STUDYING THE IMPACT OF TRANSITIONAL JUSTICE ON DEMOCRACY: CONCEPTUAL AND METHODOLOGICAL CHALLENGES

Valerie Arnould¹ and Filipa Raimundo²

Paper prepared for presentation at the 7th ECPR General Conference
4-7 September 2013, Bordeaux

I. Introduction³

In recent years, there has been an upsurge of interest in studying the impact of transitional justice (TJ) on democracy. From its inception, TJ has been seen as closely intertwined with processes of democratisation. At its core, TJ was conceived as a measure to accompany and support transitions towards democracy. Its purpose was to provide tools to leaders in emerging democracies to deal with the past and build a new polity that integrated both victims and perpetrators and would be based on new, more democratic principles.⁴ Today, TJ is no longer confined to contexts of transitions out of authoritarian rule and many transitions did not in fact see the emergence of democracy. Nevertheless, the assumption that TJ is meant to ensure that countries undergoing a transition move towards democracy rather than another form of governance continues to underlie this field of study. It is therefore essential to undertake empirical assessments to investigate how TJ has fared in meeting its foundational goal of promoting democracy.⁵

A range of empirical studies, both qualitative and quantitative in nature, have been undertaken to test normative assumptions about the relationship between TJ and democracy. However, the contradictory and inconclusive nature of these findings highlights the need for further research. The focus of this paper is on discussing some key conceptual and methodological challenges associated with assessing the impact of TJ on democracy. We argue that in developing these studies scholars have tended to underestimate the importance of conceptualisation as well as variable operationalisation and measurement, both when using the aggregated and the disaggregated forms of the two concepts.⁶ This is an important issue in the social

---

¹ Valerie Arnould is a Postdoctoral Research Fellow at the School of Law, University of East London. The author can be contacted at V.Arnould@uel.ac.uk.
² Filipa Raimundo is a Postdoctoral Research Fellow at the Netherlands Institute of Human Rights (SIM), Utrecht University. The author can be contacted at F.Raimundo@uu.nl.
³ This paper forms part of a larger project entitled ‘The Impact of Transitional Justice on Democratic-Institution Building’, funded by the Economic and Social Research Council (ESRC) and the Netherlands Organisation for Scientific Research (NWO). The project is conducted jointly by researchers based at the Netherlands Institute of Human Rights at Utrecht University and the School of Law at the University of East London. More information on the project can be found at www.tjdi.org.
⁵ We recognise that democracy itself is also not a good per se and that it can be questioned whether setting this as a goal for transitional justice is desirable. While this is an important debate, it is one that will not be addressed here. At this stage, it suffices to point out that we chose to focus on democracy as the dependent variable not because we believe this is the most desirable outcome for transitional justice but because it is one of the main claims put forward by scholars and policy-makers to justify the implementation of transitional justice measures.
⁶ As we shall see bellow, by disaggregated forms of the concept of democracy we mean individual attributes such as respect for human rights or political participation; by disaggregated forms of the concept of transitional justice we mean looking at the characteristics of individual mechanisms such as trials or truth commissions.
sciences and, in fact, the problems with conceptualisation, variable operationalisation and measurement encountered by TJ researchers are not significantly different from those in other fields of research. Twenty years after the emergence of TJ as a field of research, scholars acknowledge that ‘even studies using the same data collection and analysis method (for example, large-n regression analysis) arrive at divergent results due to differences in their conceptualisation of key variables, codification and collection of data, and model specification’.7

This paper explores a number of methodological challenges facing medium-N qualitative studies (i.e. studies that use more than four or five cases while applying qualitative methods), tackling the issues of proper conceptualisation, adequate variable operationalisation and measurement. Such medium-N studies pose specific methodological challenges related to their ability to prove causality – since the size of the N is not large enough to run statistical analysis and their goal is usually not suited for experimental design8 – or the adequate degree of in-depth analysis that is both desirable and achievable – since the size of the N raises doubts about the type of data collection methods to use and the level of detail and prior knowledge the researcher should have.

From a methodological point of view we suggest there are four questions which are particularly important for studying the relationship between such complex and aggregate concepts as TJ and democracy. First, how can we deconstruct the concept of democracy in a meaningful and measurable way that speaks to the specialised literature and adequately relates it to TJ? Even if the researcher is unable to prove causality in a definite manner, an adequate conceptualisation and operationalisation of the dependent variable is crucial for the type of causal links he/she will be able to suggest. Second, how can we assess impact by exploring and explaining the intermediate steps that link the dependent and the independent variables (what we define as pathways of impact)? If a causal link exists, the researcher should be able to identify and describe those intermediate steps. Medium-N analysis can offer an added value here as they are better suited than large-N studies to investigate these pathways of impact. Third, how can we take into account variations within TJ mechanisms (TJM) without getting lost in the detail, allowing for comparison across a wide range of cases while avoiding a simple dichotomous approach to TJMs as being either present or

8 The experimental design is thought to be one of the best methods to prove causality. By artificially creating distinctive groups and controlling for the exposure of these groups to a certain stimulus, the researcher is able to prove, with a considerable degree of certainty, whether a particular factor is responsible for a certain outcome. In the case of TJ mechanisms (TJM) implemented at the state level, the researcher does not have the ability to artificially create those distinctive groups (democracies) and to control for their exposure to the stimulus (transitional justice). At best, the experimental design can be used if the researcher is interested in assessing impact at the individual level. For more on this see for instance Gibson, James, “Does Truth Lead to Reconciliation? Testing the Causal Assumptions of the South African Truth and Reconciliation Process”, American Journal of Political Science 48 (2004): 201-207; Choi Susanne Y.P. and Roman David, “Lustration Systems and Trust: Evidence from Survey Experiments in the Czech Republic”, American Journal of Sociology 117 (2012); or Martin-Beristain, Carlos; Paza, Darío; Rimé, Bernard, and Kanyangara, Patrick, “Psychological effects of participation in rituals of transitional justice: A collective-level analysis and review of the literature of the effect of TRCs and trials on human rights violations in Latin America”, Revista de Psicología Social, 25, no. 1, (2010): 47-50. Or, as Brahms suggests, the researcher could run a quasi-experimental design using two countries, where one would have adopted a certain TJM and the other not, although finding two countries that are similar in almost all relevant variables expect for one - the implementation of a TJM – is particularly challenging. See Wiebelhaus-Brahm Eric, Truth Commissions and Transitional Societies: The Impact on Human Rights and Democracy (New York: Routledge, 2010).
absent? While small-N studies tend to go into too much descriptive detail about individual TJMs, large-N studies tend to overlook such important distinctive characteristics. Using a medium-N allows the researcher to explore if and how the particular characteristics of TJMs (such as a difference in their institutional design or mandate) may condition the outcome. Fourth, how can we deal with different types of impact and what to realistically expect from the relationship between TJ and democracy? Knowing the type of effect TJ may have on democracy is crucial to formulating the right hypotheses. This also involves knowing what can be captured when studying immediate or long-term impact.

