4 Research Design

This research design chapter is structured in four main sections. In the first, I discuss process-tracing as a method to trace causal mechanisms, and then I elaborate on my use of process-tracing in the context of a comparative case-study design. In circumstances where causal mechanisms are fairly well theorized in the literature, cross-case process-tracing helps to refine existing explanations while simultaneously contributing to their external validity (Section 4.1). Section 4.2 highlights the rationale for choosing my two cases and the implications that come along with this selection. I also reflect on the generalizability of my findings. In the section on operationalisation, I propose how the key concepts of my work can be measured with the help of specific indicators and other "traces of evidence" (Section 4.3). In the last section, I discuss issues of data collection and analysis. To provide an outlook, I rely on qualitative content-analysis in MAXQDA to analyze different types of data, including semi-structured interview transcripts, official documents, minutes of executive board meetings and parliamentary debates, as well as notes from participant observation.

4.1 Process-Tracing: Uncovering Causal Mechanisms

Process tracing has undergone a profound process of maturation throughout the last decades. In 1994, King, Keohane and Verba still claimed that process-tracing was little more than looking at intervening variables in an infinite regress toward ever shorter chains of causality (p. 86). In their edited volume, *Social Mechanisms: An Analytical Approach to Social Theory*, Hedström and Swedberg (1998) challenged this interpretation, clarifying that causal mechanisms could not be observed directly and hence differed from intervening variables. Rather, causal mechanisms were defined as social processes that link one event to another, the force that connects a cause to an effect (Hedström & Swedberg, 1998; Elster, 1998). Since George and Bennett published their important book "Case Studies and Theory Development in the Social Sciences" (2005), process tracing has enjoyed wider recognition as a method of rigorous qualitative social science research. More recently, important publications elaborated on the use of process-tracing in different research contexts (Beach & Pedersen, 2013) and issue areas (Bennett, 2013).

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

Today, shared standards evolved which moved process tracing from the realm of sophisticated narration into a recognized and rigorous method to tackle the complex problem of causality in social science research (Bennett & Checkel, 2015). Yet, the increase in studies using process tracing also led to different understandings of the term, leading to some confusion. To clarify my understanding of process tracing and to delineate it from alternatives, I briefly address the main underlying assumptions of my approach as well as its implications for data collection and analysis.

4.1.1 Underlying understandings of causality

In its most general terms, process tracing seeks to trace causal mechanisms. Thus, it assumes a certain nature of causality. David Hume famously argued that we cannot observe causal mechanisms directly (Hume, 1975). Instead, we should think of causes as regular associations (correlations) between X and Y, moving from correlation to causation by controlling for alternative explanations (Chalmers, 2013). Specifically, Hume held that three criteria need to be fulfilled to establish causality: 1) X and Y must be contiguous in space and time; 2) X has to occur before Y; and 3) X and Y are regularly correlated (Holland, 1986). In accordance with these criteria, social scientists tend to adopt a probabilistic understanding of causality. This understanding suits large-n studies, where researchers are interested in mean causal effects across a population of interest. As the causal relationship between X and Y is conceptualized as probabilistic, hypotheses are formulated accordingly (i.e., an increase in X tends to increase Y). Such a probabilistic understanding of causality seems appropriate if one is interested in mean causal effects of a given variable across a large-n population. In statistical models, so-called "error terms" account for the inherent randomness, nonlinear associations and feedback loops – in short, the complexity of our social world. Also, a probabilistic understanding of causality underlies an understanding of causal mechanisms as chains of intervening variables between X and Y (Gerring, 2011; King et al., 1994). According to this conceptualization, each intervening variable exists as an analytical unit in its own right, independent of the other intervening variables. Researching intervening processes may provide additional useful information to explain why a given X has caused Y, also allowing to increase a researcher's confidence that a given correlation is not spurious. Yet, such an approach begs the question, how a correlation between an intervening variable and a dependent variable came about. In short, a conceptualization of causal mechanisms as IVs neglects the causal linkages between these variables, thus grey-boxing the causal mechanism itself (Beach &Pedersen, 2013; Mahoney, 2001).

