



**Economics IITD**  
Humanities and Social Sciences

WORKING PAPER SERIES

IITD/EN-2021/01

## Some Questions of Ethics in RCTs

Reetika Khera

MAY 2021

Disclaimer: The opinions expressed within this article are the personal opinions of the author(s). IIT Delhi is not responsible for the accuracy, completeness, suitability, or validity of any information on this article. The information, facts or opinions appearing in the article do not reflect the views of IIT Delhi and IIT Delhi does not assume any responsibility or liability for the same.

## Some questions of ethics in RCTs

Reetika Khera<sup>1</sup>

**Abstract:** Questions of ethics in Randomized Controlled Trials (RCTs) in development economics need greater attention and a wider perspective. RCTs are meant to be governed by the three principles laid out in the Belmont Report, but often violated them, e.g. when local laws are flouted. In other cases, the framework of the Belmont Report itself has proved inadequate: for instance, when there are unintended outcomes or adverse events for which no-one is held accountable. Primarily using RCTs conducted in India, this paper highlights eight areas of concern. RCTs also have a disproportionate influence on shaping research agendas and on policy. Though ethical issues have been raised, there has been little engagement from the RCT community – a manifestation of its power in the profession. As current safeguards (such as oversight by Institutional Review Boards) have failed to protect human subjects, the concluding section discusses possible ways to resolve these issues.

### Introduction

Methodological issues in randomized controlled trials (RCTs) in economics have received a fair bit of attention (see Deaton 2019, Cartwright and Deaton, Barrett, Ravallion, Pritchett, Kabeer among others). To some extent there has been a debate on these issues between the critics and the proponents of RCTs. On ethical issues with RCTs in economics, however, while questions have been raised<sup>2</sup>, responses have been less forthcoming, or shallow and evasive.

In the US at least, RCTs are guided by the ethical principles laid out in the Belmont Report that lists Respect for Persons, Beneficence, and Justice as the three pillars of ethical RCTs ([J-PAL website](#)). The first includes allowing people to decide for themselves what to participate in. Beneficence includes ‘do not harm’ and ‘maximize benefits, minimize risks’. Justice deals with participant selection and requires that those who take the risks should receive the benefits.

This paper discusses some question of ethics arising primarily from Indian RCTs. I highlight cases where it appears that the Belmont Report principles may have been violated or the intervention raises some questions. I also point out that, at times, the Belmont Report is inadequate because it evolved for biomedical and behavioural research involving human subjects, and is now being applied to different settings. This is not meant to be an exhaustive review of RCTs that raise ethical questions. Some, but not all, of these concerns have been raised in earlier work also (Drèze 2018, Drèze 2021, Sneha P 2020).

I raise eight areas of concern: legality of interventions, adverse outcomes and accountability, costs of RCTs, the treatment of rights-bearing human beings as experiment subjects, distortion of research agendas, the lack of engagement on these issues, communication of research findings and exercise of power.

---

<sup>1</sup> I am grateful for thoughtful and detailed feedback from Angus Deaton, Pritha Dev, Jean Drèze, Aashish Gupta, Megha Jha, Ankur Sarin, Abhishek Shaw, Anmol Somanchi, Alessandro Tarozzi and Akshaya Vijayalakshmi. I am also grateful to Ananya Jain for assistance in compiling the data on budgets.

<sup>2</sup> See Abramowicz and Szafarz (2019), Bédécarrats, Guérin, Roubaud and Baele (2020), Barrett and Carter (2020), Drèze (2018, 2020), Deaton (2019) and Menon (2020).

There is a wider concern – not discussed here - about what one can learn from RCTs. These studies are often, if not always, sanitized of political and social considerations. The technocratic approach RCTs espouse have serious limitations in enhancing understanding. At their worst, such technocratic approaches undermine democratic policy-making. These broader questions – of values, objectives, one’s position in the system when giving advice, and so on - are discussed by Drèze (2018).

### **1. Legality of interventions**

Several RCTs make interventions that violate the law of the land or are at least questionable from that point of view. An early RCT used “boost” (speed money) to accelerate the acquisition of a driving license ([Bertrand, Djankov, Hanna and Mullainathan 2006](#)). In Delhi, to study whether transparency brought by the enactment of a Right to information Act 2000 would reduce corruption, one intervention arm of the experiment involved paying a bribe to get a ration card application approved ([Peisakhin and Pinto, 2010](#)). The authors acknowledge that paying a bribe raises legal and ethical issues. Eventually, they convince themselves, their IRB and journal editors and reviewers that the intervention, though illegal, does not violate any of the three principles of the Belmont Report.<sup>3</sup>

A recent experiment in Odisha involved bargaining down wages offered to casual daily wage labourers at a labour market to study “collusion” among them to maintain a wage floor ([Breza, Kaur and Krishnaswamy 2019](#)). In his critique of it, [Sarin 2019](#) rightly asks: “A natural question that should arise is: why should such an activity not be deemed illegal? And if it is, what legal and ethical frameworks allow researchers trained and working in the best universities of the world, to break the law with such impunity?”

The laws in question here exist to protect citizen’s rights (from bribery or labour rights), how/why did the researchers get IRB approval. Even if these were questionable laws, where breaking them could be viewed as acts of civil disobedience, can researchers do so using “beneficence” as a cover?

### **2. Adverse outcomes and Accountability**

RCT interventions are always made with the intention of improving outcomes, as per the understanding and knowledge of the researchers. But in some cases, the outcomes may be contrary to what was expected: e.g., in an RCT in Malawi, school years went down for some students, as a result of the intervention (Dizon-Ross 2019). In such cases, the price is paid by human subjects who were part of the RCT and often unaware that they are participants in such an experiment. Reporting of, and as a result, action to prevent, adverse events and outcomes is weak.