The paper is organised as follows. The first section provides an overview of the literature on TJ impact focusing in particular on those studies which examine the link between TJ and democracy, including those focusing on individual dimensions and attributes of these concepts. The following section elaborates on four methodological issues, which the authors believe are at the basis of a number of problems and shortcomings faced by existing studies, such as contradictory findings, dubious causal links, conclusions based on normative assumptions, and insufficient attention paid to alternative explanations. Finally, the last section will provide some brief conclusions.

II. What We (Don’t) Know About Transitional Justice Impact

The growing recourse to TJ to address legacies of mass human rights abuses has provoked an intense debate about its effectiveness. While scholars initially focused on understanding how these mechanisms work and what factors impede or facilitate the pursuit of TJ, there has been a clear shift over the past years towards the question of impact. Now that TJ has moved out of its infancy - a period marked by high hopes and bold, some would say idealistic, claims about its ability to help transitional countries build stable and peaceful polities – scholars and practitioners are becoming increasingly critical of the performance of these mechanisms. While the pursuit of accountability for mass human rights violations can be justified on moral and legal grounds, there is little agreement about the political and societal need and desirability of pursuing TJ. There has been growing scepticism whether TJ as a project is realistic, desirable and effective. In less than two decades, discourse has shifted from high-minded proclamations of faith in TJ’s transformative powers to disillusions with a project that stands accused of being overly idealistic, largely detached from victims’ needs and ineffective at radically overhauling political processes and societies. Strong disagreements about the benefits or costs of TJ can be found both in the normative literature and more recent empirical studies on TJ impact. One explanation for these disagreements is that a wide variety of goals have been attributed to TJ, such as retribution, justice, accountability, democracy, peace, redress for victims, deterrence, truth telling, reconciliation, healing, rule of law, and the removal of perpetrators.\(^9\) The result is that impact studies assess TJMs in light of very different outcomes and

consequently arrive at differing, sometimes contradictory, conclusions about their effectiveness. However, even studies looking at a similar outcome, such as democracy, come to varied and contradictory conclusions about TJ’s impact.

The relationship between TJ and democracy is complex and multifaceted, making it difficult to unravel and conceptualise their interactions. Both are ‘closely related but somewhat independent processes whose courses of action tend to follow their own logic’ and are therefore as much capable of reinforcing each other or being in conflict.\textsuperscript{10} Views diverge widely amongst scholars about whether TJ positively contributes to democratisation processes. Proponents of TJ espouse the view that TJ is necessary for democracy because it strengthens the rule of law, deters future human rights abuses, promotes reconciliation and prevents retributive violence and vigilant justice.\textsuperscript{11} The absence of justice, it is argued, in turn promotes impunity and hampers the reconstruction of civic trust and state legitimacy. By reshaping institutions, actors, identities and social behaviour – that is, by redefining the social and political order - TJ measures have the potential to positively contribute to building democracy. They are purported to do so in a variety of ways, ranging from institutional reforms and educating citizens and elites on democratic norms, to restoring relations between previously antagonistic groups, encouraging open debates about the past and thereby constructing a new inclusive political and social memory, setting out responsibilities and punishment for human rights violations, encouraging state responsiveness to the needs of victims, restoring citizen’s trust in public institutions, and providing an opportunity for new leaders to signal a break with past non-democratic practices.

These claims about the far-reaching positive impact of TJ have been contested by scholars of a more pragmatic bend. Some caution against making TJ a pre-requisite for democratisation as many paths lead to democracy and there is no pre-ordained outcome of any effort at TJ because decisions on the latter are guided by politics rather than ethics and law.\textsuperscript{12} Critics argue that TJ can only support democracy if it is sensitive to the political realities in which it operates. Thus TJsMs, and in particular retributive mechanisms such as trials, should only be pursued if political conditions are right. That is, as long as spoilers linked to anti-democratic forces have the ability to upset the democratisation or peace process, TJ should be pushed aside and amnesties favoured. Failing this, the pursuit of justice poses a risk of destabilising a democratisation process, increasing human rights abuses, heightening political and societal tensions, and

---


hampering negotiations for peace agreements and political transitions. These arguments have come particularly to the fore in debates about the International Criminal Court’s investigations in Sudan and Kenya, which have been accused of lacking sensitivity to delicate domestic political conditions and therefore hampering rather than supporting pacification efforts. While proponents of TJ recognise the constraints that politics may impose on TJ, they counter-argue that TJ policies can in turn constrain and mould the space in which political actors operate – by removing them, delegitimising them or imposing additional costs on them - and thereby limit the latter’s ability to operate as spoilers.

Pragmatists question whether TJ measures are able to achieve the goals set out for them, even when implemented in propitious contexts. Trials and truth commissions have been criticised for failing to contribute to democracy through truth-telling as they are likely to produce distorted or partial historical understandings of the past, either due to procedural constraints or political instrumentalisation by elites.

Restorative justice mechanisms, such as truth commissions or memorials, can also be harmful to democracy as their focus on building a new collective memory can revive or aggravate societal tensions, while their lack of actual punishment of perpetrators may provoke resentment, vigilante justice, and weaken trust in public institutions. Serious doubts have also been raised about the ability of trials to improve human rights practices and reduce the incentives for would-be perpetrators to commit abuses through the exercise of specific or general deterrence. Lustration in turn is rejected by some on the basis that it promotes a collectivisation of guilt, while, somewhat ironically, trials have been criticised for shielding state and societal complicity in crimes by focusing on individual criminal responsibility and thereby preventing deeper societal changes that are needed for democracy to take root.

