unbiased use of evidence by policymakers and funders for decision-making, but there are risks. While demands for more evidence in policymaking have increased, policymakers and decisionmakers continue to face internal and external constraints to evidence use. Internal constraints include individual bias – such as overconfidence and availability bias (Cojocar, Datta, et al 2021) – which limit decisionmakers’ use of evidence to update or course correct decisions otherwise based on intuition and experience. External constraints include poor timing, availability, and relevance of the evidence, as well as political pressures to favor certain decisions regardless of evidence.

The generation of evidence does not harm research participants, staff, and bystanders, but there are risks. Evidence generation often relies on data regarding socio-economic characteristics and behaviors of individuals, households, and other human subjects. As is often the case in social science research, these human subjects can have a range of vulnerabilities and liability to harm. Vulnerabilities can call into question the efficacy of an individual’s informed

¹ For reference, read “The Moral Case for Evidence in Policymaking” - <https://hewlett.org/moral-case-evidence-policy-making/>

² Definitions of research vary. According to US Health and Human Services, research is defined as “as a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.” For this paper, research refers to as a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to knowledge – regardless of extent of external validity. There are many RCTs that may produce limited generalizable knowledge, but the methods they employ are similar in terms of engagement with human subjects and collection of PII and often sensitive data.

³ For reference read “Is Good Monitoring and Evaluation the Secret to Success for World Bank Programs?” - [Is Good Monitoring and Evaluation the Secret to Success for World Bank Programs? | Center For Global Development \(cgdev.org\)](https://www.cgdev.org/publication/is-good-monitoring-and-evaluation-the-secret-to-success-for-world-bank-programs)

consent to participate in research activities due to factors such as a lack of comprehension or other forces that affect autonomy (NBAC 2001). In addition to vulnerability, human subjects can face liability to harm. When research participants experience certain vulnerabilities – such as poverty, discrimination – they may experience higher risk of harm because of the research activities. Risks can result from the participants’ engagement with the intervention (i.e., program participants can have worse outcomes after the intervention such as in Banerjee, Duflo, Glennerster 2008). Risks can also arise from the research participants engagement with the research, such as feelings of exploitation, research fatigue, or re-traumatization (Weber, Hardiman, et al 2021) or a loss of confidentiality⁴. Risks of loss of confidentiality or privacy can extend beyond participants to bystanders depending on the type of research methods. Application of certain research methods – such as satellite imagery, use of tools like Facebook or other big data sources – may reduce participants’ awareness they are even in a study which can result in loss of trust in the integrity of the scientific process. As the COVID-19 pandemic also illuminated, survey methods that require face-to-face interaction can introduce risk to the health, safety, and welfare of staff and participants alike. Other examples of risks of harm for participants and/or staff include political threats and targeting, unwarranted arrest or targeting by law enforcement, targeting for violence or theft, exposure to infectious disease/illness, unsafe road conditions or transportation, and inadequate access to food and secure lodging during data collection.

Calls for more transparent, reproducible, and ethical evidence (TREE) offer a roadmap for better practices and a means by which some threats to these conditions are mitigated. The open science movement promotes evolving practices and tools to increase research transparency and reproducibility with the goal of strengthening research credibility and reducing bias and research waste by fostering scientific integrity (Kretser, Murphy, Bertuzzi, et al. 2019). Miguel, Camerer et al (2014) highlight best practices for (i) pre-specification – through registration and pre-analysis plans (PAPs), (ii) disclosure – through standardized reporting requirements, and (iii) data and code sharing. Additional calls to better align analysis reporting with pre-specification complement this work (Laitin, Miguel, et al 2021; Claesen, Gomes et al 2021). Hoces de la Guardia, Grant, and Miguel (2020) highlight similar practices and tools important for more transparency in policy analysis. In addition, efforts to improve reporting transparency for individual studies support meta-analysis and evidence synthesis work, which can improve relevance and usefulness for decisionmakers.

Table 1: Addressing threats to research credibility with best practices in transparency and reproducibility			
Problems we want to solve	Researcher degrees of freedom, p-hacking,	Publication Bias and the File Drawer	Failure to reproduce analysis
Best practices in transparency and reproducibility to address challenges	Study registration; Pre-analysis Plans; Standardized Reporting Templates		Push-button replication; Data and code sharing; Dynamic documents

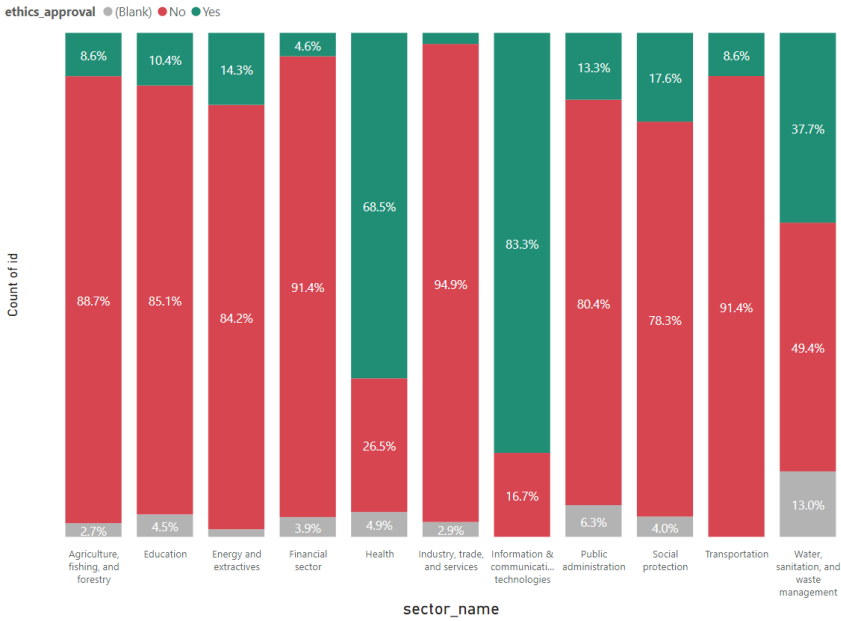
Recently, new attention has focused on the general lack of transparency regarding how research teams and their funders consider ethical issues throughout the research life cycle. In a sample of more than 9,000 publications⁵ in the 3ie Development Evidence Portal⁶, more than 6,000 report their evaluation methods as experimental, quasi-experimental, and observational. For this sample, 4.6% do not report their ethics review status, 37.3% report they completed ethics review, and 58.1% report they did not complete ethics review. As demonstrated in the chart below, there is wide variation across sectors, with a higher percentage of studies reporting ethics review in Health and Information Technology sectors, and less than 10% reporting a completed ethics review in Agriculture, Finance, and Transportation sectors. While some percentage of this sample

⁴ For a recent example, reference the discussions around the UN data sharing - <https://www.theguardian.com/global-development/2021/jun/15/un-put-rohingya-at-risk-by-sharing-data-without-consent-says-rights-group>

⁵ For the total sample, most publications do not report on Review status (39%). Of those who report Review status, 23.8% report they completed Ethics review and 37.1% report they did not complete Ethics review.

⁶ <https://developmentevidence.3ieimpact.org/>

is fully exempt from ethics review because it did not rely on engagement with human subjects, the minimal level of reporting leaves many questions – *who is deciding when a study does not require ethics review and what are the reasons for not conducting ethics review? For studies that complete ethics review, what issues were raised and discussed and how did the research team mitigate identified risks? How often was the study reviewed? How are protocols reviewed with an understanding of local context and vulnerabilities and liabilities to harm?* Although just a sub-sample of international social science research, this data suggests more can be done regarding coverage and transparency of the ethics review process.



Asiedu, Karlan, et al (2021), Evans (2021), and Khera (2021) highlight this need for more documentation and transparency of decisions related to specific issues in research ethics. Each suggests specific topics that require attention, summarized in Table 2.

Table 2: Calls for more documentation and transparency regarding specific ethical issues			
Topics and themes	Asiedu, Karlan, et al (2021)	Evans (2021)	Khera (2021)
Understand how state of equipoise and scarcity inform research methods	X	X	
Document role of researchers with respect to policy/intervention implementation	X		
Assess, monitor, and mitigate potential harms from intervention	X	X	X
Assess, monitor, and mitigate potential harms from survey methods	X	X	X
Understand and document financial and reputational conflicts of interest	X		X
Understand and document intellectual freedom to publish full results/research findings	X		
Engage with local communities (through collaboration with researchers who understand local context; feedback to participants and their communities; assessment of vulnerabilities)	X	X	X
Understand and mitigate foreseeable misuse of results/research findings by decision-makers	X		
Seek approval from Institutional Review Committee (IRB)	X	X	

Document and understand roles and responsibilities and power dynamics across researchers, funders, journals, policymakers that can affect credibility and bias			X
Understand legality of policy/intervention studied			X
Understand true costs of the research, including adequate payment for staff and participants		X	X
Assess all potential research design and methods options and assess against cost		X	X
Give careful attention to informed consent		X	X
Give appropriate credit to wide range of stakeholders		X	

Researcher teams face difficulties navigating ethical decisions throughout the research life cycle, including individual constraints and tensions between TREE best practices.