In [Bihar](#), an innovation in the flow of funds was introduced to improve the performance of the National Rural Employment Guarantee Act (NREGA) in terms of employment generation ([Banerjee, Duflo, Imbert, Mathew, Pande 2020](#)). Unfortunately, the intervention ended up worsening programme performance *and* exacerbating a pre-existing problem of delays in wage payments.<sup>4</sup> The NREGA law mandates that if workers are not paid on time (within 15 days), they are entitled to compensation. The RCT had not budgeted for such a contingency. That the government rarely honours the provision is poor justification to not budget for it (Drèze, 2021).

---

<sup>3</sup> By way of justification, they state “the officials in question would probably have been surprised to consider their money-taking activity as anything other than standard bureaucratic practice” and sidestep the question of “what right scholars have to violate the law of another country while conducting research” by stating that it is a complicated issue (pp. 270-271).

<sup>4</sup> The paper itself suggests that there was reduction in total expenditure, but earnings did not suffer, so there must have been a reduction in corruption. But whether such an interpretation holds is not obvious (Drèze, 2021).

In Udaipur (Rajasthan), an experiment sought to improve the attendance of nurses by introducing attendance ([Banerjee, Glennerster and Duflo 2008](#)). Here, attendance rates eventually got worse (down to 25%), after an initial improvement (60%), and even after the experiment ended attendance rates did not return to pre-intervention levels (already very low, less than 30%). In this case, rural residents ended up finding a nurse less often than before.

In Jaipur (Rajasthan), policemen were incentivized to perform their duty (checking people for drunken driving) by offering them a carrot by way of a coveted posting, away from their current duties ([Banerjee, Chattopadhyay, Duflo, Keniston and Singh, 2012](#)). The results were as expected: the incentive resulted in the policemen doing their duty 'better' (measured by the number of drunken people stopped by the cops). But what assurance do we have that this improved performance did not come at the cost of innocent people being stopped - and harassed - by policemen for a small bribe?<sup>5</sup> Was the possibility of such inappropriate behaviour by the policemen accounted for, and if so, was it adequately monitored and controlled? This issue is not discussed in the paper.

One could argue that ultimately these are failures of elected representatives and administrators. One could also argue that such failures are tolerated by people all the time. Yet, it is also true that bad outcomes were made worse (unintentionally) by the experiment. More importantly, in the case of non-RCT related failures, citizens can resort to established mechanisms of holding them accountable – either voting out elected representatives or by complaining about the non-performance of administrators. While these are far from satisfactory, they exist.<sup>6</sup> Who can aggrieved parties approach in the case of RCTs?

Moreover, in some cases the experiment ends up weakening already weak mechanisms of accountability. In the case of the Bihar NREGA wage experiment, the rural development department could (indeed, should) have functioned as a 'circuit breaker' (the researchers had an MOU with them). The primary responsibility of the department is towards NREGA labourers. However, warnings about the inappropriateness of the intervention from field functionaries were ignored (Drèze, 2021). The senior bureaucrat had been co-opted (by the researchers) as a co-PI and co-author.

What is being framed as a 'co-optation' here, is possibly viewed as 'collaboration' by the randomistas. No doubt there is merit in bureaucracy engaging more with academic "evidence" (broadly defined), just as there might be in bureaucracy engaging with industry. For the latter, there are nominal safeguards (e.g., a 'cooling off' period before bureaucrats can join private industry) to prevent a revolving door phenomenon; engagement with academia where pecuniary rewards might be offered needs to draw similar boundaries.<sup>7</sup> The Belmont Report cannot help here.

As questions about external validity of RCTs were raised, some randomistas considered 'on scale' experiments as the way forward (Muralidharan and Niehaus 2017). These involved collaborating

---

<sup>5</sup> Dandona, Kumar and Dandona (2006) report that a quarter of those who were caught for a traffic violation paid a bribe. When policemen in the experiment group stop more people, they increase their opportunities for such bribes. Indeed, one reason why public facing posts are preferred is because it increases such opportunities. The authors recognize that the current postings are viewed 'punishment posts' (p. 9), but not why they might be considered as such.

<sup>6</sup> In the case of the RCT, those who pay the price often do not know that their village and therefore they themselves are the subjects of an experiment. This raises questions of consent (discussed later) that the Belmont Report does not address.

<sup>7</sup> This is not the only instance of a senior government bureaucrat being a co-PI. A retired chief secretary of the state government is a co-PI in an experiment in Tamil Nadu studying pension applications by the elderly. In the Jaipur drunken driving experiment, an IPS officer was co-PI.

with state and central governments. As more bureaucrats join this bandwagon, RCTs are likely to raise conflict of interest issues.<sup>8</sup> Moreover, McClendon (2012) has pointed out, when RCTs involve public officials (as participants) it raises ethics regarding the use of their time.<sup>9</sup>

Greater transparency is also necessary. In the police reforms experiment in Jaipur, some of the interventions were designed on the basis of government recommendations (e.g., Police Reforms Commission), but the drunken driving intervention was the idea of the randomistas (p. 4).<sup>10</sup> In the [nurses experiment in Udaipur](#), those whose attendance was below 50% could have their salaries docked, or they could be suspended (if their low attendance persisted into the second month). Does such a provision for docking salaries exist in the state's laws and if so, can it be activated at the behest of a researchers request. We do not know the genesis of the idea (of using financial incentives to boost attendance). Depending on whose idea it was (the Chief Medical Officer, state health department, the local NGO and/or JPAL researchers), the issues related to accountability will vary. If researchers merely "piggy-back" on a state policy by convincing them to randomize its implementation it raises one set of questions, whereas if they suggested the intervention *and* its randomization, it would raise another set of questions (say, regarding whether an open consultative process was engaged in).