---


that vetting and truth commissions can harm the rule of law when they themselves are characterised by procedural unfairness and selectivity.\textsuperscript{18}

The problem with the assertions made by both proponents and critics of TJ is that they are highly normative but are presented ‘as empirical fact when they are merely untested hypotheses’.\textsuperscript{19} Recent empirical research has started investigating the validity of the normative claims about TJ’s impact, but the findings from these studies are inconsistent. Some studies find that TJ has a positive impact on democracy by resulting in improved rule of law standards, reducing human rights abuses, heightening civic trust, promoting reconciliation and generally increasing the level of democracy.\textsuperscript{20} Other studies however produce more mixed findings or conclude that TJ measures have had no impact on democracy as they fail to result in improved human rights practices, a reform of political institutions or bring about societal reconciliation.\textsuperscript{21} These contradictory findings can partly be explained by the fact that TJMs are unlikely to be successful across the board and that their degree of success is strongly conditioned by the context in which they operate. Different impacts are thus observed across different case studies. However, even researchers working on a same country have come to conflicting conclusions about TJ’s impact on democracy. South Africa provides a glaring example of this. Some studies find that the truth-telling process supported reconciliation by allowing individuals and various political groupings to present their versions of events, by leading people to acknowledge that others had also been victimised and leading the former apartheid political party NP to acknowledge responsibility for and distancing itself from apartheid.\textsuperscript{22} Others however, argue that the truth-telling process made a limited contribution to reconciliation as it failed to properly address the institutionalised racism and structural injustices which underpinned the apartheid regime, the limited testimonies of abusers created frustrations amongst


victims, and the truth commission did not significantly improve interracial trust. Others also contend that the truth commission had a limited impact on reforming institutions to prevent human rights abuses or increasing support for democracy amongst citizens.

More recently, scholars have undertaken medium and large-N studies to test whether general trends can be identified in terms of TJ impact, which are valid across national and regional contexts. However, these studies have also failed to produce consistent findings. A study by Sikkink & Walling looking at 19 Latin America countries finds no evidence of the claims put forward by TJ sceptics that trials have incited coups, increased violence or undermined democracy. They also observe that the highest improvement in human rights practices was found in countries that had more trials and those that combined trials and truth commissions. Other studies similarly find that TJMs such as trials, truth commissions and lustration are positively correlated to improvements in rule of law, human rights practices, level of democratisation and public trust. In contrast, the largest cross-country study undertaken to date finds that TJ has a positive impact on democracy, only if trials are combined with amnesties (with or without truth commissions). No TJM on its own seems to be likely to positively impact democracy. This, they argue, is because a combination of trials and amnesties is best able to produce an appropriate balance between accountability and stability, and allows for a sequencing of TJ policies. These findings are corroborated by Wiebelhaus-Brahm’s study looking at the impact of truth commissions, which concludes that they are negatively correlated with human rights practices and have no impact on democratic developments, and Meernik et al.’s study which finds no evidence that international trials positively impact human rights.

These contradictory findings about TJ’s impact may reflect differences in research approaches between quantitative and qualitative studies, as well as differences in how quantitative studies structure their statistic models. They are therefore not an insurmountable problem or even necessarily a negative reflection of current research. Indeed, the interdisciplinary nature of the field of TJ and the multi-faceted


character of the processes it studies means that approaching a research question from a variety of
disciplinary and methodological angles is highly instructive, and even necessary. For instance, better
understanding the effects of TJ on democracy requires an investigation of its impact at the state, society
and individual level – each of which necessitates the use of different methodological tools – and an
inquiry into whether TJ is more likely to produce a positive effect at one or the other level. Variety in
research findings about TJ impact can thus be complementary rather than contradictory. However, a
review of the existing literature has highlighted a number of cross-cutting methodological challenges
which have so far been poorly or insufficiently addressed. In the next section we present four such
challenges, which are centred on issues of conceptualisation, operationalisation and theorisation. Our aim
is to highlight methodological concerns valid across various research approaches which, in our view, are
central to improving the validity and comparability of studies on the impact of TJ.

III. Conceptual and Methodological Challenges for Studying Transitional Justice Impact

Although impact assessment is a fairly recent development in the field of TJ, there already exists great
variation in the approaches taken to it. Few authors have consciously reflected on how they choose to
define what TJ success or failure means, even though how impact assessment is done strongly conditions
the nature of the findings. On the basis of the existing TJ scholarship we can identify three broad ‘types’
of impact assessments: procedural assessments, mandate-based assessments and goal-based assessments.30
Procedural assessments are focused on assessing the efficiency of TJMs. They evaluate TJMs on the basis
of their procedural performance, such as the cost, the pace of proceedings, their ability to produce
reports/rulings, gain custody of accused, or the degree to which victims have been allowed to participate
in proceedings. Assessments may also be done on the basis of the fairness of proceedings, in particular
the degree to which a TJM conforms to due process principles and has functioned in an impartial
manner.31 Procedural assessments thus primarily aim to assess how a mechanism operates – its
institutional design, procedures and powers – and whether it conforms to ‘best practices’, rather than
focus on the outcomes produced by it.32

30 This classification presents similarities with Geoff Dancy’s categorisation of process-based evaluation and outcome evaluation,
though it does not follow his conceptual distinction between evaluation and impact assessment. Dancy Geoff, “Impact
Assessment, Not Evaluation: Defining a Limited Role for Positivism in the Study of Transitional Justice”, International Journal of
32 It should be noted though that what constitutes ‘best practices’ with regards to transitional justice is still ill-defined, particularly
outside of the context of trials. The ‘Rule of Law Tools for Post-Conflict States’ developed by the Office of the United Nations
Mandate and goal-based assessments, in contrast, are focused on investigating the outcomes produced by TJMs; however they each do so in slightly different ways. **Mandate-based assessments** evaluate TJ on the basis of their ability to fulfil their mandate. This can be the mandate specifically ascribed to them by their creators as reflected in their constitutive documents or the general mandates identified by the TJ scholarship. The latter most commonly includes truth-telling, accountability, redress for victims, institutional reform or individual healing. **Goal-based assessments**, in turn, look at the extent to which TJMs have been successful at producing broader societal transformations, such as reconciliation, deterrence, democracy, civic trust, a human rights culture, or peace. Their aim is to identify to what extent TJ has brought about changes in attitudes, behaviours and institutional practices which enable it to contribute to macro-social goals rather than TJ specific goals.