In contrast, the tracing of a causal process in a single case works with a deterministic notion of causality (Bennet, 2008). There has been a great deal of misunderstanding using the term "deterministic" in this context, so some clarification is due at this point. To put it up front, I agree with the notion that causality is never deterministic in the sense that a given cause will *always* lead to a particular outcome. Outside the laws of nature, which are by definition true (within their scope of validity), universal, and absolute (Davis, 1992), there appear to be no such laws in the social sciences. Consequently, all general statements about causality in a social scientific theory are probabilistic. What authors such as Collier (2010), Bennett (2008), Mahoney (2008), or Beach and Peder-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

sen (2013) referred to when using the term "deterministic," however, is this: Instead of examining the regular covariation between X and Y, process tracing studies seek to investigate whether X was a necessary and/or sufficient[1] cause of Y *in a particular* case" (Collier et al., 2010). Similarly, Mahoney (2008) argued that a probabilistic understanding of causality makes little sense if we are interested in a particular case. As he put it, "At the individual case level, the ex post (objective) probability of a specific outcome occurring is either 1 or 0," which means that "single case probabilities are meaningless" (pp. 415-416). This is the meaning of a deterministic understanding of causality – that the causal forces, the theorized mechanism and the outcome were either present, or not in the case under investigation. Importantly, this does not imply that only one path could lead to Y. In fact, tracing the mechanism of a given theorized process does not indicate that other factors were less important, let alone irrelevant. Neither does this understanding of causality assume that a mechanism, once triggered, necessarily produces the outcome. If scope conditions are poorly theorized and/or alter in the course of events, the mechanism may be interrupted. However, if rightly theorized, we should expect causal conditions to unfold *some* effect on Y, given that the scope conditions for the mechanism are in place. Between these causal conditions and the outcome, each part of the mechanism transmits "dynamic causal energy" (Beach &Pedersen, 2013). In line with these considerations, I adopt Bennet (2008) definition of causal mechanisms as "processes through which agents with causal capacities operate in specific contexts to transfer energy, information or matter to other entities" (p. 207). On his reading, a causal mechanism is thus composed of several parts, whereas each part of is composed of entities (or actors such as individuals, groups, or states) that engage in activities. Adopting a machine analogy, each entity can be thought of as a wheel, while activities transmit the causal force to the next entity, resulting in the movement of the whole (e.g., a tractor). In contrast to a definition of a causal mechanism as a series of intervening variables, the parts of a causal mechanism according to Bennet have no independent existence in relation to Y. Instead, "each part should be seen as an individually insufficient but necessary part of the whole" (Beach & Pedersen, 2013, p. 50).

4.1.2 Theory Testing Process Tracing

In their book *Process Tracing Methods* (2013), Beach and Pedersen distinguished between "explaining outcomes," "theory building," and "theory testing" approaches to process tracing (PT). In brief, *explaining outcome* PT is a very common method used to elucidate a particularly puzzling outcome. This approach to PT resembles historical scholarship most closely, as researchers seek to craft a minimally sufficient explanation of an event. In doing so, scholars proceed an iterative research strategy, typically combining different mechanisms in an eclectic fashion. As the aim is to explain an outcome, mechanisms are often case-specific and do not aspire to be generalizable beyond the case at hand. In contrast, *theory-building* PT is particularly suitable in situations where we know that a correlation between X and Y exists, but we are not sure about the causal mechanism

1 Without a necessary condition, the outcome would not be possible. A condition is sufficient, if it causes the outcome irrespective of other conditions.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

between them. Following this research strategy, scholars start with looking at the facts and then inductively infer the causal mechanism linking X and Y. While theory building PT partly overlaps with "explaining outcome" process tracing, the former seeks to build a mid-range theory containing a causal mechanism that is generalizable. In my work, I engage in *theory testing* process tracing, where the researcher deduces a theory from existing literature and then observes the empirical evidence to test whether a hypothesized causal mechanism was actually present in the case (George & Bennett, 2005). This third kind of process tracing shares the interest in theory rather than a particular outcome with theory building PT. Yet, it starts with theorizing the mechanism before looking at the facts, not vice versa. According to this variant, "we know both X and Y and we either have existing conjectures about a plausible mechanism or are able to use logical reasoning to formulate a causal mechanism from existing theorization" (Beach & Pedersen, 2013, p. 14). The ambition of a theory testing PT study is to test and potentially refine a theorized causal relationship that is generalizable to other cases with similar key characteristics.