Starting at research funding and design stages, research teams face a wide range of ethical dilemmas. However, researchers can face several barriers to navigating ethical issues in their work, including: (i) absence of awareness and belief that ethical concerns aren't applicable to their work; (ii) lack of awareness of the connection of their research to broader social issues; (iii) overconfidence at ability to manage ethical concerns as they arise; and (iv) expectation that reflections on ethical issues – internally or with external experts – create burden that slows down their already demanding research process (McCormick, Boyce, et al 2012; Devereaux 2014). With the introduction of TREE best practices, there are new dilemmas. For example, it may not be feasible to de-identify the data that underlies final analysis in a manner that adheres to promises of confidentiality and allows for public access to that data for transparency and reproducibility (Sturdy, Burch, et al 2017). When navigating how to be more transparent and reproducible while also considering ethical issues, researchers have more questions, including: *Which best practices improve credibility and quality even when transparency is not feasible? When is transparency preferred versus required? How is data sharing for transparency and reproducibility balanced with ethical practices, such as promises of confidentiality? Who is accountable for the quality of TREE practices?* While more documentation and transparency of this decision-making process is important, it is insufficient without defined standards, structured review, and timely feedback to research teams.

Research teams rely on Institutional Review Boards (IRBs) to define and guide ethical principles, but there are gaps. Research activities involving human subjects – particularly when the collection of personally identifiable information (PII) and sensitive data is required – often pass-through IRB review to determine if the research is designed and implemented in alignment with research regulations typically governed by foundational principles of ethics - beneficence, justice, and respect for persons (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979).

- **Respect for persons.** *“Respect for persons incorporates at least two ethical convictions: first, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy are entitled to protection. The principle of respect for persons thus divides into two separate moral requirements: the requirement to acknowledge autonomy and the requirement to protect those with diminished autonomy.”* This principle guides the informed consent process, including information provided, ensuring comprehension of information provided, and allowing for the voluntary nature of participating in the study. Alignment with this principle requires an understanding of various forms of research participant vulnerabilities (NBAC 2001 and Appendix 1).
- **Beneficence.** *“Persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by making efforts to secure their well-being. Such treatment falls under the principle of beneficence. In this document, beneficence is understood in a stronger sense, as an obligation. Two general rules have been formulated as complementary expressions of beneficent actions in this sense: (1) do not harm*

and (2) maximize possible benefits and minimize possible harms.” This principle guides how research teams determine there is balance between benefits and harms.

- **Justice.** Justice informs why “*the selection of research subjects needs to be scrutinized in order to determine whether some classes (e.g., welfare patients, particular racial and ethnic minorities, or persons confined to institutions) are being systematically selected simply because of their easy availability, their compromised position, or their manipulability, rather than for reasons directly related to the problem being studied. Finally, whenever research supported by public funds leads to the development of therapeutic devices and procedures, justice demands both that these not provide advantages only to those who can afford them and that such research should not unduly involve persons from groups unlikely to be among the beneficiaries of subsequent applications of the research.*” The principle of justice is about fairness in the distribution of benefits and burdens introduced by the research activity (MacKay, 2020).

The IRB review consists of reviewing a research protocol, the data collection instruments, and the informed consent(s) and assent(s) (as relevant) to assess alignment with these principles and regulations.

There are at least four forms of IRBs that research teams may engage with on these reviews: (i) *Academic*– Researchers based within a university are often required to submit their research materials for IRB review to their Academic IRB; (ii) *Institutional* – Researchers based within research organizations may have an internal IRB that has been established by the research institution; (iii) *Country-based*– An IRB, or multiple IRBs, may be established at the country, state, or other level. Research teams may be required to submit their research materials for IRB review to the Country-based IRB⁷ regardless of if they are required to go through an Academic IRB or not; and (iv) *Independent, for-hire* – There are many independent IRBs available for-hire by research teams, particularly when there is no Academic, Institutional, or Country-based IRB for the relevant research.

This is further complicated by sub-categories of qualifications of IRBs (such as United States Health and Human Services (HHS) Registered⁸ and Association for the Accreditation of Human Research Protection Programs (AAHRPP) Full and Qualified Accreditation⁹) and the multiple levels of IRB review (full, expedited¹⁰, and exempt¹¹). When research teams have complex compositions with research team members based in academic and non-academic settings, it is not always clear which and how many IRBs must review the protocols. This diverse landscape of IRBs in terms of type, level, and qualifications produces variation in terms of the level and quality of review conducted on any given research activity.

Research teams working in lower-middle income countries (LMICs) experience additional variation during IRB reviews (Grady 2015). For example, IRBs governed by United States regulations (Federal Policy for Protection of Human Subjects, Revised Common Rule 2018¹²) can classify certain research activities as exempt or expedited review, despite direct engagement with human subjects and the collection of PII and sensitive data. In addition, IRBs in high and middle income countries may have IRB members with limited experience in the LMIC context where the research activity takes place and may not understand local constraints and context (Gilman and Garcia, 2004) which can result in a full IRB review that does not adequately consider vulnerabilities and liability of harm of human subjects or the research team. On the other hand,

⁷ <https://www.hhs.gov/ohrp/international/compilation-human-research-standards/index.html>

⁸ <https://ohrp.cit.nih.gov/search/search.aspx?styp=bsc>

⁹ <https://www.aahrpp.org/learn/find-an-accredited-organization>

¹⁰ <https://www.hhs.gov/ohrp/regulations-and-policy/guidance/categories-of-research-expedited-review-procedure-1998/index.html>

¹¹ <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/common-rule-subpart-a-46104/index.html>

¹² <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html>

Country-based IRBs may require more resources and more capacity building to respond to the complexity and increase in volume of review (Davies, 2020).

There are efforts to strengthen the IRB process internationally (for example, see Kruger 2014). However, even as IRB processes improve, IRB review may not consider issues beyond research regulations, including TREE issues raised in Table 1 and Table 2.

There are lessons from clinical research ethics that can inform ethical requirements for social science research. Despite the establishment of high-level foundational research ethics principles that guide research regulations and IRB review, research teams often look for more guidance on how to implement practices that align with these principles. The clinical research ethics literature offers lessons that are applicable and adaptable to social science research. As presented in MacKay (forthcoming (a)), specific ethical requirements to consider for public policy research include:

- **Social value.** The aim of public policy research is to determine either the efficacy or the effectiveness of an identified policy or intervention to improve targeted outcomes. Targeted outcomes are defined as the “*specifications of the type and amount of goods or services governments have a duty to provide, and the outcomes they have a duty to realize*” (MacKay, 2020). Therefore, the research is expected to produce social value when it is reasonable to expect the research to produce generalizable knowledge relevant to the development of cost-effective policies or interventions for the realization of target outcomes.
- **Favorable risk-benefit ratio.** In line with the beneficence principle, research teams must weigh the expected benefits against risks (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). Risks should be viewed as possible harms to individuals, or other units (households, communities, etc.) as a result of the research. The severity of a risk should be defined as the combination of the potential impact of the harm on an individual (or other unit) and the probability that the harm could occur. As those summarized in Phillips (2021), harms can include:
 - Intervention harm - Interventions may create unintended consequences where targeted outcomes decline rather than improve. Such harms can extend to bystanders as well (family members, community members), requiring an assessment of the potential aggregate harm from the intervention (Desposato, forthcoming).
 - Data harm – Inappropriate and improper data management that fails to protect the confidentiality of research participant may cause harm. Unauthorized disclosure of sensitive data – such as income, bank account information, savings, assets, health status – can result in targeting of research participants for exploitation, violence, or other harm.
 - Exclusion harm – When a study requires withholding treatment from a control group, exclusion harm can become a concern. Those excluded from the treatment group may experience harm from not participating in a beneficial intervention, particularly if they may have received the intervention in the absence of the study. There is also the potential for control group participants to experience jealousy or outrage because they are excluded from the study, which may affect community cohesion and security.
 - Outcome harm. These harms can result when research findings are used – or misused – to enact programs or policies that make people worse off or result in stigmatizing groups of people.

With these points in mind, a favorable risk-benefit ratio considers the following three conditions (adapted from Emanuel, Wendler, and Grady 2000): (i) risks to participants and affected bystanders are minimized; (ii) benefits to participants and society are enhanced; and (iii) the potential benefits to participants and society outweigh the risks to participants and affected bystanders.

- **Standard of care.** The standard of care is the level of care or goods to which research participants are entitled. Research participants should not therefore be subject to an intervention that is expected to be inferior to the standard of care intervention. MacKay (2020) argues that the standard of care for any study is the intervention to which participants have a

claim to be subject. Since governments have a duty to implement the most effective policies they can to realize target outcomes with the resources they have at their disposal, MacKay (2020) argues that the standard of care is the best proven morally and practically attainable and sustainable policy (BPA policy). The standard of care requirement thus requires that government agencies not assign people to interventions reasonably expected to be inferior to the BPA policy.