A number of methodological challenges currently faced by TJ impact studies relate to the lack of dialogue between these different approaches. At present, a lot of disparate findings can be gleaned from the TJ scholarship on what TJ does and what outcomes it produces. However, because assessments are done on such varied basis (the most obvious reflection of this being the wide variety of dependent variables that are used across studies) it has proven difficult to draw comparative and cumulative conclusions. Our view is that advancement in the field of TJ impact assessment will not merely be achieved by refining the methods that are used, but could greatly benefit from trying to bring together the various strands of existing research. Procedural-, mandate- and goal-based assessments focus on different aspects and effects of TJMs, and therefore inevitably produce different findings, sometimes even when looking at a same TJM. However, a degree of complementarity exists between them which should be further explored. Focusing on the interstices and connections between findings produced by the three approaches to impact assessment is likely to improve coherence and comparability between research findings, as well as develop a more refined understanding of what TJ really does.

First, there is the need to bring together procedural and outcome assessments. The former are useful for telling us whether a TJM has functioned well, and is therefore an important tool for undertaking specific

---


project assessments or for investigating institutional reforms that policy makers should pursue.\textsuperscript{35} Outcome-focused assessments, in contrast, are more appropriate tools for investigating the effectiveness of TJ. Nevertheless, such studies can benefit from integrating procedural-based assessments in order to not only establish if a TJM is effective but also under what conditions it is most likely to be so. Thus one way of building a bridge between current research findings is by jointly investigating how TJ contributes to achieving a certain objective and whether particular characteristics of a TJM conditions its ability to achieve this goal. As we discuss below, treating TJMs as single units obfuscates the fact that the unique features of a mechanism may be key in explaining variations in impact. Bringing together observations about how a TJM works (their institutional design, mandate and mode of implementation) and the outcomes it produces may well prove central in improving our understanding of its impact.

A second way of bringing together the different approaches to impact assessment consists of building bridges between mandate- and goal-based assessments. Greater effort should go into investigating whether the specific mandates assigned to TJ are effective at producing broader societal changes. For instance, a truth commission may have been relatively successful at producing recommendations on institutional reforms and create pressures on the government to implement some of these reforms. However, it remains an open question whether the latter have actually changed behaviour in a way that promotes democracy or changes in human rights practices.\textsuperscript{36} In turn, researchers should reflect on whether the macro-social goals ascribed to TJ – which it should be noted, are largely founded on normative claims and aspirations – are realistic in light of the mandates that are given to these mechanisms. Better understanding the connection between TJ’s specific mandates and macro-social goals requires that close attention be paid to the conceptualisation and operationalisation of the dependent and the independent variable, the identification of pathways of impact, and the timeframe adopted for analysis. We argue below that both qualitative and quantitative studies have shown deficiencies on these fronts.

1. Conceptualisation of the dependent variable

The methodological rigour required for impact assessment demands a clear clarification from the outset of the question ‘impact on what’. While existing studies generally state what their dependent variable is (i.e. the phenomenon on which they expect TJ to have an impact), they don’t always pay close attention to its conceptualisation and operationalisation.\textsuperscript{37} This problem is particularly acute in qualitative studies focused on outcome assessment, as they often fail to define the dependent variable clearly or do so in a

\textsuperscript{35} The Assembly of State Parties of the International Criminal Court, for instance, set up a ‘Study Group on Governance’ in October 2012 with the mandate to carry out a procedural-based assessment of the court’s operations to ‘strengthen the institutional framework and improve its efficiency and effectiveness’, looking in particular at means to expedite criminal proceedings before the court.

\textsuperscript{36} Wiebelhaus-Brahm Eric, \textit{Truth Commissions}, 11, supra n.6.

\textsuperscript{37} Conceptualisation in its broadest sense means defining the meaning of a concept while operationalisation refers to the identification of sub-variables or indicators that allow for the empirical measurement of the concept.
subjective manner. Quantitative studies usually fare better when it comes to operationalising the dependent variable, but face particular challenges in their measurement of variables and pay insufficient attention to conceptualisation as they rely on indices where the concept is predefined. Defining concepts such as democracy, reconciliation, truth or healing is notoriously difficult as they are abstract, ambiguous and amorphous and therefore subject to a variety of interpretations.\textsuperscript{38} It is possible for two studies looking at, for instance, the impact of truth commissions on reconciliation to conceptualise and operationalise reconciliation in very different ways.\textsuperscript{39} Where for some reconciliation will be achieved when victims and perpetrators are able to live side by side within the local community on the basis of trust and mutual acceptance, for others reconciliation means political tolerance and intergroup cohabitation. The difficulty of qualifying processes such as democracy and reconciliation does however not exonerate researchers from engaging in the exercise of defining and operationalising them.

When dealing with concepts such as democracy or reconciliation the challenge lies with their multifaceted nature. Democracy, for instance, involves different levels of agency (state, society, communal, individual), different practices (institutional performance as well as the attitudes and beliefs of elites and citizens) and different spheres of action (political, economic, judicial, social, cultural). Furthermore, there is a lack of agreement amongst scholars about the procedural and substantive requirements of democracy, that is the conditions that need to be filled for democracy to effectively exist and which enable us to distinguish a formal from a substantive democracy. While minimal definitions of democracy postulate that democracy exists if competition for power is managed through regular and free elections and rule of law prevails, maximalist definitions consider that democracy requires the realisation of social justice and a market-based economy. In the absence of a single, authoritative definition of democracy, it may be more fruitful to focus on forming what Adcock and Collier term a systematized concept: a specific formulation of a concept that is linked to the goals and context of the research undertaken.\textsuperscript{40} This will usually involve focusing on particular dimensions or components of the concept rather than trying to build a comprehensive definition of the concept. A first step therefore in the conceptualisation of contested concepts is the identification of units of analysis.\textsuperscript{41}

In the case of the concept of democracy, initial questions that could be asked are ‘what level of impact should I look at’ or ‘which sphere of action should I study’, in order to capture different forms of TJ impact. For instance, investigating the state-level or individual-level impact of TJ on democracy will