The answers to these questions help determine who is to be held accountable and questions regarding the exercise - and possible abuse - of power.<sup>11</sup>

### 3. Costs

An important tenet – implicitly, if not explicitly - of the philosophy of the randomistas is that policy-making without "evidence" would be wrong. Similarly, RCTs should themselves be evaluated for cost effectiveness (Pritchett 2018 and Kapur 2018).

Questions that could have been adequately answered by other methods, are posed in expensive experimental setting because of the unspoken hierarchy of methods. For instance, two studies of the integration of Aadhaar based biometric authentication (ABBA) in the PDS come to almost exactly the same conclusions (Drèze, Khalid, Khera and Somanchi, 2017 and Muralidharan, Niehaus and Sukhtankar, 2020). While the former cost less than Rs. 5 lakh, tens of thousands of dollars were spent on the latter.

There are three cost related questions – overall costs of RCTs, who funds them and what the money is spent on. RCTs are expensive. The [Odisha Scabs](#) study cost just over Rs. 2 crores (\$279,974), the equivalent of 100,000 days of work at the local wage. A [cash transfer](#) study in Chandigarh, Dadra and Nagar Haveli and Pondicherry was just under Rs. 2 crores (\$196,798). The [Bihar NREGA wages](#) study

---

<sup>8</sup> On the other hand, some critics have been concerned about RCTs that bypass local governments entirely. Either way, doing experiments 'on scale' raises this question.

<sup>9</sup> A (non-RCT) study in Madhya Pradesh, sent fake patients to public health providers ([Das, Holla, Mohpal and Muralidharan](#), 2015) taking away precious time from real patients.

<sup>10</sup> "Since performance pay and explicit financial incentives were totally alien to the police managerial system, even senior officers who appreciated incentives in principle thought that adopting them would require infeasible legal changes. The researchers' contribution as pseudo-consultants was to help identify a window where it was possible to provide some incentive within the existing organization." (p. 4)

<sup>11</sup> For instance, are local researchers (more or less) likely to get an appointment and hearing with state government officials than Indian/non-Indian researchers based at foreign universities.

was GBP 19,000 from IGC and USD 279,000 [from 3ie](#). Total budgets for RCTs are not easy to find, but see Table 1.<sup>12</sup>

In the US, the NIH and NSF are among the biggest funders (their budgets were also the easiest to find). Bilateral aid agencies (e.g., DfID) are also big donors. Funds also come from large private donors – philanthrocapitalists and their foundations. What are the implications of research agendas being increasingly shaped by philanthrocapitalists, donors not accountable in the same way as, say, government-funded research might be (see Deaton 2019 and Giridhiradas 2018).

The third question relates to what the money is spent on. Information on cost breaks-up for RCT would help. To illustrate, while the ethics of teacher surveillance in this manner are questionable, an early RCT provided digital cameras in non-formal education schools, to monitor their hours of work. In the early 2000s when this was done, digital cameras were expensive, not quite ubiquitous as they are now. The schools where these cameras were provided had only the most basic of teaching aids (not much more than a blackboard, chalk and seating arrangements). Perhaps this is merely distasteful, and certainly not specific to RCTs. One also hears of low pay to field investigators, short duration contracts with breaks of a few days or weeks in order to avoid paying benefits. Pay ratios (say, salary of the PI to that of the lowest paid employee) could be made public. Development research, in general, is fraught with such concerns.<sup>13</sup>

#### 4. Rights-bearing human beings as experiment subjects

Sometimes, the questions that RCTs seek to answer, and the interventions they make in order to answer those questions, appear not to prioritize people as rights-bearing human beings, in violation of the spirit of the “Respect for persons” principle of the Belmont Report.<sup>14</sup>

The Odisha wage bargaining study involved bargaining down the wage offer made to poor exploited labourers, knowing that the offer may violate the minimum wage law. This is a clear violation of the ‘do no harm’ principle. Children are classified as ‘vulnerable populations’ in the Belmont Report, for whom special protections are supposed to be in place. They have often been the subjects of RCTs (e.g., on education or nutrition), and it is not clear whether special protections were put in place.

When entire villages or blocks are in the treatment arm of an experiment, how does informed consent work? Whether it is an NREGA worker, or a person entitled to subsidized grain from the PDS, participants are not informed that they are subjects of a study evaluating the effectiveness of a new technology. In such a situation, there is no question of seeking informed consent (required by the Belmont report). *This* question of consent has not been addressed in the Belmont Report. When I raised it at a meeting, consent was interpreted as seeking it from survey respondents for participating in the surveys (see Menon, 2020 and Giradeau, Caille, Le Gouge, Ravaud, 2012).

Two RCTs involved introduction of a new technology (the NREGA wage payment experiment in Bihar and the introduction of ABBA in the PDS in Jharkhand). Both RCTs were latching on to policy decisions that had possibly already been made by state governments. However, in the absence of an

---

<sup>12</sup> Budgets are not declared on the AEA registry, but published papers often provide grant details that allow budgets to be traced.