- **Fair randomization.** Research methods that depend on withholding the intervention from a control group for a selected time (through randomization or another treatment assignment method) require scrutiny. Two conditions inform when it may be ethically appropriate to withhold treatment from a group of equally eligible participants: equipoise and scarcity (Friedman (2014); MacKay (2018); MacKay (2020); Asiedu, Karlan, et al (2021); Evans (2021)). The first condition is equipoise – uncertainty regarding the efficacy and/or effectiveness of the proposed intervention or policy on targeted outcomes. The second condition is scarcity – when there are insufficient financial, administrative, or other resources to provide the treatment to all equally eligible participants. Under scarcity, the following conditions should be met when using randomization to assign treatment and control group: (i) the people subject to randomization all have equally strong claims to the intervention; (ii) the people subject to randomization have stronger claims to the intervention than those not subject to randomization; and (iii) no person’s claim is left unmet longer than is necessary due to legitimate scarcity.
- **Fair subject selection.** Fair subject selection is a cornerstone of the justice principle. Researchers must recruit and select participants in a way that fairly distributes benefits and burdens (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979; MacKay and Saylor 2020). As presented in MacKay and Saylor (2020), fair subject selection requires consideration of:
 - Fair *inclusion* – Selection of research participants must be sufficiently inclusive to ensure the research fairly benefits members of society. The population enrolled in research should reflect the diversity of the population which is likely to be subject to any subsequent intervention.
 - Fair *opportunity* – When participation in the study is ex-ante net beneficial (such as a cash transfer scheme), prospective participants must be granted a fair opportunity to participate in research.
 - Fair *burden sharing* – When it is unknown that participation in the study is net beneficial, the burdens of participation in public policy research must be shared fairly. Burdens include opportunity cost for participating in long survey, biomarker data collection, and other potential harms flagged above such as feelings of exploitation, re-traumatization experienced as study participants. This is particularly relevant to control group participants who bear the burden of the research but not the potential benefit of the intervention.

Since there may be tensions across these factors, MacKay (forthcoming) provides additional guidance:

- Design inclusion and exclusion criteria to answer the scientific question in a way that fairly benefits members of society (fair inclusion).
- Among potential participants meeting inclusion criteria and not meeting exclusion criteria, set and meet goals for enrollment of potential participants to ensure research fairly benefits members of society (fair inclusion).
- Ensure all prospective participants satisfying inclusion and exclusion criteria have a fair opportunity to participate, and refrain from targeting disadvantaged prospective participants for research which is ex-ante net burdensome (fair opportunity and fair burden sharing).
- **Informed consent.** Informed consent is a cornerstone of the respect for persons principle (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979) and is a standard requirement for ethical conduct of research to respect participants’ autonomy and provide them with sufficient information on the risks and benefits

of participating in the research. However, MacKay and Chakrabarti (2019) presents circumstances where informed consent may be waived in ethical research. Such circumstances require three conditions: (i) when there is a strong justification for not obtaining participants' informed consent; (ii) infringement of participants' autonomy rights is minor; and (iii) the expected social value of the research is sufficient to outweigh the infringement of participants' rights (Gelinas, Wertheimer, and Miller 2016). Any decisions by a research team to waive informed consent should be carefully assessed with these in mind, as well as with community engagement.

- **Community engagement.** Community engagement is critical for research design, implementation, and dissemination. It involves researchers collaborating with various stakeholders – participants, community members, policymakers, and/or local organizations – to design and/or implement various aspects of the research (Anderson and Spelley forthcoming). Community engagement promotes social value by ensuring research is designed to answer relevant questions for stakeholders and examine interventions that are relevant to improving the lives of participants. It supports assessing the risk-benefit ratio by providing clear and contextual understanding of vulnerabilities and potential harms that can result from the research. In addition, integrating community engagement in the research life cycle acknowledges that people have rights to participate in and influence decision-making processes that affect their interests. MacKay (forthcoming(a)) offers two conditions for this requirement: (1) provide materially relevant information to affected individuals and provide them with an opportunity for meaningful feedback; and where community rights are in play, (2) secure consent from the authorized community representatives regarding the design and conduct of the research. However, as discussed below, a challenge is understanding who are the appropriate representatives and any additional risks that may arise from insufficient or inappropriate community engagement (i.e., such as only meeting with village leaders or certain groups that may represent only a certain profile of the participant population).

This literature provides a solid foundation for supporting the social science research community in establishing common language and standards for ethical requirements.

There are additional review mechanisms in the clinical research setting that complement IRB review and can inform similar social science review mechanisms. For clinical research that is determined to be minimal risk, IRB review may be sufficient. However, for more complicated and potentially higher risk studies, additional review mechanisms exist, including (i) Study Monitoring Committees, (ii) Data and Security Monitoring Committees, and (iii) Independent Medical Monitors. The aim across committees is similar – create independent bodies that are responsible for timely and continuous assessment of the risks facing the study. These Committees can be within the same research institution as the original research team but must be independent of the research team itself. Such committees are important to the independent review and risk management of study from design to completion of the study. These committees are in addition and complementary to IRBs¹³.

Just as the open science movement has moved to change the norms, incentives, and institutions within science to realign with original scientific values of communism, universalism, disinterestedness, and organized skepticism (Merton 1942), there is a need to challenge the norms, incentives, and institutions of research ethics. As stated in Devereaux (2014): “A broader conception of ethics is needed that would include reflection on these matters and encourage scientists to consider not only the ethical and social consequences of their work, but also the ways in which lab culture and institutional reward structures may themselves undermine objectivity and rigor.” Researchers and research organizations have a stronger role to play in

¹³ For reference, see NIH Policy for Data and Safety Monitoring (<https://grants.nih.gov/grants/guide/notice-files/not98-084.html>), FDA Policy for Clinical Trial Data Monitoring (<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/establishment-and-operation-clinical-trial-data-monitoring-committees>) and NINDS Guidelines for Monitoring Clinical Trials <https://www.ninds.nih.gov/Funding/Apply-Funding/Application-Support-Library/NINDS-Guidelines-Data-and-Safety-Monitoring#Resp%20DSMB>

operationalizing organized skepticism through the intersecting lens of transparency, reproducibility, and ethics. Afterall, researchers are likely to know the ethical and methodological challenges faced in design, implementation, and dissemination decisions better than other actors. Rather than outsourcing ethical judgement to IRBs (Zechmeister, 2013; Desposato, 2014) or relying on transparency to outsource quality assurance to the broader scientific community, we propose providing researchers with the resources to document, monitor, and mitigate risks to the ethical conduct of research through a broader conception of ethical research.

Table 3: Addressing threats to research credibility and ethics with TREE Review Framework

Problems we want to solve	Researcher degrees of freedom, p-hacking,	Publication Bias and the File Drawer	Failure to reproduce analysis	Improved oversight of ethical risk assessment and mitigation	Inconsistent communication on and adherence to promises of confidentiality
TREE practices to address challenges	Study registration; Pre-analysis Plans; Standardized Reporting Templates		Push-button replication; Data sharing	Training in protection of human subjects; adhering to IRB requirements; establishing ethical requirements, conducting timely, independent review of ethical decision-making	Informed consent; responsible data stewardship; Data de-identification

The Transparent, Reproducible, and Ethical Evidence (TREE) Review Framework proposed here is a means by which the International Initiative for Impact Evaluation (3ie) operationalizes this system to address gaps in existing IRB review and establish a broader conception of what ethical conduct of research means. The following summarizes the objectives and guiding principles, requirements, and practices of the framework, then presents the pilot test of the process and tool within the 3ie portfolio, concluding with a set of lessons learned.

Objectives

The goal of policy-relevant research is to resource research practices that produce credible, unbiased evidence that is meaningful for decision-makers and conducted in an ethical manner. To achieve this goal, the proposed TREE Review Framework aims to achieve three objectives: (i) establish ethical standards, (ii) better integrate TREE best practices into research workflow; and (iii) establish timely, independent review of risks facing the research team to meet our goal.

As with clinical research monitoring mechanisms, the TREE Review Framework is not a replacement for IRB review and clearance, nor is it a substitute for any local IRB and/or ethical review requirements. IRB reviews remain required for research activities to pass through for alignment with ethical principles and regulations. In addition, IRB review remains a requirement for many research funders and journals.

Ethical standards

These requirements extend from the foundational research ethics principles of beneficence, respect for persons, and justice and build upon the literature summarized above. Each requirement is a key factor for aligning with ethical principles and supporting the goal of research activities: to produce credible, unbiased evidence that is meaningful for decision-makers and conducted in an ethical manner.

Use transparency as a tool

It is important to remember that transparency is not the goal or outcome – transparency is a tool for supporting research practices that to produce credible, unbiased evidence that is meaningful for decision-makers and conducted in an ethical manner. Best practices in research transparency and reproducibility - pre-specification, standardized reporting, de-identified data and code

sharing, and push-button replication are well-defined. Even if full transparency is not feasible, these best practices facilitate achievement of the goal.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Research team and field staff's training in protection of human subjects.* The preference is for all data handlers and those who directly engage with human subjects to be trained in foundational protection of human subjects. For example, 3ie requires all staff and consultants to complete training through [Protecting Human Research Participants | PHRP Training](#).
2. *Study registration.* The preference is for the study design elements to be registered prior to baseline data collection. When research teams also develop a time-stamped pre-analysis plan, it is acceptable for the study to be registered after data collection starts, but preferably before final publication. Examples include: [RIDIE](#), [ClinicalTrials.gov](#), [OSF](#), [AEA RCT Registry](#) and [EGAP](#).
3. *Pre-specification with a thorough a pre-analysis plan.* The preference is for the research team to prepare and time-stamp a pre-analysis plan prior to receiving any data. The PAP should describe the hypotheses to be tested and specifications, as well as the sequence of all planned statistical analyses. It should clearly describe primary and secondary outcomes, covariates, and any planned subpopulation analysis. It should also be clear who leads and documents any necessary revisions to the pre-analysis plan, and if/when the PAP will be published.
4. *Standardized reporting templates.* The preference is to establish agreement on the standardized reporting templates for analysis reporting. These templates should be consistent with best practice reporting standards published by the [Equator Network](#).
5. *Alignment between final reporting and pre-specification.* The preference is to ensure the results of all pre-specified models are presented. Exploratory analysis should be distinguished from analysis of pre-specified models. Interpretation of results should clearly articulate findings based on pre-specified models compared to those based on exploratory analysis.
6. *IRB procedures.* The preference is to document all potentially required IRBs – academic, institutional, country, and independent – based on research team and country context. The research team should document the IRB(s) review level – full, expedited, exempt – as well as any issues raised during IRB review.
7. *Responsible data workflow and management.* The preference is to establish a Data Management Plan that clearly articulates which data handlers may access which data (identified vs. de-identified) in accordance with the informed consent. The DMP should define how data will be collected, stored, transferred, and shared in a manner that aligns with promises of confidentiality in the informed consent.
8. *Reproducible workflow.* The preference is to establish internal reproducible workflow procedures which coordinate data de-identification with independent push-button replication (PBR) or computational reproducibility checks (i.e. ability to run the code on the data and produce the same results presented in the paper). The extent to which results are reproducible outside the study team will depend on the feasibility of conducting PBR on de-identified data that can be shared publicly or through restricted access.