\textsuperscript{40} Adcock Robert and David Collier, “Measurement Validity: A Shared Standard for Qualitative and Quantitative Research”, \textit{American Political Science Review} 95 (2001): 530-532.

require focusing on different dimensions of democracy – and therefore a different conceptualisation of
democracy – as well as the use of different research methods. Identification of these units of analysis can
occur in different ways. One can, for instance, choose to identify the different groups of actors whose
behaviours and attitudes matter to democracy, making a distinction between elites and the population, or
between the judiciary, the executive, political parties and the security forces. Alternatively, units of analysis
can be based on what the democracy literature has identified as core democratic principles, such as rule of
law, participation, competition, accountability and civic culture. Studies can chose to focus on one
particular level of analysis or engage in a multilevel analysis, allowing for a comparative investigation of
TJ’s contribution to varied dimensions of democracy. This will ultimately produce not only a more fine-
grained understanding of TJ impact but also more useful policy-relevant findings by pointing towards
fields of action where TJ is most likely to be effective. Once choices are made with regards to the unit of
analysis, more fine-grained conceptualisation and operationalisation can be undertaken. For instance, if
the unit of analysis is the judiciary, we can draw from the democracy literature an approximate definition
of what a democratic judiciary means (independent, impartial, respecting due process) and start reflecting
on how changes on these fronts can best be measured and traced over time.

One of the main criticisms levied against TJ is that the expectations that have been placed on it are
unrealistic. While we do not believe it is unrealistic to assume that TJ may have an impact on democracy
or reconciliation, and that it is worth empirically investigating to what degree this is true, it is evident that
TJ is unlikely to impact all dimensions of democracy. One of the problems with existing impact
assessment studies is that they often do not formulate realistic conceptions of the outcomes TJ can
produce, and therefore also use unrealistic measurements for TJ’s impact. When conceptualising and
operationalising variables the focus should be on those factors which we can reasonably expect TJ to have
an impact on. This has posed particular challenges for quantitative studies. The tendency within
quantitative studies has been to define democracy as a single unit, rather than focus on particular
subcomponents. Most have consequently relied on cross-national indices that provide aggregate scores
for democracy, such as Polity IV or Freedom House, which include components or attributes of
democracy on which TJ is unlikely to have an impact. As Brahm points out: “Polity’s measure focuses
generally on the degree to which political institutions both facilitate and constrain open political
competition and the free and fair adjudication of disputes in a court of law. Truth commissions, though,
have contributed little to the construction of the postconflict political system such as electoral rules and
checks on executive power”. Similarly, the Cingranelli-Richards Human Rights Data Project (CIRI),
which is often used for measuring improvements in human rights, aggregates 15 types of human rights,
not all of which TJ is likely to have an impact on, while the World Bank’s World Governance Indicator

---

43 Exceptions are Choi and David, “Lustration Systems”, supra n.6; Horne, “Assessing Impact”, supra n. 18; Weiffen, “The
forgotten factor”, supra n.18.
25.
on rule of law measures an extremely broad set of components of which many are unrelated to TJ such as contract certainty, property rights or the incidence of organised crime.\textsuperscript{45}

The result is that measurements used for the dependent variables are sometimes ill-suited to assess the impact of TJ, as they overestimate the factors on which TJ can have an impact. Moreover, when building their democracy indexes, indices sometimes include data on TJMs (such as human rights trials). In such instances, the finding that TJ contributes to democracy becomes tautological. At the same time, indices on democracy may exclude other hard to quantify components of democracy on which TJ is likely to have an impact, such as civil-military relations.\textsuperscript{46} This poses a significant problem in terms of measurement validity but also highlights a problem inherently faced by quantitative studies in studying the impact of TJ: they rely on indices where definitions of the dependent variable are predefined and circumscribed by the fact that they can only include indicators that are quantifiable. While this facilitates comparisons across studies, as studies using the same index rely on a same conceptualisation of the dependent variable, the definition they adopt of complex social processes such as democracy tends to be overly technical and exclude other dimensions of the phenomenon. They are consequently likely to fail capture some of the ways in which TJ contributes to democracy.

The reliability of the findings produced by large N-studies thus also suffers from poor conceptualisation and operationalisation, which cannot readily be addressed by improving statistical models. Rather, like qualitative research they can benefit from a more critical reflection on the conceptualisation of the dependent variable and on how existing measures are appropriate for the task at hand. Coppedge and Gerring, for instance, express strong doubts about the adequacy of existing democracy indices for evaluating the effectiveness of efforts at promoting democracy.\textsuperscript{47} When indices enable researchers to access disaggregated data on the various components, as is the case with CIRI, the problem is less acute. However, where indices do not allow for disaggregation serious questions can be raised about their adequacy for operationalising the dependent variable. Where large N-studies do rely on such indices, they could consider using alternative indices alongside it to test findings. Quantitative studies which rely on existing democracy indices could also be complemented by qualitative studies that focus on dimensions of democracy that are not included in the indices.

2. **Exploring pathways of impact**

Closely related to the problem of conceptualisation is the issue of defining the pathways through which TJ impacts macro-social processes such as democracy. As many researchers have pointed out, proving


\textsuperscript{46} Civil-military relations are, for instance, not included in the Democracy Barometer, the Bertelsmann Transformation Index and Freedom House.

empirically that TJ causes democracy is ‘a Sisyphean task [as] stable democratic practices [...] are the product of many simultaneously operating processes’. The same can be said of TJ causing reconciliation, individual healing or improved human rights practices. One way of dealing with this problem is by peering into the black box rather than focusing on the macro-process as a single unit. As we discussed earlier, researchers should focus on identifying the various dimensions of the phenomenon under study and develop realistic conceptions of the dimensions of the dependent variable which TJ is able to impact. The question is therefore not whether TJ causes democracy but whether it contributes to reshaping institutions,behaviours, identities and relations (between individuals, and between individuals and the state) which underpin democratic developments. This requires a clear understanding of the processes through which TJ produces change; that is an identification of the pathways of impact.