4.1.3 Methodological Limitations

As all methods, theory-testing PT also comes with limitations. To begin with, the generalizability of the theorized mechanism to other cases is limited. The more complex the theorized causal mechanism and the more its unfolding is contingent on specific scope conditions, the less generalizable our results. Also, theory testing PT is "blind" for alternative processes happening at the same time. While process-tracing that seeks to explain outcomes scrutinizes different paths leading to an outcome, "theory testing PT" seeks to establish a maximum degree of confidence that a theorized mechanism was present (or not), while remaining agnostic about alternative mechanisms which were at work simultaneously. Theory testing PT thus cannot test the relative explanatory power among competing mechanisms and evaluate which one mattered the most. Instead, scholars are bound to limit themselves to report that a theorized mechanism was present, and that it mattered for the outcome. This limitation also holds for my work. However, once we have a good understanding of the causal mechanisms at work affecting certain outcomes, future research will be better equipped to test the relative strength of competing explanations through large-n studies with more confidence as they may know the underlying causal mechanisms of their variables.

4.1.4 Process Tracing in Comparative Case Study Designs

My *main research aim* is to test and refine a causal mechanism. According to Gerring (2011), the choice between fewer and more cases comes with a number of trade-offs. Prominent among these trade-offs is the focus on causal mechanisms (small-n) versus causal effects (large-n). Given my main research aim, I engage in a comparative small-n study with two cases. While some authors argue that studying a causal mechanism works best in single case study (Beach & Pedersen, 2013), single case studies come at huge costs in terms of external validity. Moreover, while single case-studies have a place in *generating* novel hypotheses or causal mechanisms, they are not suitable to test them

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

(Gerring, 2011). In contrast, comparative case study designs allow me to enhance external validity (albeit only moderately), without comprising my ability to investigate both cases in depth. Given that process-tracing studies already enjoy very high degrees of internal validity, shifting the balance toward more external validity seems promising. A cross-case study design thus corresponds to my formulation of a generalizable causal mechanism which is theorized to work across cases of TSM engagement towards MDBs. This approach helps to avoid idiosyncratic, case-specific factors from the outset and to enhance theoretical parsimony. Most importantly, the most-similar case study design allows me *test* a causal mechanism across cases with very similar characteristics, yet different outcomes. This design thus enables me to refine the theoretical formulation of the causal mechanism andto specify the scope conditions under which it operates (Wunsch, 2016). Process tracing embedded in comparative designs thus contributes to the formulation of mid-range theories (George & Bennett, 2005; Bennett & Checkel, 2015). In sum, process-tracing is a resource-intensive enterprise and the inclusion of a second case multiplies the effort required. Yet, a most similar case study design fits my research aim best. In addition, it enhances the external validity of my findings and therefore promises considerable gains that outweigh the costs overall.

4.2 Case Selection

While *case selection* is a crucial part of any research design, it is especially important for qualitative studies where the limited number of cases does not allow for random sampling (Seawright & Gerring 2008). George and Bennet (2005) define a case as "an instance of a class of events," (p.5) and a case study as "the detailed examination of an aspect of a historical episode to develop or test historical explanations that may be generalizable to other events." (p.17). Following this definition of a case study, the researchers interest does not lie with an historical episode as a whole (e.g. an MDB reform process), but with specific aspects of that episode. I am interested in a specific causal mechanism involving mixed TSM tactics. Given this research interest, my universe of cases consists of all instances where transnational social movements tried mixed tactics to demand political and institutional change at MDBs. In line with my most-similar case study design and my interest to refine the scope conditions of my causal mechanism, I looked for cases with the same sequence of TSM tactics, largely similar scope conditions and yet different outcomes: TSM engagement for human rights accountability toward the World Bank from 1988 - 1994 (Case 1), and from 2011 – 2016 (Case 2). The selection of these two cases fulfils the basic requirements of a most similar case study design.