Maximize social value and meaningful use

Maximizing social value and meaningful use will require early and continuous engagement with the research stakeholders – implementation partners, policymakers, other decision-makers, and research participants. Best practices in transparency and reproducibility also increase the social value and use of the research. Pre-specification, standardized reporting, responsible data

management, and reproducible workflow are all intended to reduce research waste and increase social value and useability of research.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Problem diagnostic and theory of change.* The preference is for a shared and data-driven understanding of the problem targeted by the intervention. It is recommended there is a clear theory of change that connects how the proposed intervention will address the identified problem(s). It is recommended the research team clearly documents if the research will study the full theory of change or only certain components of the theory of change.
2. *Motivation.* The preference is to have a shared understanding of the motivation of the research to inform and shape the research questions. It is recommended that the research team defines if the motivation is primarily learning – to fill an evidence gap – or accountability – working with a government partner to document achieved results – or both.
3. *Evidence consumers.* The preference is to establish as early as feasible who the targeted evidence consumers are and ensure they are included in various stages of research design, implementation, and dissemination to facilitate social value and use. This group may change over time and requires monitoring to determine what effects change should have – or not – on the study.
4. *Incentives for misuse.* There is the possibility that results or partial results are deliberately misinterpreted or misused by researcher stakeholders. While there may be little research team members can do ex-ante to identify and mitigate this risk, this should be monitored and documented.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, value and prioritize community engagement, use fair methods, and ensure fair subject selection.

Balance power and align incentives

Power dynamics between various stakeholders involved in international social science research – particularly in the context of specific vulnerabilities (NBAC 2001) - can create potentially misaligned incentives that impact the ability of the research team to deliver credible, unbiased findings and the ability and/or motivation of policymakers and implementers to use the evidence for unbiased decision-making. For example, when the intervention and the study have the same funder, are there potential incentives in place that affect how willing the funder is to accept negative or null findings regarding the intervention they funded? What governance systems are established to ensure these potential incentives are documented and mitigated? Best practices in transparency and reproducibility can support research teams, particularly use of pre-specification and agreement on standardized reporting that align with pre-specified analysis plans.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Research stakeholders' power dynamics and relationships.* It is recommended that the research team documents power dynamics that may affect various stakeholders: between research team and research participants; within the research teams between field staff and headquarters staff; between funders and research team; etc.
2. *Vulnerabilities.* It is recommended that the research team examine any vulnerabilities that can affect autonomy of research participants, as well as staff. Rather than define specific groups as vulnerable, such as children, pregnant women, Appendix 1 provides an overview of types of vulnerabilities and their definitions.

3. *Researcher role in intervention.* The preference is to clearly document the research team's role in intervention design and implementation decisions. Is the research team completely independent? Or is the team partially or fully responsible for intervention design and implementation?
4. *Restrictions on reporting.* The preference is for clarity and agreement from the beginning on any process required for the research team to publish findings from pre-specified analysis plans. There should be a clear understanding of any review processes and the autonomy of the research team to publish – or not – complete results.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, maximize social value and meaningful use, value and prioritize community engagement, ensure informed consent and protection of confidentiality.

Preserve standard of care

All research stakeholders – research teams, funders, policymakers, implementing partners – should have a shared, documented understanding of the current standard(s) of care available to the participant population. This assessment should ensure the intervention studied does not introduce a standard of care that is lower than what study participants are entitled prior to the intervention.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Standard of care.* The preference is for the research team to assess and document the standard of care and determine how the proposed intervention may affect the standard of care for eligible participants, considering treatment and control groups. Research teams should note there is debate on what should define the standard of care – the de facto standard of care (status quo) and de jure standard of care (global best standards) – and may want to provide their definition of the standard of care used in their assessment. It is recommended that research teams focus research on interventions that aim to preserve or improve the standard of care, and not study or introduce an intervention that reduces the standard of care. Any case where the study examines an intervention that is expected to be inferior to the standard of care (reducing benefits), should be done through community engagement, with full and transparent documentation.
2. *Deceit.* Sometimes deceit may result in evidence that has social value and meaningful use. The preference is for the research team to assess and document any deceit required in the intervention and research design and implementation. This assessment should be done through community engagement, with full and transparent documentation.
3. *Legality.* Not all laws are moral or the best policy for targeted outcomes. Studying illegal behavior or activity may bring social value and meaningful use that justifies illegal activities. The preference is for the research team to clearly assess and document the legality of the intervention they are studying and their study methods, particularly when the research team is partially or fully responsible for intervention design and implementation. This assessment should be done through community engagement, with full and transparent documentation.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, value and prioritize community engagement, use of fair methods, ensure fair subject selection, ensure favorable risk-benefit ratio and accountability.

Value and prioritize community engagement

Research teams are often outsiders to the research participant populations. Engaging with the research participant community is a critical cornerstone in the ethical conduct of research, and supports the research team to design, implement, and disseminate evidence that is credible and meaningful.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Research participant representation.* The preference is for the research team to establish a mechanism by which the research participant population(s) are represented at study design, implementation, and dissemination milestones. This can include the development and prioritization of research questions, pilot testing of instruments and informed consent, and sharing and interpretation of analysis findings.
2. *Research participant feedback loop.* The preference is to establish a feedback loop for two types of information. First, if data is collected that informs the health, safety, and welfare of the participants (individual, household, community, etc.), this information should be reported and shared as soon as possible. For example, if the research team conducts anemia biomarker testing or water quality testing in a home or communal water source, the results should be shared immediately with the participant population. Second, the research team should assess how it can share the findings of the research with the research participant population to foster trust and shared respect for the value of their contributions to the research¹⁴.

Alignment with this requirement also depends on alignment with ensure fair subject selection and ensure appropriate informed consent and protection of confidentiality.

Use fair methods

Critics of certain social science methods pose the question - is it ethical to withhold an intervention from a population (such as through random assignment) to assess the effectiveness of that intervention? Others would pose a counterquestion - is it ethical to fund interventions and policies worth millions of dollars and not know if they are as effective as another intervention or policy? Research teams are tasked with demonstrating they have carefully thought through the ethics of their research methods to strike the right balance between these two sides.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Alignment of methods with state of project – efficacy vs. effectiveness.* The research team must understand the state of the intervention design and implementation to determine if the research is an efficacy study – establishing proof-of-concept – or effectiveness study – understanding impacts of intervention in real-world conditions. This assessment should inform the research team’s application of specific research methods, such as answering formative vs. summative research questions, establishing a counterfactual, examining implementation fidelity, etc.
2. *Alignment of methods with intervention details.* The research team must assess intervention design and implementation details that inform selection and prioritization of research methods. This requires understanding the unit of implementation, selection criteria, intervention outputs, theory of change, etc. The evaluation methodology, particularly when

¹⁴ For example, refer to Value and Validation - <https://medium.com/busara-center-blog/value-and-validation-113750e7c0ad>

methodology requires assignment of groups to treatment and control, should be informed by and align with the intervention design and implementation details.

3. *State of equipoise*. The research team should assess if there is uncertainty regarding the benefits of one intervention or policy compared to another. This uncertainty may be a question of *efficacy* – extent to which an intervention does more good than harm under ideal circumstances compared to another – or a question of *effectiveness* – whether an intervention does more good than harm when provided under usual circumstances compared to another. When there is uncertainty, this can inform the social value and fairness of withholding treatment from a control group to study the intervention’s impacts.
4. *State of scarcity*. The research team should assess if there is scarcity of resources – financial, administrative, bureaucratic, etc. – that prevent all eligible participants from receiving the intervention at the same time. The research team should also assess if there are any specific characteristics – gender, poverty level – that prioritize treatment assignment within the eligible population. Any study methods should adhere and respect the intervention selection details. When there is scarcity, this can inform the social value and fairness of withholding treatment from a control group to study the intervention’s impacts.
5. *Data quality*. The research team should understand and document all methods required to assess, produce, and monitor the quality of data required for the research. When collecting primary data – such as household surveys – this requires careful consideration of the data collection firm budget, timing, field staff composition, and oversight. When extracting secondary data – such as existing administrative data – this requires careful assessment of the quality of the data, its distributions, and likelihood of any bias or other risk factors that affect the quality of the data and subsequent analysis.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, preserve standard of care, value and prioritize community engagement, ensure fair subject selection, ensure appropriate informed consent and protection of confidentiality.