<sup>13</sup> One researcher familiar with these practices said “Few surveyors are paid minimum wages for skilled workers, for instance, and workplace conditions are pathetic. Talk to any surveyor and they will share their woes. People are paid peanuts, and then they skimp on food and live in terrible bug-infested hotels for 100 rs a day. Then they get sick, but none of the higher-ups who have way higher salaries care.”

<sup>14</sup> There are bunch of RCTs studying sanitation and toilet habits of people: [Demand for sanitation for Kenyan urban slums](#), [Community toilet use in slums](#), [Creating a toilet habit](#).

RCT, the rollout, instead of being randomized to make a study possible, might have followed more sensible placement rule - e.g., such as availability of electricity and internet connectivity in the village (Ravallion [2014](#) and [2020](#)). For instance, in Chhattisgarh, the rollout of PDS smartcards began with a few urban PDS outlets in one city, gradually spread to other urban areas, and then to rural areas. Similarly, in Jharkhand, a new cash transfer experiment for the PDS, was piloted in Nagri Block near the capital city of Ranchi. In the case of ABBA, the RCT forced a randomized rollout, whereby village residents were subject to the vagaries of uncertain technologies sooner than they might have been.

## 5. Distorting research agendas

In 2018, the British Medical Journal carried a spoof article titled "[Look before you leap](#)". It reported the results from an RCT that aimed to study the effects of the use of a parachute on mortality and traumatic injury when jumping from an aircraft.

At times, expensive development RCTs are conducted to ask banal questions (e.g., giving eyeglasses to children with poor eyesight to study the effect on learning outcomes; see Ziliak 2016). Some RCTs ask questions where you learn little that is new: for instance, Oh's (2020) RCT 'establishes' that there are rigidities in the labour market based on caste identity. What constitutes a banal or new is clearly a hard question, but there appears to be little thought to this at the time of approval.

As Deaton (2019) put it "The imposition of a hierarchy of evidence is both dangerous and unscientific." The hierarchy of evidence has tilted research towards RCTs. Not all questions, however, can be answered through RCTs. For instance, Kumar (2016) shows how two prominent RCTs on improving the public education system, in some ways, missed the essence of the issue. Attention to detail and establishing causality is a strength of RCTs, but it can also be a weakness. They tend to undervalue (even ignore) broader understanding of issues (see Drèze 2018).

An ongoing RCT in Tamil Nadu titled "[The Causal effects of old age pensions](#)" seeks to study whether when assisted with the application process (the intervention), excluded old age pensioners who succeed in getting it, have better health and nutrition outcomes and/or financial security.<sup>15</sup>

There is a case for treating pensions as an entitlement given the vulnerabilities faced by the elderly. If the researchers are interested in studying whether pensions impact financial security and health outcomes, is this the best (or only) way to do so? Should assistance be held back from some pensioners in order to study the financial and health benefits? When administrative capacity is low and there is a felt need for supporting pensioners, shouldn't such assistance be provided to all pensioners, irrespective of possible positive effects? Earlier studies show that pensioners face problems in applying due to lack of clarity regarding the process, low awareness about eligibility and procedures (Chopra and Pudusserry 2014 and Marulasidappa, Raonka and Sabhikhi 2014).<sup>16</sup>

---

<sup>15</sup> The abstract on the [AEA registry](#) states: "We will assign households with a member who is likely to be eligible for the pension but not currently receiving it to either a treatment or control condition. Those assigned to treatment will receive assistance in applying for the pension. We will track outcomes of the elderly and their family members for several years following the intervention."

<sup>16</sup> This drives them to middlemen, resulting in corruption (otherwise, the programme is largely corruption free). Could the introduction of a helper, as envisaged by this RCT, increase such questionable practices? What precautions are necessary in the case of state-appointed (middlemen) or assistants. More information on how the idea evolved is important. Was the appointment of an assistant a state policy proposal that is being evaluated by this RCT, and for this reason, being held back to some people for the duration of the experiment? It might be that the study seeks to provide evidence that such appointments are good policy. If the policy lesson is that such assistance be provided only if health and financial benefits accrue should these outcome variables alone determine the decision regarding an extra appointment?

Of course, validating an intuition or common sense requires research. But is an RCT the most cost effective method of doing so? There is pressure to use RCTs because they have often been projected as the gold standard. Often, expensive RCTs are conducted in contexts where empirical research is sorely lacking. Privileging evidence from RCTs, over other evidence, raises questions because expensive RCTs displace other research.

## 6. Lack of engagement

There is inadequate discussion, reflection and engagement from the RCT community on these issues. For instance, one question that has arisen repeatedly is how did the IRB give permission for these questionable interventions? Authors rarely register that questions are being raised about their practices. For instance, Sarin (2019) wrote about the violation of the ‘do not harm’ policy in the Odisha wages RCT. There has been no acknowledgement of the concerns from the authors of the study. Sometimes the publication of an RCT based paper *after* the damage has been done, leads to discussion (primarily on social media). Even that is primarily one-sided.<sup>17</sup>

Sometimes researchers appear to understand that they are on sticky ground, but instead of reconsidering the intervention, they are at pains to find a way around it (e.g., Peishakin and Pinto). In other experiments (say, NREGA wage payment in Bihar), one response has been that the state government had decided to make the intervention anyway, or that the state does not pay compensation either.<sup>18</sup> Another response is that such concerns arise with other forms of research as well. Indeed, some of the concerns raised here do arise with other methods as well. If the lack of accountability is an issue in other forms of research it should be condoned in the case of RCTs too? Can two wrongs make a right?