In most similar case-study designs, the researcher knows the causes and outcomes of both cases. Since we do not yet have a good understanding of the precise connection between a sequence of movement tactics involving different arenas of contention (MDB and nation states) and their effect on the socialization of MDBs, I chose two cases with similar causes and scope conditions, yet different outcomes. My comparison thus allows me to test the theorized causal mechanism of movement influence on two cases where the mechanism predicts political and institutional change. At the same time, I have

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

variation across both cases with regard to their dependent variables. This circumstance allows me to compare both cases in a most-similar cross-case comparison (George & Bennet, 2005, pp. 81-83). According to this case selection strategy, both cases should ideally be similar in all respects except for the outcome and one independent variable – the variable explaining the difference. In real life, only approximations of this similarity exist. In the discussion, I elaborate on how this limits the confidence in my findings. Still, my most similar cross case study is promising to refine the theorized causal mechanism and to specify the scope conditions for its application (George & Bennet, 2005, pp. 119-120).

To select my two cases, I initially mapped my universe of cases broadly by reviewing secondary material on TSM engagement for human rights accountability at international financial institutions (IFIs) which have in common that member states have differential influence according to their shares. To begin with, several studies exist on TSM engagement towards the IMF (O'Brien et al., 2000), among which only few specifically studied TSM engagement towards accountability (Scholte, 2008, 2011). Few studies systematically study TSM engagement towards human rights accountability at Regional Development Banks – a notable exception is that by Pallas and Uhlin (2014) on the safeguards policy reform at the Asian Development Bank. According to several authors, however, most instances of TSM engagement for human rights accountability has been towards the World Bank (Nelson, 1995, 1996; Tallberg et al. 2013; Weaver, 2008). Arguably, this comes as no surprise: first of all, the World Bank is an MDB that scores particularly high on social relevance due to its reach and impact. In contrast to regional MDB such as the Asian Development Bank or the African Development Bank, the World Bank possesses global membership. Its impact can partly be measured in financial terms, that is, the lending volume at its disposal, which exceeds that of other MDBs (see Ch.8)[2]. Moreover, the World Bank is widely regarded as a norm-setting organization, defining the standards of appropriate behavior for other MDBs throughout the last decades (Kapur, 2006; Kapur et al., 2011). As a consequence, whenever new general policies and institutional designs develop at the World Bank, there is a rather high likelihood that they will be copied by other MDBs in its organizational environment (Borges & Waisbich, 2014).

Other reasons relate to the World Bank's organizational culture - a comparatively open, non-hierarchical culture that invites outside influences but also encourages norm entrepreneurs from within (Vetterlein, 2015; Weaver, 2008). Partly as a reaction to social movement demands, the World Bank bureaucracy developed sophisticated mechanisms of interaction with civil society organizations already in the early 1980s. In 1982, it created the World Bank NGO Committee as well as an NGO Unit specifically charged with NGO relations (Nelson, 1995, p. 56). The continuous opening up of the World Bank also took place on the project level: While 21% of the projects involved civil society in 1990, a vast majority of 82% had components of civil society involvement in 2010 (Tallberg et al., 2013, p.5). This high degree of openness at the institutional and project level enabled comparatively high degrees of TSM engagement with the World Bank on different levels a time when other MDBs still offered no or very little access to civil society. This openness is also relevant theoretically, as several authors pointed to the fact that

2 Estimates are based on Annual Reports 2016 of World Bank, AfDB, IADB, EBRD, AIIB

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

it enabled the World Bank bureaucracy to continuously learn and improve its counter mobilization (Anderl, 2018; Rich, 2013).

Among the instances of TSM engagement for World Bank policy and/or institutional reform, three cases stand out involving particularly high levels of TSM mobilization and high degrees of counter mobilization. They have also been particularly far-reaching in terms of their impact. These three cases are: TSM engagement for the creation of a World Bank Inspection Panel in the early 1990s (Rich, 1994; Fox & Treakle, 2003), TSM engagement towards enhanced transparency standards for projects in the extractive industry sector (Colchester & Caruso, 2005; Anderl, 2018) and that concerning the recent Safeguards policy reform (2012 – 2016). While the second case concerned a specific branch and type of projects, the first and the third cases concerned the policy and institutional infrastructure of World Bank human rights accountability. According to Christopher Pallas, the mobilization for increased human rights accountability at the World Bank involving NGOs and activists from around the world in these two cases arguably represented "the first truly global civil society reform effort" (Pallas, 2013, p. 14). Crucially, though, the World Bank Inspection Panel case and the World Bank Safeguards Reform offer what is theoretically required for my most-similar case study: not only the *cause* of the mechanism, but also the sequence of tactics employed by the TSM were largely identical across these two cases. Moreover, several scope conditions appeared very similar: (a) the actor demanding change stayed in place with only some changes to its internal make-up and hence the movement's organizational as well as moral/epistemic resources, (b) the issue remained constant, (c) discursive opportunities were only slightly greater in the second place, and (d) access to decision-makers was good. Similarly, the degree of power asymmetries on the Board of Directors due to voting shares did not change substantially over the course of the years, with the United States and European member states holding the majority). Concerning *outcomes*, the political and institutional reforms toward increased human rights accountability marked a great success in the early 1990s. In fact, the establishment of an Inspection Panel that was able to accept citizen complaints and to review the World Bank's compliance with social and environmental safeguards was revolutionary at the time, as it provided citizens with a standing in front of MDBs for the first time. In contrast, the comprehensive reform of the social and environmental safeguards from 2011-2016 led to a weakening of the policies, the increasing limitation of the Inspection Panel's scope of authority. Also, the establishment of parallel institutions next to the Inspection Panel indicated that mixed social movement tactics did not unfold their power, resulting in social movement failure overall.