Ensure fair treatment of participants

Fair subject selection aims for justice in the distribution of the burden and benefits of the research. In addition, it aims to maximize social value by producing appropriately generalizable knowledge on the impacts of the intervention on the eligible participant population.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Alignment of research participant selection with intervention selection*. It is recommended the research team aligns inclusion and exclusion criteria for the research with inclusion and exclusion criteria for the intervention. Any differences between the two should be well documented and explained in terms of any risk of bias introduced by the differences in the research participants and the intervention participants.
2. *Fairness to control group*. When the research methods rely on a control group, the research team should examine the burdens of the research placed on the control group and compare them to the expected benefits. It is recommended that research stakeholders examine how the control group may be prioritized to receive the intervention if (i) scarcity is resolved and (ii) the research demonstrates the intervention leads to superior outcomes.
3. *Participant payment*. The research team should examine the time burden placed on research participants (some studies require 2+ hours of each individual participant’s time) and opportunity costs to participate. Relying on community engagement, local IRBs, the research team should assess appropriate participant payments in the form of reimbursement, compensation, or incentives. Working within the local context and culture, the research team

can also examine the preference to cash vs. in-kind payments. This assessment should be done to understand risks of undue influence over participants to participate when they would rather not vs. risks of exploitation of often vulnerable populations.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, preserve standard of care, value and prioritize community engagement, ensure appropriate informed consent and protection of confidentiality.

Ensure appropriate informed consent and protection of confidentiality

Informed consent is not just a document that needs to be read and signed (or waived). Informed consent is a process that requires meaningful engagement between the research team (often represented by interviewers) and the research participant to discuss the objectives of the study, what will be collected, when, why, and how. It is also the opportunity to provide adequate information to the participant regarding how their data will be protected and who will have access to what data and for what purposes. Building trust at this stage around data use and protection and being clear about who and how the data may be shared is fundamental to using transparency as a tool in an ethical manner.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Use of secondary data.* An important way that research teams can reduce the costs of research is leveraging existing data sources. When this is feasible, it is recommended the research team think carefully through how the original data was collected, what informed consent took place at that stage, if re-consent is required for the new analytic purposes, etc. Additionally, the research team should determine how the link between this secondary data and the full research data may affect de-identification efforts and re-identification risk. Even if the research data is de-identified by the study team, does the existence of this secondary data increase risk that others who access it may re-identify the research data?
2. *Consent and assent requirements and comprehension.* The research team must assess and document all research participants for whom data will be collected and determine if consent and assent is required. For studies that require data collection from a male head of household, female household member, and biomarkers on children, there will be multiple consent and assent forms required. Additionally, the research team should work through community engagement to determine when community consent (such as from a village leader) is necessary and assess any implications this can have on vulnerable populations (if the village leader consents, will households feel obligated to participate?).
3. *Promises of confidentiality.* The research team may not need to promise confidentiality if the data collected is observable or known or otherwise not sensitive, confidential data. The data needs should be assessed to determine if promises of confidentiality are required. The promise of confidentiality should clarify exactly who across the research stakeholders – data collection team, research team, implementing partners, others – will have access to identifiable data and who will have access to de-identified data (if relevant).
4. *Disclosure risk and mitigation in data collection.* In many contexts, the highest risk – in terms of probability and impact – from a disclosure is in the field and local area where the data collection takes place. If unauthorized disclosure occurs in the local setting, there is a higher probability that harm may come to the participants based on data disclosed, such as income, assets, savings, personal details regarding relationships and other sensitive data.
5. *Disclosure risk and mitigation in data storage, transfer, and sharing.* There is an increasing amount of data available from diverse sources that can make research more effective and

efficient. However, the existence and use of these interrelated data can increase risk of misuse or risk of loss of confidentiality to research participants. The more data connects and links to any individual or household or other populations, the higher the potential re-identification risk through indirect identifiers (Sweeny 2000) though this can be mitigated through disclosure limitation and understanding of linkage documentation (Barth-Jones 2012). In the context of growing demand for transparency and open data, there is a need to ensure balance between data sharing and ensuring adherence to promises of confidentiality. For these reasons, the research team should carefully assess and document how it will minimize the risk of unauthorized disclosure of PII and sensitive data through its secure data storage, transfer, and sharing practices.

Alignment with this requirement also depends on alignment with other requirements: use transparency as a tool, maximize social value and meaningful use, balance power and align incentives, preserve standard of care, and value and prioritize community engagement.

Ensure favorable risk-benefit ratio and accountability

Implementing interventions intended to change behavior and improve social and economic outcomes is risky. One motivation for research is to ensure adverse outcomes of these interventions are measured, reported, and mitigated in future efforts. However, in addition to the risks introduced by the interventions, social science research itself introduces risks for many stakeholders – staff, participants, and bystanders. The issue of who is accountable for when risks materialize also requires assessment and scrutiny.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Research staff risk-benefit ratio and accountability.*
2. *Field staff risk-benefit ratio and accountability.*
3. *Research participant risk-benefit ratio and accountability.*
4. *Bystander risk-benefit ratio and accountability.*

Table X presents an overview of the types of risks, potential harms, and potential risk mitigation strategies for striking this favorable risk-benefit ratio.

Table X: Summary of Risks, Potential harms/adverse events, and risk mitigation efforts to consider for a favorable risk-benefit ratio		
Risk	Example(s) of Harm/Adverse events	Risk Mitigation Efforts
<p>1. Intervention design and/or implementation violates ethical principles (beneficence, respect for persons and/or justice)</p> <p>NOTE: The risk can extend to PERCEPTION. When others PERCEIVE a violation of ethical principles, there remains a high risk to researchers and others for harms described.</p>	<ul style="list-style-type: none"> • Unethical Intervention design and implementation can lead to physical, economic, emotional, social, or legal harm(s) for intervention participants, research participants, bystanders • Studying unethical interventions can pose reputational harm to researchers • Studying unethical interventions can lead to reduced funding for research • Researchers become too risk averse to study controversial interventions and these interventions are allowed to continue without evidence to inform course corrections which result in harmful policies that affect the physical, economic, emotional, social, or legal status of populations of interest 	<ul style="list-style-type: none"> • Carefully examine the standard of care and how the intervention(s) studied may affect the standard of care in the 'preserve standard of care' section • Carefully examine the social value of the study and whether there is a favorable risk-benefit ratio • Ensure all key research team members are trained in foundational principles of protection of human subjects • Ensure review of research protocol and informed consent by appropriate Institutional Review Board(s) • Complete timely and continuous ethical review using resources like TREE Review questionnaire

Table X: Summary of Risks, Potential harms/adverse events, and risk mitigation efforts to consider for a favorable risk-benefit ratio		
Risk	Example(s) of Harm/Adverse events	Risk Mitigation Efforts
	<p>Example: A study in Nairobi examining effects of various mechanisms to increase payments for water supply, including cutting off water access, resulted in various discussions around the ethics of the study (see discussion on ethical considerations here).</p>	<ul style="list-style-type: none"> Establish mechanisms for participants, bystanders, staff to report harms/adverse events Be accepting of risk – some interventions may be unethical, and how interventions are improved or removed may be research demonstrating the adverse effects. However, this requires additional emphasis on the need to use transparency as a tool for how the researcher navigates ethical concerns and abides by research ethical principles.
<p>2. Research design and/or implementation violates ethical principles (beneficence, respect for persons and/or justice)</p> <p>NOTE: The risk can extend to PERCEPTION. When others PERCEIVE a violation of ethical principles, there remains a high risk to researchers and others for harms described.</p>	<ul style="list-style-type: none"> Unethical research design and implementation (failure to provide sufficient information for informed consent, poorly designed or executed data collection) can lead to physical, economic, emotional, social, or legal harm(s) for research participants as result of study itself Findings from poor quality or biased research design can inform programs/policies that affect the physical, economic, emotional, social, or legal aspect of population of interest Poor quality or biased research design and implementation can pose reputational harm to researchers Unethical research practices can lead to reduced funding for research <p>Example: A study in Montana on the effect of different communications flyers to increase voter turnout was likely illegal and failed to consider ethics of impacts on community.</p>	<ul style="list-style-type: none"> Ensure all key research team members are trained in foundational principles of protection of human subjects Ensure review of research protocol and informed consent by appropriate Institutional Review Board(s) Complete timely and continuous ethical review using resources like TREE Review questionnaire Establish mechanisms for participants, bystanders, staff to report harms/adverse events Be prepared to end a study if it is determined the study cannot be conducted in a way that aligns with ethical principles
<p>3. Research studies a solution that does not align with the context/problem</p>	<ul style="list-style-type: none"> Studying interventions that don't align with problems can bias the interpretation of results of the study. For example, the study can determine the intervention 'didn't work' when it was not aligned with context/original problems. This can affect future programming of policies or interventions that affect the economic and social status of research participant and bystander populations The potential risks and actual costs of the research may outweigh the benefits, violating 'beneficence' principle Research waste results in less funding for research 	<ul style="list-style-type: none"> Ensure there is a strong, data-driven understanding of problem the intervention will address in the 'maximize social value' assessment
<p>4. Results of study are not useful or used by decision-makers</p>	<ul style="list-style-type: none"> The potential risks and actual costs of the research outweigh the benefits, violating 'beneficence' principle Research waste results in less funding for research 	<ul style="list-style-type: none"> Ensure motivation of study and evidence consumers are identified early and engaged throughout research life cycle through 'maximize social value' assessment