In any case, some ethical concerns arise with RCTs only, primarily because of the involvement of human subjects and treatment (or the lack of it) affects people in the treatment (control) arm. When participants are asked to pay bribes (especially when the local laws forbid it), is it enough to say that the implications of such an action were discussed with the participants, and they were assured of legal aid should such a need arise? What about the corrupting influence of the intervention on an already corruption-ridden system? Or, the possibility that the intervention legitimizes morally corrupt behaviour (by condoning or endorsing the giving of bribes) by ‘respected’ members of society?

The reluctance to engage raises ethical issues means that the ethical issues persist, and there hasn’t been enough (or any) course correction. It is one thing to have made a questionable intervention in the early 2000s, but such interventions continue to date. It is an awkward conversation; perhaps explains the silence around it. As this [twitter thread](#) points out, the debate has been stuck in a rut.

A small beginning has been made: for instance, J-PAL has a detailed (possibly since 2020) page on ethics. There is a need to go further. The discussion of [IRB proposals](#) requires that ‘adverse events’ are reported.<sup>19</sup> However, the stakes are often high – in terms of money and publication-related tenure decisions. Existing safeguards, relying on self-reporting, have not worked. At times, it appears to suggest strategies for reducing the liability of the researcher: ***“The more involved you are in the design, administration, targeting, or any other aspect of the intervention you are studying, the***

---

<sup>17</sup> There are some exceptions. Recently, in response to outrage on social media to an RCT in Nairobi the authors immediately issued a statement. See also [Lyal](#) (2020).

<sup>18</sup> Imbert, personal communication.

<sup>19</sup> See [this](#) on defining unanticipated problems and adverse events, and when they should be reported to the IRB.



***more the intervention itself falls under the scrutiny of the IRB and under your obligation as a researcher.*** (Bold in original)<sup>20</sup>

## 7. What they find and what they say

“I have often been struck by the contrast between the care that goes into running an RCT and the carelessness that goes into advocating the use of its results” (Deaton 2019).

The communication of RCT results has also been a problem in some cases. In the ABBA in PDS in Jharkhand, the intervention (biometric authentication) did not reduce corruption, contrary to the authors’ anticipation ([Muralidharan, Neihaus and Sukhtankar 2020](#)). Worse, the intervention increased exclusion and transaction costs. The researchers conducted additional endline surveys. One of those studied the outcome of *another* reform (“reconciliation”) introduced by the government (partly because biometric authentication didn’t reduce corruption). Using extrapolations and counterfactuals, the authors present the *possibility* of reduced corruption as evidence of it; exclusion persisted.

The details are laid out carefully in their NBER paper. In popular writings, however, the message that goes out (a) does not distinguish between the results from the original RCT and those from the “additional” quasi-experimental study, (b) the caveats and qualifications regarding interpretations of the non-RCT results are more or less missing; (c) sharing the popular writings, a covering email from one of the authors takes the message further from what one actually learns from the study. According to Muralidharan et al, it becomes a matter of a trade-off between reducing corruption and minimizing exclusion (these issues are discussed in Drèze, Khera and Somanchi, 2020a and 2020b).

What is highlighted, suppressed, diluted or underplayed in popular presentation of findings is important, especially in the case of RCTs, because of the hierarchy of evidence mentioned earlier. In DBT cash pilots in Puducherry, Chandigarh, Dadra & Nagar Haveli, subsidized food through the PDS was replaced with cash transfers. In Puducherry, the state government rolled back the intervention (providing cash instead of food) within three months, because people were unhappy with cash. Food was reintroduced. A few months later, the state government reintroduced cash, along with food. These policy changes meant that it is impossible to study treatment effects in Puducherry. Yet, one result that is highlighted as an ‘overall’ finding is that ‘beneficiary preference for DBT over in-kind PDS benefits grew over time’ (Muralidharan, Niehaus and Sukhtankar, 2017). The fact of the treatment being messed up because of policy changes by the state government is mentioned, but easily lost.<sup>21</sup> It is not mentioned at all in the [Hindustan Times](#) article based on the study.

The main finding – that less than half (47%) of beneficiaries received the cash transfer, and that it managed to reach only 78% in Round 3 (p. 11) – is underplayed. The authors blissfully suggest that

---

<sup>20</sup> In similar vein: “Note that even if you **do not identify any risks**, IRBs may still require you to mention risks and benefits in the consent form, e.g., by saying “there are no risks or benefits to you from this study.” There is no need to try and bend over backwards to convince yourself, the IRB, or the subjects that there is absolutely zero risk. Depending on your study population and context, this can make subjects actually more suspicious or at least confused: if there are no risks, and so far the study sounded completely harmless, why is this explicitly mentioned? Thus, on balance it is actually not necessarily going to reduce participation or harm your relationship with the subjects if you mention small risks. It can help the subjects see that you thought carefully about this but that these are the only risks you came up with.”

<sup>21</sup> When the UT-wise preferences are reported (pp. 13-14), Puducherry is the one where the preference for cash is highest (rising from 32% to 77%), where they mention that “since PDY still provides grains through a state scheme, perhaps subsidy salience becomes less important than the other two UTs” (p.15).

“DBT-based reforms holds long-term promise”; they recommend that “given the implementation challenges”, the government should experiment with a “choice-based DBT” as it allows for a ‘politically and ethically risk-free approach to policy experimentation” (p. 6).

Such instances of dissonance between the findings and the advocacy, raises questions about whether it is the result of “friendly research” (Khera, 2019b). RCTs are expensive and donors want to see results. Do fundees face pressures? This is an important area about which little is known.

## **8. The question of power**

“Even in the US, nearly all RCTs on the welfare system are RCTs done by better-heeled, better-educated and paler people on lower income, less-educated and darker people.” (Deaton 2020.)