The mandates and goals that are presently ascribed to TJ have largely developed from normative claims and aspirations – they identify what we hope TJ can achieve rather than what we know they achieve. Advocates of TJ make claims about their positive effects – such as reconciliation, healing, deterrence – on the basis of untested, contentious or misconstrued assumptions. Critics of TJ similarly make claims about the potential harmful effects of TJ – such as provoking political instability or undermining peace talks – on the basis of mere assumptions. As Mendeloff observed, most of these assumptions are neither factually nor theoretically founded and “rest far more on faith than sound logic or empirical evidence”. Empirical studies have sought to test the validity of some of these assumptions but with mitigated success. The reason for this is that they themselves have relied on these flawed assumptions and poor theorisation. Consequently, they fail to establish clear links between the outcomes produced by TJ and its effects on broader social and political processes.

Impact studies looking at TJ’s impact on democracy often build their hypotheses or questions on observations about what TJ does - such as fact-finding, creating a shared history, reforming institutions, removing perpetrators from positions in government institutions, giving a voice to victims – without due attention paid to what this means in terms of democracy. This reflects the problem we highlighted earlier of a poor conceptualisation of the dependent variable but also the lack of ‘dialogue’ between mandate- and goal-based assessments. Because we don’t know whether truth-telling does effectively contribute to democracy, we cannot take the measure of a TJM’s success in fulfilling its truth-telling mandate as a measure of its positive impact on democracy. To truly measure TJ’s impact one would need to ascertain (a) how it performed in terms of truth-telling and (b) whether this truth-telling positively or negatively impacted dimensions of democracy such as political participation and civic trust. Indeed, truth-telling can help strengthen democratic values and diffuse conflicts about the past but it can just as likely contribute to entrenching victimisations and political polarisations. Similarly, the removal of perpetrators or

---

collaborators to the previous regime from state institutions can help restore civic trust in institutions or enhance political animosities.

Researchers should thus pay closer attention to investigating the links between the mandates of TJ, their direct outcomes, and the broader goals pursued through it. This can be done through a deductive approach: based on the literature, build theories about which dimensions of democracy truth-telling might have an impact on and how, and then systematically test these. However, an inductive approach might be more suited for the purpose of identifying pathways of impact. Because of the normative nature of the assumptions about TJ, and the fact that it is likely to exert impact in an indirect manner and produce unintended consequences, impact studies might be advised to refrain from building their hypotheses around what TJ does. Guiding questions could instead be formulated around the dimensions of the dependent variable which TJ may impact. From there, it is then possible to observe through which means TJ has impacted these, and establish empirically whether and how truth-telling contributes to producing broader societal transformations or whether it has other, previously, unidentified ways of producing change. Or, to the contrary, we may observe that the mandates which have so far been attached to TJ are, in fact, largely ill-suited to producing the hoped for societal transformations.

3. Disaggregating TJMs

While our discussion so far has focused on the importance of conceptualising the dependent variable, we also believe that a revised approach to conceptualisation of the independent variable (i.e. TJMs) is needed. One of the possible explanations for the contradictory findings produced by existing studies is that they tend to treat TJMs as a single unit. No single truth commission or amnesty, however, is similar to other truth commissions or amnesties but instead present distinctive characteristics. In their large-N study, Olsen, Payne and Reiter acknowledged that ‘given the widespread adoption of TJ to address a range of past violence, the scope of these mechanisms varies dramatically across cases’. Unfortunately though, the authors did not take this ‘dramatic’ variation into account when building their database. Neither did they reflect on how such ‘dramatic’ variation could account for differences in the observed outcome. Instead, the authors coded each year in which the TJM was in existence for each country – which gave them the possibility to generate panel data to allow for measurement over time – but classified each TJM in five broad categories (trials, truth commissions, amnesties, reparations and lustration) – which prevented them from taking variations within TJMs into account. As a result, the study is unable to investigate whether such ‘dramatic’ variation has an effect on TJMs’ ability to strengthen democracy and, more importantly, whether taking this into account would have led the authors to arrive at different conclusions about TJ impact.

50 Olsen, Payne and Reiter Transitional Justice, 2, supra n.25.
51 The authors seem to acknowledge this shortcoming when they suggest that ‘later research could begin to create meaningful typologies of truth commissions to capture their variation’, 35.
Knowing whether differences in the individual characteristics of TJMs conditions the nature of their impact on democracy is particularly important as several studies have highlighted that there exist great differences in how TJMs are designed, what their true goals are, and how they operate.52 In her oft-quoted study on truth commissions, Priscilla Hayner mentions several aspects that, according to the author, can influence the outcome, and consequently, the impact such commissions can have on democracy. She, for instance, argues that ‘perhaps more than any other single factor, the person or persons selected to manage a truth commission will determine its ultimate success or failure’.53 The author suggests that truth commissions not characterised by a strong leadership are likely to fail, as they are more likely to suffer from internal divisions, fail to mobilise the necessary funding, or be unable to complete their tasks. Similarly, a study by Taylor and Dukalskis suggests that the degree of ‘publicness’ of a truth commission conditions the degree of impact that it will have on democracy.54 As a result, lumping all truth commissions together into a single category without considering their institutional performance may result in missing a central explanation as to why they were (un)able to positively impact democracy. In a similar vein, Louise Mallinder shows how differences between amnesty laws may explain their ability to impact democracy or particular aspects of democracy: ‘if a dictatorial regime claimed it was amnestying its opponents to promote national unity, but its real aim was rather to entrench its own power, it is unlikely that the amnesty will result in a more harmonious society’.55 This does not mean that amnesties will never contribute to peace and ‘harmony’, but rather that studying amnesties across a wide range of countries may require that a distinction is made between, for instance, those amnesties that are approved by the autocratic elite and those that result from negotiations (and were therefore approved by the old and the new elite and/or included in a transitional agreement). The conclusion could be that amnesties approved as a result of negotiations are more likely to contribute to peace and ‘harmony’ than those resulting from an autocratic elite’s attempt at maintaining control over power.

In reflecting on the characteristics of TJMs which may influence their impact, researchers may find that certain instruments which are usually placed under one and the same broad TJM category should in fact be treated distinctly. Again using Mallinder’s example, one may want to distinguish between amnesties directed at political prisoners (i.e. the victims of the regime) and those directed at the former elite or fighters (i.e. the wrongdoers). The reason for this is that because they target radically different groups of actors their degree of societal acceptance or contentiousness will likely vary, which in turn is likely to condition the nature of their impact on peace, reconciliation, or trust in institutions. This being said, decisions about whether or not to include both kinds of amnesties in a study should be guided by the choice made with regards to unit of analysis, as discussed earlier.