Table X: Summary of Risks, Potential harms/adverse events, and risk mitigation efforts to consider for a favorable risk-benefit ratio		
Risk	Example(s) of Harm/Adverse events	Risk Mitigation Efforts
		<ul style="list-style-type: none"> Ensure quality of research through ensuring fair methods and fair treatment of participants
5. Research team and/or research funders have incentives to prioritize positive results, downplay null or negative results	<ul style="list-style-type: none"> Selective reporting and/or specification searching can result in biased findings and biased evidence base Bias can affect future programming of policies or interventions that affect the economic and social status of data provider populations Selective reporting and/or specification searching can pose reputational harm to researchers Research waste results in less funding for research 	<ul style="list-style-type: none"> Define research questions and outcomes through pre-specification in registration and pre-analysis plans Assess power dynamics and incentives across stakeholders Use transparency as a tool to mitigate misaligned incentives and allow for assessment of credibility of findings
6. Research participants experience specific vulnerabilities and/or power dynamics that affect their sense of autonomy to participate or refuse to participate in the research	<ul style="list-style-type: none"> Research participants feel unable to be truthful during data collection, leading to bias in data Research participants feel coerced or unduly influenced to provide private, sensitive data Provision of private, sensitive data can lead to physical, economic, emotional, social, or legal harm(s) for research participants <p>Example: UNHCR shared detailed database of the Rohingya refugee population with Myanmar's government.</p>	<ul style="list-style-type: none"> Assess and acknowledge vulnerabilities and power dynamics in 'balance power and align incentives' assessment Limit collection of study data with any other purpose (such as intervention eligibility) Define data purpose and use in informed consent De-identify data before sharing if defined and allowed through informed consent process
7. Data collected but not used or needed	<ul style="list-style-type: none"> Unnecessary identifiable information collected and stored that is linked to private, sensitive data can be breached or have other unauthorized disclosure that leads to physical, economic, emotional, social, or legal harm(s) Research waste results in less funding for research 	<ul style="list-style-type: none"> Define data needs with the pre-analysis plan or other pre-specification tool to clearly map data needs to research questions and analysis models Clearly define data that is collected for 'unknown' purposes and assess risk and need Support these efforts with a data inventory (example - Data Inventory Map - Fillable Template - This Is Amos)
8. Data scope creep and/or misuse that results in data used for something other than purpose for which it was collected	<ul style="list-style-type: none"> Data that is used for purposes outside the original purpose can lead to physical, economic, emotional, social, or legal harm(s) Under certain regulations, this data misuse can be illegal or result in legal repercussions for the research team Misuse of data can cause reputational harm to researchers <p>Example: On April 20, 2010, Arizona State University (ASU) agreed to pay \$700,000 to 41 members of the Havasupai Indian tribe to settle legal claims that university researchers improperly used tribe members' blood samples in genetic research.</p>	<ul style="list-style-type: none"> Define data needs with the pre-analysis plan or other pre-specification tool to clearly map data needs to research questions and analysis models (same as above) Define who has access to what data, when, and for what purpose in the Data Management Plan Ensure alignment between Data Management Plan and information provided to the data provider in the Informed Consent and/or Data Sharing Plan Under US HHS regulations, carefully consider use of 'broad consent' during informed consent

Table X: Summary of Risks, Potential harms/adverse events, and risk mitigation efforts to consider for a favorable risk-benefit ratio

Risk	Example(s) of Harm/Adverse events	Risk Mitigation Efforts
<p>9. Incomplete data resulting from the study team not considering all data needs required to fulfill research objectives</p>	<ul style="list-style-type: none"> • Incomplete data can bias the results of the study leading to false positive or negative findings that affect future programming of policies or interventions that affect the economic and social status of data provider populations • Inability to fulfill research objectives can pose reputational harm to researchers • Research waste results in less funding for research 	<ul style="list-style-type: none"> • Define data needs with the pre-analysis plan or other pre-specification tool to clearly map data needs to research questions and analysis models (same as above) • Ensure alignment with research sample and intervention sample through fair subject selection assessment
<p>10. Poor Quality Data (primary and/or secondary data sources)</p>	<ul style="list-style-type: none"> • Poor quality data can bias the results of the study leading to false positive or negative findings that affect future programming of policies or interventions that affect the economic and social status of data provider populations and can pose reputational harm to researchers • Research waste results in less funding for research <p>Example: Data obtained from an implementing partner used to inform influential study later identified as fraudulent data.</p>	<ul style="list-style-type: none"> • Ensure data quality management and assessment practices are established for primary and secondary data collection and detailed under fair methods assessment
<p>11. Loss of confidentiality – <i>with or without additional linkage to private, sensitive data</i> – is a risk driven by several potential actions, including:</p> <ul style="list-style-type: none"> • Intruders breaching systems and accessing raw data • Unintentional leakage or unintentional disclosure of either the raw data or of the information/ knowledge resulting from analysis of the data can occur by (a) of a member of the project team; (b) of known third parties (e.g., government, research partners); who have requested 	<ul style="list-style-type: none"> • Confidential data obtained by nefarious actors can result in physical, economic, emotional, social, or legal harm(s) to research participants • Under certain regulations, loss of confidential data may be illegal or result in legal repercussions for the research team • Irresponsible data management can cause reputational harm to researchers <p>Example: UNHCR shared detailed database of the Rohingya refugee population with Myanmar’s government; Biometric data on Afghan security forces potentially available to the Taliban.</p>	<ul style="list-style-type: none"> • Align data collected with data needs – if do not need to collect direct identifiers, do not (see above) • Ensure staff members have completed the information security awareness and data protection training (Talk with IT team) • Ensure the Data Management Plan describes how to: <ul style="list-style-type: none"> ○ Limit access of identifiable data to staff based on need ○ Encrypt files and devices storing confidential data ○ Store direct identifiers separate from other data as appropriate • Ensure there is a data breach plan that is up to date and known by research team

Table X: Summary of Risks, Potential harms/adverse events, and risk mitigation efforts to consider for a favorable risk-benefit ratio		
Risk	Example(s) of Harm/Adverse events	Risk Mitigation Efforts
<p>or may have access, or who may be motivated to get access in order to misuse the data and information; or (c) by unknown third parties (e.g. due to hackers or other bad actors)</p> <ul style="list-style-type: none"> Anonymized or de-identified datasets combined with other datasets to re-identify participants 		<ul style="list-style-type: none"> Ensure the data sharing strategy is assessed for re-identification risk and appropriate sharing strategy (public, restricted, or no access)

Alignment with this requirement also depends on alignment with all other requirements.

Ensure favorable cost-benefit ratio and accountability

Resources for interventions are not infinite, and neither are resources for research. If money is invested in one research activity, it is not invested in another and this requires an understanding of how the proposed research methods will maximize benefits, minimize harm, while also balancing costs. To do this, research teams must estimate the full cost of conducting the research, including TREE best practices discussed above. This requires properly accounting for fair¹⁵ compensation to research staff members, appropriate payment (such as reimbursement, compensation, or incentives) for research participants for their time and the value of their input into the research activity, necessary communication campaigns with research participants to ensure comprehension of study objectives, the informed consent process, and findings of the research. As the research community recognizes the need to address power imbalances and empower research participants in the design, implementation, dissemination, and learning from the research, additional activities for the ethical conduct of research likely require additional costs.

However, these costs should also be placed in the context of the interventions and policies they study. Research costs – even when considered expensive – are often a fraction of the amount spent on the types of interventions and policies studied. Returning to the moral argument for investments in evidence generation to inform decision-making, the ethical conduct of programmatic and policy decision-making should require measuring the efficacy and effectiveness of such programs.

To achieve alignment with this requirement, research teams should assess, document, and monitor:

1. *Research costs compared to other methods.* It is recommended the research team does a full budget including all labor and materials required for TREE best practices discussed above. When feasible, the research team should assess applicability of various methods and cost-saving alternatives – relying on existing secondary data vs. primary data; trade-offs in the number of data collection rounds; trade-offs for sample size adjustments and outcomes measured. When feasible, presenting research funders with a high, medium, and low research budget can provide decision-makers with an understanding of the true cost of conducting research and the trade-offs required if funders are unwilling or able to fully fund the research.

¹⁵ What is fair compensation and how is it set? This clearly varies by context, but organizations should consider how wages have been set and if they follow methods for establishing fair wages. For example, reference ILO project on indicators and methodologies for setting wage standards - https://www.ilo.org/global/topics/wages/projects/WCMS_826265/lang--en/index.htm

2. *Research costs compared to intervention or policy costs.* It is recommended research teams place the total budget for the research in context of the intervention cost and future potential investment decisions for the studied intervention. For example, a \$1.5 million research activity for a \$500K pilot intervention for which there are no current plans to scale-up is very different than a \$1.5 million research activity for a \$10 million pilot that may result in \$300 million of future investment.

Alignment with this requirement also depends on alignment with all other requirements.

TREE Review Framework

Questionnaire

The TREE Review Questionnaire is the tool by which research teams and independent review committees can document, assess, and monitor these issues and alignment with the ten ethical requirements.

Use by research teams. The questionnaire can be used by research teams as a guide for defined, and often evolving, TREE best practices. Research teams can use the questionnaire to document decision-making throughout the research life cycle around the core set of best practices and defined ethical requirements and ensure relevant documentation (research protocol, informed consent, field manuals, etc) are complete. Research teams do not need to answer all probing questions. If certain questions are answered in other corresponding documentation (in an IRB protocol, Data Management Plan, etc.) the research team can reference the relevant document. In this way, the TREE Review Questionnaire can be a living document that supports research team management and oversight. In addition, completing the questionnaire can support future efforts to generate an Ethics Appendix at the time of final publication of research results (Aseidu, Karlan, et al 2021).