Peishakin and Pinto (2010) dismiss the possibility of their work being an example of ‘neocolonialism’. They “reject the dichotomy often drawn between corruption in developed and developing countries”. The fact remains that studies with such interventions are rarely conducted in a developed country.

The exercise of “soft” power is evident in the relationship between the RCT researchers and bureaucracy and politicians. In the nurses study (Udaipur) or NREGA payments (Bihar), we do not know to what extent the policy idea was locally conceived and discussed, and to what extent, local partners received nudges (say, encouraging remarks in casual conversations) from the foreign-based researchers. In Duflo (2017), the (co-author) bureaucrat who failed in his duty as a bureaucrat is lauded as a ‘rare creature – a bureaucrat-plumber’, whereas the Kerala officers who ultimately refused to conduct an RCT because they deemed it discriminatory, are derided.<sup>22</sup> Another glimpse of such power comes from the tweets of two chief ministers after the Nobel prize was announced - either flaunting their association with J-PAL (Delhi) or inviting them to conduct RCTs in their state (Chhattisgarh).<sup>23</sup>

Power manifests itself in stifling discussion of the issues raised here (see Khera 2019b also). Many randomistas are influential in the editorial processes at top journals. In conversations with Indian economists, they helpfully share information on questionable practices, but do not raise these questions themselves, possibly because their papers under review at those very journals. Younger researchers familiar with the nitty-gritty and day-to-day issues of running experiments are not in a position to raise questions because their letters of recommendations depend on their employers.

## **Concluding remarks**

The questions here are for those who conduct RCTs, for funders, for the IRBs of universities, journal editors and reviewers of such studies. Several of these have been raised earlier, but often from outside economics or on social media, which has perhaps made it easier for them to be ignored.

In letter, RCTs are largely able to justify the interventions on ethical grounds, but in spirit, a violation of existing principles is not uncommon. What is evident is that the Belmont Report does not provide an adequate framework for RCTs of the kind that are undertaken in the field of economics. One obvious reason is that the Belmont Report was for experiments in “controlled” settings, and that itself does not often hold. RCTs in development economics have inverted the role of the Belmont Report. Something that was meant to provide protection to human subjects from experimentalists,

---

<sup>22</sup> The “mistake” the Kerala bureaucracy made was to request the researchers to liaise directly with the concerned officials (a professor and a doctor) instead of the ‘top official’ (Ayyappan, 2019).

<sup>23</sup> African researchers have raised questions about who gets authorship.

is treated as a shield for the experimentalists from human subjects. As McDermott and Hatemi (2020) suggest “While institutional review board (IRBs) guidelines may raise ethical awareness, they exist to provide legal protection for institutions....they do not primarily exist for the purpose of protecting participants, nor do they necessarily accomplish this goal.” (p. 2).

In their current form, IRBs are ineffective. They rely almost entirely on the information (on risks, etc.) submitted to it by the researchers themselves. But as in the cases discussed here, the researchers themselves were unaware of the risks (either because they didn’t understand the context well enough, didn’t anticipate them or simply overlooked them). While children and pregnant women are recognized as ‘vulnerable populations’, we find RCTs among them (especially children). We do not know whether the researchers were asked any questions (either at the time of approval or after questions were raised), and how they justified their interventions and study design, in particular cases. IRBs rely on self-regulation and trust (a good thing), but given the questionable experiments, some safeguards ought to have been put added.

Another area that needs urgent attention is reporting adverse events. When adverse events or unintended outcomes occur, it is unclear how many report them to IRBs. One could argue that, say, the greater delays in wage payment in the NREGA study in Bihar, hunger deaths from September 2017 onwards during the PDS-Aadhaar study in Jharkhand were adverse events that should have been reported. A related concern is whether the researchers themselves were able to learn of such adverse events. For instance, in the bribes for ration cards study in Delhi: did any of those who paid a bribe ever get into trouble? Similarly, in the Jaipur drunken driving experiment, we do not know if any of those who passed by the control group checkpoints unnecessarily harassed.

Across disciplines, researchers have been suggesting ways of resolving these ethical issues – from prevention, e.g., Sarin (2019) suggests a moratorium on RCTs until some of these issues are resolved, to greater transparency (Karlan and Udry 2020) and damage control, e.g., Ruxter and Muldon (2019) are interested in what might be done if ethically questionable or compromised research manages to get published. Barrett and Carter (2020) warn of the inadequacy of the current safeguards and advise a greater role for IRBs and journals. Local IRBs, could create an additional cross check, but have tended to become part of the rubber-stamping culture in obtaining approvals. Greater transparency of IRB processes (including questions asked) might help. Oversight bodies (such as NIH’s Office of Human Subjects Research Protections) and training groups (such as the Collaborative Institutional Training Initiative) could be more proactive for RCTs in development economics and political science settings. J-PAL, as the leading RCT association, and AEA, as the pre-eminent association for economics, need to take the initiative urgently.