The suggestion that disaggregation of TJMs matters for impact studies is corroborated by small-N studies – which have the added value of providing detailed information about the reasons why, under what circumstances, and with what consequences certain TJMs were adopted. The issue of implementation, beyond formal rules, has received particular attention in such studies. As one scholar argues: ‘For the case of Aceh the question of how the amnesty was granted is thus of particular significance. (...) loosely defined amnesties have the potential to cause significant problems in the post-settlement period. Thus, the inclusion of what appeared to be a blanket amnesty without any specified exclusion in the MoU, but which actually turned out to include some limitations, caused what may have been unnecessary tension in the implementation phase’.56 Thus it is not only important to consider how differences in institutional design and mandate may produce difference outcomes, but also how the manner of implementation conditions the impact a TJM will have.

While coding all relevant variations in TJMs in a large-N study could prove difficult – also because of the limited number of theories available to create testable hypotheses – a medium-N analysis enables the researcher to take such TJM variations into account by relying on a series of research sub-questions defined around TJM characteristics we expect to matter rather than working from rigidly formulated hypotheses. This, however, requires the formulation of a systematic method for collecting and analysing the data for each of these questions, taking into account the characteristics which, according to the literature, are likely to have a higher potential in explaining impact. At this stage, it is important to emphasise that we are not arguing that researchers should not build an a priori comparative framework and reasonably well defined research design. Quite to the contrary, while a medium-N allows the researcher to delve into the detail of TJMs, failing to arrive at a systematic means for collecting and organising such data would unduly limit the ability to compare across cases.

Ideally, a most similar/dissimilar comparative design should be used to test these differences across TJMs of the same type, comparing how they were designed, what influenced their implementation process, and what outcome they had. We suggest this should be done in three steps. An important first step is to ensure that the study relies on strong theories when building his/her hypotheses or expectations. Should we expect a truth commission that produces a partial and biased narrative of the past to have a positive effect on citizen’s trust in institutions? Should a set of trials resulting in impunity be expected to have a positive effect on a country’s human rights culture? Should reparative measures that compensate certain victims’ groups to the detriment of others be expected to have a positive effect on peace and reconciliation? The answer to all the above questions, even if counter-intuitive, may actually be yes. It may be that the mere fact that the government was willing to sponsor a truth commission gives citizens a positive sign of commitment and willingness to address past wrongdoings. It may be that the mere fact that trials took place, regardless of their outcome, represents a shift in how the state deals with criminals.

and protects human rights. It may also be that victims who have not been compensated in the first round of reparations recognise that economic constraints forces a state to pursue a progressive implementation of reparations, and that as peace and democracy develops further opportunities for compensation will emerge. What is important is that the researcher reflects on all these things and supports his/her choices with strong theories.

The second important step is to draw from the literature a systematic list of factors which may account for the type of variation that is likely to influence the ability of the TJM to impact democracy. Backer, for instance, highlights the importance of understanding the initiation, implementation, and impact when comparing TJ processes, while Skaar suggests focusing on the mode of establishment of the TJM and the manner of implementation.\(^{57}\) In a similar vein, we propose that researchers take into account the institutional design and mandate, the input, and the output of TJMs. By institutional design and mandate we mean TJMs’ formal characteristics, such as their jurisdictional scope, composition, mandate, powers and limitations. Because there is no guarantee that the formal characteristics will correspond to the actual practice of the TJM, it is important to also take into account other aspects beyond formal rules. In other words, the researcher should take into account the input or the external factors which may shape the functioning and implementation of the TJM. This includes, for instance, a mechanism’s resources, the timing of its adoption, the degree of societal participation and international involvement, as well as the actors involved in its establishment. All this should contribute to the final output produced by the TJM, which in itself is likely to influence the impact it may have on democracy. Indeed, the degree to which the outcome produced by a TJM is widely accepted by the elite and/or population is likely to strongly condition its impact on democracy. It is therefore important to take into account all the aspects which may shape the perception and the acceptance of the outcome, be it a report, verdict, law, apology, or monument.

Finally, a third step should be to try to create a typology of TJMs – something which Olsen, Payne and Reiter have also suggested.\(^{58}\) This would allow researchers to move from the current common system of classification of TJMs as single units to a more fine-grained level of differentiation of TJMs. To the extent that a typology is ‘a coordinated set of categories or types that establishes theoretically relevant analytical distinctions’,\(^{59}\) it would prove useful for a variety of purposes, in particular to test TJ’s impact on democracy. Where the literature offers enough evidence on the importance of particular characteristics of a TJM, it is possible to build a typology at the start of the study and then test the distinctive impact of each type on a particular aspect of democracy.\(^{60}\) Where the literature is still insufficiently developed, creating a typology could become the end goal of the research. This would constitute an extremely


\(^{58}\) Olsen, Payne and Reiter; Transitional Justice, 160-161, supra n.25.

\(^{59}\) Brady Henry and David Collier, Rethinking Social Inquiry (Lanham, MD: Rowman and Littlefield), 356.

\(^{60}\) An example of such a study is Choi and David, “Lustration Systems”, supra n.6.
valuable addition to the field of TJ. One possible way of doing this would be to summarise all variations into two major categories, such as efficiency (based on the design and input, such as composition, resources, timing, duration) and fairness (based on the output, such as level of bias, acceptance, follow up activities). The typology would then emerge from the cross-tabulation of these two major dimensions leading to four types of TJMs: failed TJM (not efficient and not fair), contested TJM (efficient but not fair), controversial TJM (not efficient but fair) and successful (efficient and fair). The challenge here is to arrive at a typology that not only serves the purpose of the research but also of a broader audience. Typologies are meant to help us make sense of reality. Although each researcher could, in theory, come up with his/her own typology, this field would gain significantly from having one or two strong and theory-grounded typologies that could prove useful for a variety of TJ studies.