Use by independent review committee. The TREE Review Questionnaire can also be the input into an independent assessment of how well the research design and implementation aligns with the 10 ethical standards presented above. The study team can submit all existing documentation for review and assessment, including but not limited to the research protocol, informed consents, questionnaires, and field manuals. The

Review Committee

When the research team or its organization prefers an independent review of the research team's documentation and responses to the TREE Review Questionnaire, a Review Committee can be established. Ideally, the Review Committee would consist of:

1. *Independent TREE representative(s).* There should be one-two peer reviewers who represent the same general level of expertise as the research team (5+ years' experience designing and implementing international social science research projects and overseeing relevant data collection activities). These reviewers should have specific training and understanding of key TREE best practices as presented here, as well as an understanding of risks facing the ten ethical requirements.
2. *Sector expert reviewer.* As with any peer review – for funding, for publication – independent reviews benefit from specific sectoral expertise. When feasible, a reviewer with a specific expertise in the sector studied – energy, water, sanitation, female empowerment – should be included on the Review Committee to bring relevant experience and expertise through the sector lens.
3. *Regional expert reviewer.* As with any peer review – for funding, for publication – independent reviews benefit from regional or other contextual expertise. When feasible, a reviewer with a specific expertise in the region and/or population studied should be included on the Review Committee to bring relevant experience and expertise through the contextual lens.
4. *Participant population representative.* Community engagement is critical at all stages of the research life cycle. Ideally, the Review Committee would include a representative from the participant population. This representative should be selected based on ability to describe participants' vulnerabilities, comprehension of the informed consent statement and process, risks of harm, and provide input on the research team's risk mitigation strategies.

Review Process

There are critical milestones in a research lifecycle that can benefit from an independent review of the research team's assessment and decision-making regarding best practices in TREE: (i) funding decision, (ii) before data collection (baseline, interim, final), and (iii) before final publication. For each stage, the TREE Review Framework can follow a three-step process:

- **Step 1: Research Team assessment (4 weeks).** The research Program Manager should initiate the TREE review on behalf of the research team. In addition to completing the TREE review, the research team should submit the (i) informed consent statement(s), (ii) study-related questionnaire(s), (iii) research protocol, and (iv) any data management and sharing documentation.
- **Step 2: TREE Review Committee assessment (4 weeks).** The Review Committee should receive all materials including the completed TREE Questionnaire. The Committee should have at least 2-3 weeks to review materials and may meet ahead of the TREE Review meeting to reach consensus on the Review Committee assessment and feedback to the research team.
- **Step 3: TREE Feedback and Follow-up (2+ hours).** The Review Committee and the research team should meet to review the assessments, provide feedback, and discuss and document risk mitigation strategies as relevant. The independent reviewers may support the research team on additional follow up, such as reviewing revised informed consent statements, questionnaires, and protocols.

For some organizations and research teams, it may be preferable to conduct an initial TREE review capacity building session prior to initiating the review process.

Outputs

The TREE Review Framework is expected to produce the following outputs:

1. **Study materials** – Updates and/or revisions to existing study materials that document specific decisions and protocols in place to align study design and implementation with the ethical standards. These include but are not limited to the research protocol, informed consents, questionnaires, field manuals. Ideally these materials are publicly available alongside any other study report and documentation.
2. **TREE Review Questionnaire** – The TREE Review Questionnaire can be a one-stop-shop for summarizing the study design and implementation factors and decisions that inform alignment with the ethical standards. Ideally, the Questionnaire is published alongside other study materials, including but not limited to the research protocol, informed consents, questionnaires, field manuals, and analysis reports.

Appendix 1: Understanding vulnerability

TABLE 4.1: Taxonomy of Vulnerability for Research Participants ¹⁶			
Vulnerability	Definition	Potential Causes	Ethical Research Practice
Cognitive Vulnerability	The research subject does not have the capacity to deliberate and decide whether to participate in the study	Immaturity (through age, other cause), dementia, certain types of mental illness, disability; educational deficits and unfamiliarity with the language; situational mental distress/crisis	Mitigated through proper Informed Consent : plain-language, advance directives (where incapacity is anticipated), supplementary educational measures to ensure comprehension, and the proper use of surrogates and advocates
Juridic Vulnerability	The research subject is liable to the authority of others who may have an independent interest in the research subject's participation	Prisons and the military, where wardens and officers have legal authority over prisoners and enlistees; Children under the authority of their parents, Students subordinated to Professors, Institutionalized persons subject to the authority of custodians, women legally subject to their husbands;	Mitigated through proper Informed Consent : devise a consent procedure that will insulate the research subject from the hierarchical system to which he or she is subject. This is particularly challenging if the researcher/project team is a part of the hierarchical system (so program beneficiaries who are surveyed by their benefactors).
Deferential Vulnerability	The research subject exhibits patterns of deferential behavior that may mask an underlying unwillingness to participate	May be driven by social and political pressures to follow/defer to others despite own desire to not follow/defer (<i>often present with juridic vulnerability</i>)	Mitigated through Sample Recruitment/ Screening and Informed Consent : Inclusion Criteria/Sample Selection may require input of local informants or consultants to devise a process that eliminates as much as possible the social pressures a research subject feels. Informed consent mitigation same as above.
Allocational Vulnerability	The research subject is lacking in important social goods that will be provided because of participating in the research	When participation in the research can provide research subject a social good - money, housing, medical care, childcare, burial benefits, opportunities to benefit the community, freedom – that they otherwise do not have access to	Mitigated through Sample Recruitment/ Screening and Compensation : The Inclusion Criteria/Sample Selection may require input of local informants to determine whether or not the offering of research participation may introduce undue influence; Project Teams must also carefully consider Compensation packages to limit their under or over-value and may need to consider not just their research sample, but also neighboring communities /individuals/households that are excluded and may feel resentment for the exclusion.
Infrastructural Vulnerability	The political, organizational, economic, and social context of the research setting does not possess the integrity and resources needed to manage the study	Research subjects have access to research requirements (phone, transport); Project teams have access to research requirements (skills for specific biomarker tests, psychological tests, etc.; electricity, transport, safety)	Mitigated through Study Design : The study design/protocol should be carefully reviewed for local context and cultural sensitivities.
Medical Vulnerability	The research subject has been selected, in part, because he or she has a serious health-related condition for which there are no satisfactory remedies	When (i) illness is severe and (ii) no safe, effective, and otherwise satisfactory treatments are available, patients can be primarily driven to participate based on false hope for benefits	Mitigated through Study Design and Informed Consent : Given the interests and aspirations of both parties (and the poor bargaining position of the research subject) work toward fair division of the benefits and burdens of cooperation and design the study to maximize the likelihood of subject benefit based on medical intervention found to be safe and effective; communicate benefits and their probabilities for success through Informed Consent.

¹⁶ This reference is extracted from Hoces de la Guardia and Sturdy (2019) and adapted from NBAC 2001.

Appendix 2: TREE Review Questionnaire

The [TREE review Questionnaire is available here.](#)

References

- Asiedu, Edward, Dean Karlan, Monica P. Lambon-Quayefio, and Christopher R. Udry. A Call for Structured Ethics Appendices in Social Science Papers NBER Working Paper No. 28393 January 2021, Revised May 2021. DOI 10.3386/w28393
- Banerjee, A. V., Duflo, E., & Glennerster, R. (2008). Putting a band-aid on a corpse: incentives for nurses in the Indian public health care system. *Journal of the European Economic Association*, 6(2-3), 487-500.
- Barnett, C. and Munslow, T. (2014) Workshop Report: Framing Ethics in Impact Evaluation: Where Are We? Which Route Should We Take? IDS Evidence Report 98, Brighton: IDS. Accessed online on 6/26/2021 - Workshop Report: Framing Ethics in Impact Evaluation: Where Are We? Which Route Should We Take? - Institute of Development Studies (ids.ac.uk)
- Barth-Jones, Daniel, The 'Re-Identification' of Governor William Weld's Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now (July 2012). Available at SSRN: <https://ssrn.com/abstract=2076397> or <http://dx.doi.org/10.2139/ssrn.2076397>
- Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, Fowler J, Waldo A, Weibel N, Patrick K, Klemmer S, Melichar L Reimagining Human Research Protections for 21st Century Science *J Med Internet Res* 2016;18(12):e329. doi: [10.2196/jmir.6634](https://doi.org/10.2196/jmir.6634)
- Broockman, David & Kalla, Joshua & Aronow, Peter. Irregularities in LaCour (2014). [10.31222/osf.io/qy2se](https://doi.org/10.31222/osf.io/qy2se).
- Claesen Aline, Gomes Sara, Tuerlinckx Francis and Vanpaemel Wolf 2021. Comparing dream to reality: an assessment of adherence of the first generation of preregistered studies. *R. Soc. open sci.* 8211037211037 <http://doi.org/10.1098/rsos.211037>
- Center for Global Development (2006). 'When will we ever learn? Improving lives through impact evaluation.' Washington, DC: CGD. Available at: <https://www.cgdev.org/publication/when-will-we-ever-learn-improving-lives-through-impact-evaluation> (accessed 21 December 2021).
- Cojocar, Laura, Saugato Datta, Abigail Sellman, Cecilia Shang. November 2021. Behavioral Insights on the Use of Evidence for Local Policymaking in South Africa. ideas42. Available at: https://www.ideas42.org/wp-content/uploads/2021/12/Behavioral-Insights-on-the-Use-of-Evidence_ideas42.pdf (accessed 21 December 2021)
- Davies SE. (2020) The introduction of research ethics review procedures at a university in South Africa: review outcomes of a social science research ethics committee. *Research Ethics*. 2020;16(1-2):1-26. doi:10.1177/1747016119898408
- Desposato, Scott. (2014). Ethical Challenges and Some Solutions for Field Experiments. <http://www.desposato.org/ethicsfieldexperiments.pdf>
- Desposato, Scott. Forthcoming. "The Ethical Challenges of Political Science Field Experiments." In *The Oxford Handbook of Research Ethics*. edited by Ana Iltis and Douglas MacKay. New York: Oxford University Press.
- Devereaux, Mary. (2014). Rethinking the Meaning of Ethics in RCR Education. *Journal of microbiology & biology education*. 15. 165-8. 10.1128/jmbe.v15i2.857.
- DFID (2016) Review of Ethics Principles and Guidance in Evaluation and Research. Accessed online 6/26/2021 - <https://www.gov.uk/government/publications/review-of-ethics-principles-and-guidance-in-dfid-evaluation-and-research>
- Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA*. 2000 May 24-31;283(20):2701-11. doi: 10.1001/jama.283.20.2701. PMID: 10819955.