**Table 1 Selected list of budgets for RCTs conducted in India**

RCT title, year	Cost and funder
<a href="#">Improving Third-Party Audits and Regulatory Compliance in India</a> , 2008-2011	<a href="#">NSF</a> \$321,283
<a href="#">Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunization campaigns with and without incentives</a> , 2005-2007	<a href="#">Mac Arthur Foundation</a> \$280,000
<a href="#">What helps children to learn?</a> Evaluation of Pratham's Read India program in Bihar & Uttarakhand, 2008-2010	William and Flora Hewlett Foundation, <a href="#">Planning</a> ; \$90,000; <a href="#">main</a> \$1,225,000
<a href="#">Helping the Ultra-Poor Use Microcredit in Murshidabad, India</a> , 2007-2011	<a href="#">Ford Foundation</a> \$259,750
<a href="#">Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India</a> , 2007-2009	<a href="#">Stanford OTL Research Incentive</a> \$25,000, <a href="#">Stanford Presidential Fund</a> \$75,000; NIH ( <a href="#">1</a> , <a href="#">2</a> ) \$155,481
<a href="#">Do Management Interventions Last? Evidence from India</a> , 2008-2017	<a href="#">NSF</a> : \$215,380 and World Bank \$150,000
<a href="#">Child control in education decisions: an evaluation of targeted incentives to learn in India</a> , 2007	<a href="#">Russell Sage</a> \$5,000
Improving Labor Market Opportunities to Increase Women's Employment and Education in India, 2003-2006	<a href="#">William F. Milton Fund</a> \$34,695
<a href="#">Anemia and School Participation</a> , 2001-2002	<a href="#">NSF</a> \$240,680
<a href="#">Formal Rainfall Insurance for the Informally Insured in India</a> , 2010	International Growth Centre at LSE \$112,000
<a href="#">Who Gets the Job Referral? Evidence from a Social Networks Experiment</a> , 2009	<a href="#">Russel Sage</a> \$4,971
<a href="#">Discrimination in Grading in India</a> , 2007	<a href="#">NAE/Spencer Post-Doctoral Fellowship</a> \$50,000
<a href="#">Can Informational Campaigns Raise Awareness and Local Participation in Primary Education in India?</a> , 2005-2006	<a href="#">World Bank Trust Fund for Environmentally and Socially Sustainable Development</a> \$205,356
<a href="#">Bridge Classes and Peer Networks Among Out-of-School Children in India</a> , 2006-2007	<a href="#">NSF</a> , <a href="#">ISERP Seed Grant</a> \$353,855 and \$14,640
<a href="#">Scabs: The Social Suppression of Labor Supply</a> , 2015-2018	<a href="#">NSF</a> \$279,974
<a href="#">Enhancing Local Public Service Delivery: Experimental Evidence on the National Rural Employment Guarantee Scheme in Bihar (BNREGA)</a> , 2012-2013	IGC: £19,000 <a href="#">The International Initiative for Impact Evaluation (3ie)</a> : \$279,000

## References

Abramowicz, M., & Szafarz, A. (2019). Ethics of Randomized Controlled Trials: Should economists care about equipoise? Working Papers CEB 19-017, ULB -- *Université Libre de Bruxelles*.

Ayyappan, R (2019), Why Kerala remained cold to the queries of Nobel winners Duflo and Banerjee, Manorama, 16 October. <https://www.onmanorama.com/kerala/top-news/2019/10/16/kerala-remained-cold-nobel-winners-esther-duflo-abhijit-banerjee.html>

Baele, S. J. (2013). The Ethics of New Development Economics: Is the Experimental Approach to Development Economics Morally Wrong?. *Journal of Philosophical Economics*. 7(1).

Banerjee, A. V. and Duflo, E., (2011). Poor Economics. *Public Affairs*.

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.

Banerjee, A. V., Chattopadhyay, R., Duflo, E., Keniston, D., Singh, N. (2012). Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy and training. *National Bureau of Economic Research*. No. w17912. Published as *American Economic Journal: Economic Policy*, 13 (1): 36-66

Banerjee, A. V., Duflo, E., Imbert, C., Mathew, S., Pande, R. (2020), E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India. *American Economic Journal*. Published as *American Economic Journal: Applied Economics*, 12 (4): 37-72. Available here: <https://economics.mit.edu/files/12528>

Barrett, C. B., & Carter, M. R. (2020). Finding our balance? Revisiting the randomization revolution in development economics ten years further on. *World Development*, 127.

Bédécarrats, F., Guérin, I. and Roubaud, F. (eds.) (2020). Randomized Control Trials in the Field of Development: A Critical Perspective. *Oxford University Press*

Bertrand, M., Djankov, S., Hanna, R., & Mullainathan, S. (2006). Does corruption produce unsafe drivers? (No. w12274). *National Bureau of Economic Research* .

Breza, E., Kaur, S., & Krishnaswamy, N. (2019). Propping Up the Wage Floor: Collective Labor Supply without Unions . *National Bureau of Economic Research Working Paper No. w25880*.

Chopra, S., & Pudussery, J. (2014). Social security pensions in India: An assessment. *Economic and Political Weekly*, 68-74.

Dandona, R., Kumar, Anil K., and Dandona, L. (2006), Risky Behaviour of drivers of motorized two wheeled vehicles in India, *Journal of Safety Research*, Vol 37, No. 2, pp. 149-158.