To what extent are my findings generalizable? I believe that they are to instances of TSM engagement towards MDBs, provided that TSM employ the same sequence of tactics under similar scope conditions. As briefly mentioned above, the World Bank largely shares its mission, scope of work, funding model and governance structure – and thus a set of relevant scope conditions of TSM-MDB interaction - with other multilateral development banks such as the Inter-America Bank for Development (IADB), the Asian Bank for Development (ADB), the European Bank for Reconstruction and Development (EBRD) and the African Development Bank (AfDB). Crucially, among all aforementioned institutions, voting power is dependent on deposits, leading to highly

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

asymmetrical powers of influence among member states. Moreover, all these MDBs share that liberal democracies are the largest member states: while the US holds the largest shares in the World Bank by far, Japan leads in the ADB and Nigeria in the AfDB. Outside the circle of MDBs with the aforementioned structural composition, my findings should be read with caution. Already the different mandate of the IMF (i.e. macro financial stability) or the fact that China is the most influential member state at the AIIB means that my findings cannot easily be transferred to these institutions. Beyond TSM engagement towards MDB, my findings only bear relevance if liberal democratic member states are in a position to dominate IO decision-making due to their *de facto* influence (and despite a formal focus on consensus). While most international organizations adopt a "one-country-one-vote principle", recent empirical research shows that power inequalities among member statesare currently increasing, rather than diminishing (cp. Viola et al., 2015), which suggests that my findings could be generalizable to an increasing number of IOs outside the realm of MDBs in the future.

4.3 Operationalization

Before evaluating the strength of the hypothesized causal mechanism to my case study, I develop indicators for all key features of the mechanism, starting with an operationalization of the socialization outcome (limited vs. comprehensive human rights accountability). I then move to the cause triggering the mechanism (joint TSM activism), and to an operationalization of parts I – III of the mechanism. Beneath each part, I operationalize the relevant scope conditions. Section 4.3 on operationalization concludes with an elaboration on my rules of aggregation, such as the way in which I add the values of single indicators of a tactic, scope condition, or outcome, into an aggregate value (low, medium or high).

4.3.1 Socialization Outcome: Human Rights Accountability

In the conceptualization of accountability above, I related limited and comprehensive accountability to the respective degree of legalization of the norm. Specifically, *limited accountability* is characterized by a low degree of obligation and precision of human rights and transparency policies, as well as a low degree of delegation to institutional oversight bodies. *Comprehensive accountability*, in contrast, is characterized by high values on each of these dimensions. To determine low versus high values on the dimensions, I operationalize the degree of *obligation* by looking at the bindingness of human rights and transparency policies. Either, they are highly binding (comprehensive) or rather enjoy the status of guiding principles or voluntary guidelines (limited). I operationalize the degree of *precision* of human rights and transparency policies by assessing whether they define specific circumstances under which they apply, and which specific actions result given that the circumstances of their application exist (comprehensive), or whether they are formulated in vague terms, leaving a lot of room for discretion (limited). Regarding the *scope* of relevant policies (scope 1), human rights and transparency policies may cover all activities by the MDB (comprehensive), or only apply to some as-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