4. Short-term vs. long-term impact

A great number of studies focusing on the impact of TJ on democracy have begun to explore what they claim to be the ‘long-term’ impact of TJ, yet there is little agreement on what sort of long-term impact one should expect TJ to have. More importantly, there has been little reflection on the methodological challenges of studying long-term impact. Studying the long-term impact of a certain policy or measure on institutions, behaviours and beliefs requires that researchers answer two important questions. First, what sort of impact do we expect the policy or measure in question to have and, more importantly, whether we expect it to persist over time or whether the nature of its impact may change over time? Second, what other factors, external to the factor we are analysing, may lead to changes in the same institutions, behaviours or beliefs? Failing to answer these questions poses the risk that the study will focus on a kind of impact that in reality does not persist and thereby overlook the real factors causing the change (or lack thereof) in the dependent variable. As Backer argues, ‘the quick bursts of activity on the TJ front could possibly cause an adjustment of attitudes, incentives, relationships, and practices, at least in the short term. Yet it is debatable whether such effects will be sustained over the longer term in the absence of other initiatives’.

In theory, any potential impact may be mitigated or enhanced with the passage of time depending, again, on the question ‘impact on what’. A possible way for dealing with this is to group the possible effects into two major types: the contextual and the structural aspects of democracy. Contextual aspects of democracy are essentially those attitudes previously mentioned by Backer, in particular trust in institutions. This is not to say that trust in institutions will never be influenced by the long-term impact of particular TJMs. What we argue is that the causal link is likely to be significantly more difficult to measure and prove, even in single case study, let alone in comparative analysis. Trust in institutions is a contextual phenomenon – associated with the political and economic context in which it is captured – and it tends to be heavily

---

dependent on the performance of said institutions. If this changes, levels of trust are also likely to change. Assuming that TJMs are likely to have an impact on how citizens trust institutions (such as the judiciary, the army, the parliament or political parties), its measurement, for instance through a survey, should occur soon after the TJM has produced its results (such as a law entering into effect, a report, a verdict, an apology, or a memorial). In reality, scholars rarely have the financial means to conduct their own survey and acquire the necessary baseline data or have the opportunity to take part in an existing survey that would allow them to establish that correlation, knowing with a degree of certainty that they can reject the null hypothesis.

Beyond the immediate implementation of a TJM it may be difficult, albeit not impossible, to ascertain that TJ is in fact responsible for certain levels of trust in particular institutions. Let us take the example of trials. One possibility would be to ask citizens if they believe human rights violations have been punished in their country. Ideally, this would be done by selecting a number of countries and dividing them into two groups: those that punished their wrongdoers and those that did not. Then scholars would ask citizens in both groups of countries if they believe human rights violations are punished in their countries. The hypothesis would be that citizens in countries where past human rights violations were tried before courts would be more likely to believe human rights violations have been punished, which could be seen as a proxy for their degree of trust in the judiciary and the rule of law. However, in this instance it would still be difficult to know if when people answer the question they are in fact reflecting on the TJM adopted in the past. What is more, the TJ literature suggests that TJMs tend to be adopted in a sequential manner, making it even more difficult for scholars to know which TJMs citizens have in mind, if any at all.

Structural aspects of democracy include beliefs, norms, and institutions characterised by authoritarian legacies. If we expect TJ to have a positive impact on these structural dimensions of democracy, it probably means we previously observed that their performance is not ‘perfectly democratic’, and as such is probably marked by the legacies of repression or any other legacies associated with the previous regime or conflict. Legacies can be defined as ‘the structural, cultural and institutional starting points (…) at the outset of the transition’. Possible examples include military reserved domains, low levels of political participation and different forms of undemocratic behaviour. We expect TJMs to have a positive impact on democracy by diminishing the presence of such legacies through the removal of military actors from positions of power, demonstrating political will and a commitment to democratic practices and therefore encouraging citizens to participate in politics, or demonstrating through practice that the three branches of government are equally committed to respecting human rights, the rule of law, and compensating the victims of previous human rights violations.

One possible way to test whether a particular TJM has had a prolonged impact on a particular dimension of democracy is to collect data on, for instance, levels of respect for human rights, the number of

---

individuals becoming members of human rights organisations, the inclusion of previously marginalised
groups in politics, etc. This data is likely to be available from a number of existing databases on various
aspects of democracy. This type of data would then allow the researcher to draw a graph to illustrate the
performance of that particular democracy dimension over the years. Then it should be reasonably easy to
locate in time the moment at which the TJMs produced their outcome (not the moment at which they
were initially implemented, but the moment at which a report was released, a monument was inaugurated,
a lustration law entered into effect, etc.). Whether the line in the graph grows (or declines) before or after
the date at which the TJM was adopted – and always relying on a strong theory – gives an indication that
a correlation may exist between the TJM and the observed change in the dimension of democracy, which
can then further be explored qualitatively. Following an analysis of the above mentioned pathways of
impact, the researcher should be able to build a reasonably solid argument about the likelihood that a
TJM has had a positive effect on the particular democracy aspect being studied.

IV. Conclusions

Although empirical studies on the impact of transitional justice on democracy have become increasingly
widespread, the findings produced by these studies have been contradictory and inconclusive. This paper
has presented a number of conceptual and methodological challenges linked to problems of
conceptualisation, operationalisation and theorisation, which we believe explain the limited reach of
existing research findings, focusing in particular on medium-N qualitative studies.

We explored four major issues: how to operationalise the dependent variable, how to identify pathways of
impact, how to disaggregate the independent variable(s), and how to distinguish between short-term and
long-term impact. In the first case, we suggested that conceptualisation should rely on a disaggregation of
the concept of democracy (i.e. identifying its key component elements in order to build units of analysis),
bearing in mind the goals of the research. Secondly, we suggested paying closer attention to the links
between the dependent and the independent variables. In order to move away from establishing simple
correlation, future research needs to focus on identifying the pathways through which transitional justice
impacts democracy. Thirdly, we suggested exploring the characteristics of TJ in three steps: departing
from strong theories, identifying the relevant factors, and building a typology. Finally, in the fourth case
we suggested distinguishing between structural and contextual phenomena before analysing the short or
long-term impact of TJMs on democracy or attributes of democracy.

We believe more medium-N comparative studies are necessary in this field as they allow the researcher to
overcome the insufficiencies of small-N studies (which tend to be excessively descriptive) and large-N
studies (which often suggest causal links based on weak theories and spurious statistical results).

64 Polity IV, Freedom House, Bertelsmann Transformation Index, Varieties of Democracy, Democracy Barometer, to give but a
few examples.