- [Das, J. Holla, A., Mohpal A, and Muralidharan](#), K. (2015). Quality and Accountability in Healthcare Delivery: Audit-study Evidence from Primary Care in India, NBER working paper 21405. Published as *American Economic Review*, 106 (12): 3765-3799.
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21.
- Dizon-Ross, R. (2019). Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments. *American Economic Review*, 109 (8): 2728-65.
- Drèze, J. (2018). Evidence, Policy, and Politics. *Ideas for India*, 3 August, 2018.
- Drèze, J. (2020). Policy beyond evidence. *World Development*, 127.
- Drèze, Jean (2021), The Perils of Embedded Research, forthcoming in *Ideas for India*.
- Drèze, J., Khalid, N., Khera, R., & Somanchi, A. (2017). Aadhaar and food security in Jharkhand: Pain Without Gain?. *Economic & Political Weekly*, 52(50).
- Drèze, J., Khera, R., & Somanchi, A. (2020a). Balancing corruption and exclusion: A rejoinder. *Ideas for India*, 28 September 2020.
- Drèze, J., Khera, R., & Somanchi, A. (2020b). Balancing corruption and exclusion: Comment on the response. *Ideas for India*, 19 November 2020.
- Duflo, E. (2017). The Economist as Plumber. *American Economic Review*, 107(5), 1-26.
- Friedman, J. (2014). Should impact evaluation be justified by clinical equipoise or policy equipoise? World Bank Blog, 17 April 2014. Retrieved from [Should impact evaluation be justified by clinical equipoise or policy equipoise? \(worldbank.org\)](#).
- Girardeau, Bruno, Caille, Agnes, Amelie, Le Gouge, Ravaud, Phillippe (2012), Participant Informed Consent in Cluster Randomized Trials: Review, *PLOS One*. DOI: <https://doi.org/10.1371/journal.pone.0040436>
- Giridharadas, Anand (2018), *Winners Take All*, The Elite charade of changing the world.
- Glewwe, P., Park, A., & Zhao, M. (2012). Visualizing development: eyeglasses and academic performance in rural primary schools in China. *Journal of Development Economics*, 122, 170-182.
- Kapur, Devesh. (2018), *The Study of India in the United States, India in Transition*, 29 June, <https://casi.sas.upenn.edu/iit/deveshkapur2018>
- Karlan, Dean and Udry, C. (2020), A Call for Structured Ethics Appendices in Social Science Papers, <https://www.poverty-action.org/blog/call-structured-ethics-appendices-social-science-papers>
- Khera, R. (2019a). Fashions, framing and expertise: ethical hazards for economists. *Journal of Global Ethics*, 15(1), 45-54.

Khera, R. (2019b), Viable solutions to PDS portability are being ignored in the push for Aadhaar, *Indian Express*, 28 November.

Kumar, S. M. (2016). RCTs for better policy? The case of public systems in developing countries. *Economia Politica*, 33(1), 83-98.

Marulasiddappa, M., Raonka, P., & Sabhikhi, I. (2014). Social Security Pensions for Widows and the Elderly. *Indian Journal of Human Development*, 8(1), 49-63.

Mukherjee, S. (2014). Smart Cards Cut Fund Leakage in MGNREGA by 12%: Study. *Business Standard*, 03 April, 2014. Retrieved from [Smart cards cut fund leakage in MGNREGA by 12%: Study | Business Standard News \(business-standard.com\)](https://www.business-standard.com/article/economy/smart-cards-cut-fund-leakage-in-mgnrega-by-12-study-1140401).

Lyall, J. (2020). Group Rights in an Era of Large-Scale Field Experiments, Working Paper.

Muralidharan, K. and Niehaus, P. (2017). [Experimentation at scale](https://doi.org/10.3386/w26744), *Journal of Economic Perspectives*, 31 (4), 103-124.

Muralidharan, K., Niehaus, P., Sukhtankar, S. (2017). "Direct Benefits Transfer in Food: Results from One Year of Process Monitoring in Union Territories." UC San Diego.

Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2018). We need a choice-based approach in the Public Distribution System. *Hindustan Times*, 03 October, 2018. Retrieved from [We need a choice-based approach in the Public Distribution System - columns - Hindustan Times](https://www.hindustantimes.com/india-news/we-need-a-choice-based-approach-in-the-public-distribution-system-columns-hindustan-times/).

Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2020). *Identity verification standards in welfare programs: experimental evidence from India*. National Bureau of Economic Research Working Paper No. w26744.

Oh, S. (2019). Does Identity Affect Labor Supply?. *Job Market Paper, Columbia University*.

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.

Pritchett, Lant (2018), The Debate about RCTs in Development is over. We won. They Lost. <https://www.youtube.com/watch?v=ct-lWYNEdFY>

Ravallion, M. (2014). Taking Ethical Validity Seriously. *World Bank Blogs*, 17 March, 2014. Retrieved from [Taking Ethical Validity Seriously \(worldbank.org\)](https://blogs.worldbank.org/development/taking-ethical-validity-seriously).

Ravallion, M. (2020). Highly prized experiments. *World Development*, 127, 104824.

Ruxter and Muldon (2019), Unethical work must be filtered out or flagged, *Nature*, 572, pp. 171-172. doi: <https://doi.org/10.1038/d41586-019-02378-x>

Sarin, A. (2019). Indecent Proposals in Economics. *The India Forum*, 07 November, 2019. Retrieved from [Indecent Proposals in Economics, The India Forum](https://www.theindiaforum.org/2019/11/07/indecent-proposals-in-economics/).

Sneha, P. (2020). RCTs for policy in India-Ethical Considerations, Methodological Concerns and Alternative Approaches (Working Paper). *IDFC Institute*.

Vivalt, E. (2020). Using Priors in Experimental Design: How Much Are We Leaving on the Table. In: *Randomized Control Trials in the Field of Development : A Critical Perspective*. Oxford Publications.

Yeh, Valsdotir, Yeh, Shen, Kramer, Strom, Secemsky, Healy, Domeier, Kazi, Nallamothu. (2018). Look before you leap. *BMJ* 2018; 363:k5094. <http://dx.doi.org/10.1136/bmj.k5094>

Ziliak, S. T., & Teather-Posadas, E. R. (2016). The unprincipled randomization principle in economics and medicine. *The Oxford Handbook of Professional Economic Ethics*, 423.