Evans, David K. 2021. "Towards Improved and More Transparent Ethics in Randomised Controlled Trials in Development Social Science." CGD Working Paper 565. Washington, DC: Center for Global Development.

Friedman, J. (2014). Should impact evaluation be justified by clinical equipoise or policy equipoise? World Bank Blog, 17 April 2014.

Gelinas, Luke, Alan Wertheimer, and Franklin G. Miller. 2016. "When and Why Is Research Without Consent Permissible?" *Hastings Center Report* 46, no. 2: 35-43.

Gertler, Paul & Sebastian Galiani & Mauricio Romero, 2018. "How to make replication the norm," *Nature*, *Nature*, vol. 554(7693), pages 417-419, February. <https://doi.org/10.1038/d41586-018-02108-9>

Gilman, RH, Garcia, HH (2004) Ethics review procedures for research in developing countries: a basic presumption of guilt. *Canadian Medical Association Journal De L'Association Medicale Canadienne JAMC* 171(3): 248–249.

Grady C. (2015). Institutional Review Committees: Purpose and Challenges. *Chest*, 148(5), 1148–1155. <https://doi.org/10.1378/chest.15-0706>

Hallonsten O. On the essential role of organized skepticism in science's "internal and lawful autonomy" (Eigengesetzlichkeit). *Journal of Classical Sociology*. March 2021. doi:10.1177/1468795X211000247

Hoces de la Guardia, F. and J. Sturdy. 2019. "Best Practices in Transparent, Reproducible, and Ethical Research", for the Inter-American Development Bank Technical Note series. <http://dx.doi.org/10.18235/0001564>

Hoces de la Guardia, Fernando, Sean Grant, and Edward Miguel. "A Framework for Open Policy Analysis." *Science and Public Policy*, no. scaa067 (December 3, 2020). <https://doi.org/10.1093/scipol/scaa067>. An open-access version is available [here](#).

Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005 Aug;2(8): e124. doi: 10.1371/journal.pmed.0020124. Epub 2005 Aug 30. PMID: 16060722; PMCID: PMC1182327.

Karlan, Dean and Jacob Appel. *Failing in the Field: What We Can Learn When Field Research Goes Wrong*. Princeton: Princeton University Press, 2016. muse.jhu.edu/book/58060.

Khera, Reetika, Some Questions of Ethics in RCTs (February 7, 2021). Available at SSRN: <https://ssrn.com/abstract=3780908> or <http://dx.doi.org/10.2139/ssrn.3780908>

Kretser, A., Murphy, D., Bertuzzi, S. *et al.* Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. *Sci Eng Ethics* 25, 327–355 (2019). <https://doi.org/10.1007/s11948-019-00094-3>

Kruger, Mariana. (2014). Research Ethics in Africa: A Resource for Research Ethics Committees. 10.13140/2.1.2560.8646.

Laitin, David

D., Edward Miguel, Ala' Alrababa'h, Aleksandar Bogdanoski, Sean Grant, Katherine Hoerberling, Cecilia Hyunjung Mo, Don A. Moore, Simine Vazire, Jeremy Weinstein, Scott Williamson Proceedings of the National Academy of Sciences Dec 2021, 118 (52) e2106178118; DOI: 10.1073/pnas.2106178118

Mackay, Douglas. 2015. "Standard of Care, Institutional Obligations, and Distributive Justice." *Bioethics* 29, no. 4: 262-273.

Mackay, Douglas. 2018. "The Ethics of Public Policy RCTs: The Principle of Policy Equipoise." *Bioethics* 32, no. 1: 59-67.

Mackay, Douglas, Averi Chakrabarti, Government Policy Experiments and Informed Consent, *Public Health Ethics*, Volume 12, Issue 2, July 2019, Pages 188–201, <https://doi.org/10.1093/phe/phy015>

- MacKay, Douglas. 2020. "Government Policy Experiments and the Ethics of Randomization." *Philosophy & Public Affairs* 48, no. 4: 319-352. <https://doi-org.libproxy.lib.unc.edu/10.1111/papa.12174>
- MacKay, Douglas, and Katherine Witte Saylor. 2020. "Four Faces of Fair Subject Selection." *The American Journal of Bioethics* 20, no. 2: 5-19.
- MacKay, Douglas. Forthcoming(a). "The Ethics of Public Policy Experiments: Lessons from Clinical Research Ethics," in *The Oxford Handbook of Research Ethics*, edited by Ana S. Iltis and Douglas MacKay, Oxford University Press, October 2020. DOI: [10.1093/oxfordhb/9780190947750.001.0001](https://doi.org/10.1093/oxfordhb/9780190947750.001.0001)
- MacKay, Douglas. Forthcoming(b). "Fair Subject Selection in Clinical and Social Scientific Research," in *The Oxford Handbook of Research Ethics*, edited by Ana S. Iltis and Douglas MacKay, Oxford University Press, October 2020. DOI: [10.1093/oxfordhb/9780190947750.001.0001](https://doi.org/10.1093/oxfordhb/9780190947750.001.0001)
- McCormick JB, Boyce AM, Ladd JM, Cho MK. (2012). Barriers to considering ethical and societal implications of research: perceptions of life scientists. *AJOB Primary Res.*3340-50
- Merritt MW. Health researchers' ancillary care obligations in low-resource settings: how can we tell what is morally required? *Kennedy Institute of Ethics Journal*. 2011 Dec;21(4):311-347. DOI: [10.1353/ken.2011.0019](https://doi.org/10.1353/ken.2011.0019). PMID: 22187929; PMCID: PMC3576820.
- Merton RK. 1942. The Ethos of Science, *J. Legal and Political Sociology*. 1: 115-126. Reprinted In: Merton RK, Sztomka P., editor., editors. *Social structure and science*, Chicago: University of Chicago Press, 1996.
- Miguel, E., C. Camerer, K. Casey, J. Cohen, K. M. Esterling, A. Gerber, R. Glennerster, D. P. Green, M. Humphreys, G. Imbens, D. Laitin, T. Madon, L. Nelson, B. A. Nosek, M. Petersen, R. Sedlmayr, J. P. Simmons, U. Simonsohn, M. Van der Laan. 2014. "Promoting Transparency in Social Science Research." *Science*, [10.1126/science.1245317](https://doi.org/10.1126/science.1245317).
- National Bioethics Advisory Commission [NBAC]. *Ethical and Policy Issues in Research Involving Human Participants. Volume II: Commissioned Papers*. Rockville, MD: National Bioethics Advisory Commission [NBAC]; 2001: G1 - G13 (Accessed 11/14/2021 - <https://aapcho.org/wp/wp-content/uploads/2012/02/Kipnis-VulnerabilityinResearchSubjects.pdf>)
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. [Bethesda, Md.]: The Commission. (Accessible online here <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html>)
- Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science* 394:aac4716. <https://doi.org/10.1126/science.aac4716>
- Peisakhin, L., & Pinto, P. (2010). Is transparency an effective anti-corruption strategy? Evidence from a field experiment in India. *Regulation and Governance*. 4(3), 261- 280.
- Phillips, Trisha. 2021. "Ethics of Field Experiments." *Annual Review of Political Science* 24: 277-300.
- Sturdy, J., Burch, S., Hanson, H., & Molyneaux, J. (2017, March 3). Opening Up Evaluation Microdata: Balancing Risks and Benefits of Research Transparency. <https://doi.org/10.31222/osf.io/s67my>
- Sweeney, Latanya. 2000. "Simple Demographics Often Identify People Uniquely." Carnegie Mellon University, Data Privacy. Type of Work: Working paper. [Project website](#)
- Weber S, Hardiman M, Kanja W, Thomas S, Robinson-Edwards N, Bradbury-Jones C. Towards Ethical International Research Partnerships in Gender-Based Violence Research: Insights from Research Partners in Kenya. *Violence Against Women*. October 2021. doi:10.1177/10778012211035798

Wood BDK, Müller R, Brown AN. Push button replication: Is impact evaluation evidence for international development verifiable? PLoS One. 2018 Dec 21;13(12): e0209416. doi: 10.1371/journal.pone.0209416. PMID: 30576348; PMCID: PMC6303036.

Zechmeister, Elizabeth. 2013. Ethics and Research in Political Science: The Responsibilities of the Researcher and the Profession. In *Ethical Challenges in Political Science Experiments*, ed. Scott Desposato.