pects of MDB activities, and contain major escape clauses, thus leaving large parts in the MDB's portfolio unregulated (limited). Then, I operationalize the degree of *delegation* to independent oversight bodies by looking at three critical factors. I examine whether oversight institutions exist in the first place, whether they are independent by virtue of a sufficient and independent budget as well as sufficient and highly qualified staff and the quality of their competencies, like those to investigate the relevant issues and to yield effective sanctions. Finally, I assess the *scope* of complaint mechanisms (scope 2) by observing whether the jurisdiction of oversight mechanisms is encompassing and/or whether access to the mechanism easy (comprehensive) or whether jurisdiction is restricted (e.g., only relating to part of the portfolio) and/or access difficult (limited). The following table summarizes the operationalization of my outcome:

Table 4: Operationalization of Limited vs. Comprehensive Accountability

	Comprehensive Accountability	Limited Accountability	Operationalization
Obligation	Binding Human Rights and Transparency policies	Nonbinding Human Rights and Transparency policies	Are policies formulated in binding documents and language vs. voluntary guidelines/ aspirational principles?
Precision	Detailed and specific Human Rights and Transparency policies	Vague Human Rights and Transparency policies	Policies define circumstances of their application and consequences vs. vague terminology
Scope 1 (Policies)	Encompassing coverage of substance Encompassing application (apply to all MDB activities and staff)	Restricted coverage of substance Restricted application (apply to some MDB activities or staff)	Human Rights policies entail civic, political, socio, economic and cultural rights. Transparency policies cover a public right to information at all stages of the policy cycles. Policies cover all activities + staff.
Delegation	Independent assessment of complaints Body authorized to take binding decisions	Third party is dependent on MDB Body makes nonbinding decisions/recommendations	Sufficient and independent *budget*? Sufficient and highly qualified *staff*? Authorized to take binding decisions vs. nonbinding recommendations?
Scope 2 (Complaints)	Encompassing Jurisdiction (third party investigations cover all MDB activity) Wide Access	Restricted Jurisdiction (third party investigations cover only fractions of MDB activity) Narrow Access	Oversight body has jurisdiction over all MDB activity. Can affected individuals file a complaint, OR only states (on behalf of those affected)? Are formal requirements for complaints high or low?

Source: own illustration.

4.3.2 The Cause: Joint Transnational Social Movement Activism

The cause triggering my causal mechanism is joint transnational social movement activity, involving community-based and internationally oriented constituencies. Such concerted TSM activity presupposes a minimum degree of coordination. Without such coordination, individual movements across different locations act on their own behalf.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

They may still fight for the same cause and use similar tactics. Yet, the sum of activities of individual constituencies alone would not be sufficient to justify the assumption of collective actor hood. Of course, movement subunits will never coordinate every single decision among one another. Similar to other collective actors, their subunits retain some degree of autonomy. To justify the transnational social movement operating as a collective actor, I expect to see coordination above a certain threshold. Given my research interest, I hold that communication about, and joint determination of the overriding strategic approach toward the target organization defines this threshold. In the empirical material, I thus expect to find traces of routinized communication between local and international movement actors.

4.3.3 Disruptive Movement Tactics towards the MDB (Part 1)

In part one of the mechanism, we should expect movements to employ disruptive tactics toward the MDB (Part 1a), which causes decision-makers of liberal member states to worry about the MDB's legitimacy and, following a logic of consequences, to open up for improved movement access (Part 1b).

Part 1a: Movements employ disruptive tactics toward the MDB
In line with my conceptualization of disruptive tactics, I operationalize the use of such tactics toward the MDB as all instances of movement activity outside established channels of engagement. The list of such activities includes but is not limited to *non-violent* forms of protest such as[3]:

- boycotts,
- demonstrations,
- breaches of obligations, agreements and/or the law,
- sit-ins,
- hunger strikes,
- scandalization (e.g. confrontational media campaigns),
- protest letters (a flood of letters or joint letters).

They may also include *violent* forms of protest, such as:

- riots,
- violence against the police or other representatives of the target institution,
- the destruction of property or symbols related to the target institution.

Part 1b: Decision-makers of liberal member states worry about the MDB's legitimacy and open up for improved movement access
I assess the worries of decision-makers among liberal member states by looking at traces that indicate such worries in press statements, protocols of board meetings, parliamentary debates and interviews, including my own. Such worries can thus be

3 This list builds on but also expands existing classifications (Gertheiss & Herr, 2017, p. 12).

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

expressed directly (e.g., if a decision-maker says, "I am worried") or indirectly (e.g., they demand a fact finding mission and an action plan from the MDB on the highest level). The opening up of decision-makers to the movement is then assessed in terms of increasing access, which is also a scope conditions for movement influence for part II (see below).

Scope Conditions (Part 1)

The relevant scope conditions theorized for such successful disruption are the availability of organizational resources, an issue salient enough to justify disruptive tactics in the eyes of decision-makers and the wider public, as well as support from the organizational environment. While the degree of counter-mobilization also matter at this stage (in fact it is the one scope condition that matters throughout the whole process), it is most important in part 3 of the mechanism and will hence be discussed in that context (see below).

I operationalize ***organizational movement resources*** by looking at the connections movement protagonists have to supporting networks and, particularly, to large social movement organizations (SMOs).

Operationalizing the ***nature of the issue*** requires a look at several dimensions. I assess the degree of issue specificity in a qualitative manner. An issue counts as "specific," if it refers to a particular MDB policy and/or a specific MDB project. In contrast, the issue counts as "general" if it comprises features of the MDB (or MDB environment) as a whole. In addition, a causal chain between MDB behavior and movement demands is "short" if movement demands addressing particular MDB behavior and establishing a direct connection between that behavior and the suffering incurred. The vulnerability of an MDB to an issue counts as "high," if the issue addressed by the movement directly relates to the MDB's mandate (i.e., development). The respective "sovereignty costs" of reform demands to member states will be assessed by analyzing whether the issue involves aspects that are traditionally located at the core of Westphalian sovereignty. According to Cogan et al. (2016), territorial matters as well as those involving the relation between the state and its citizens (e.g., security, immigration, citizenship and human rights) score highest.[4] Finally, the issue possesses a particularly "high resonance potential" if it involves bodily harm to vulnerable individuals and/or legal equality of opportunity.

I operationalize ***support from the organizational environment*** as referring to those organizations with which the MDB works closely and/or with which it shares a common identity and purpose. As to the former, I look at the main cooperating partners of the MDB in question as indicated in its Annual Reports and evaluations. As to the latter, I assess an organization's identity in virtue of its mandate. Where in doubt, I also consult

4 In a different conceptualization, Tallberg et al. (2013) hypothesize that sovereignty costs differ according to stages of the policy cycle. Specifically, they are highest on issues that involve decision-making, i.e., processes where states agree to internationally valid policies that limit their autonomy (Tallberg et al., 2013). However, this focus on stages of the policy cycle is different from the sovereignty costs associated with the issue.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

the lists of participants of joint workshops and conferences at the MDB's annual meetings with a common goal. The closer an organization is to the MDB in question, the more likely it is that the MDB will be compared to this organization. I assess the degree of *support for movement demands* from these organizations by looking at their "talk" (i.e., in organizational documents and press statements) as well as their "actions" in terms of policies and institutional reform.

4.3.4 Conventional Tactics toward Member States (Part 2)

According to the second part of the causal mechanism, the movement engages in conventional tactics toward decision-makers of liberal member states (Part 2a). Following a logic of appropriateness, these decision-makers are persuaded by movement arguments and accept their demands (Part 2b) to become active vis-à-vis the MDB. The relevant scope conditions for this part of the socialization process to work are MDB *crisis*, *access* and *moral/epistemic authority*.

Part 2a: Conventional tactics toward decision-makers of liberal member states
When using conventional channels of engagement toward liberal member states, movement representatives either actively invite decision-makers (e.g., to conferences, workshops or debates), or they accept invitations (e.g., to participate in parliamentary hearings or expert briefings). In both cases, movement representatives use conventional channels to persuade decision-makers of their aims. More specifically, they communicate their frames. These frames typically consist of the following four elements: (a) a problem definition, (b) the causal attribution, (c) the evaluative nature of a statement, and (d) the action that is required to solve the problem (Entman, 1993). To observe each of these components, I engage a qualitative content analysis in MAXQDA to analyze movement statements made in conventional (inside) channels toward decision-makers of member states in view of these four dimensions. Instead of looking at all movement statements, I am only interested in those involving the demand for an accountability reform at the World Bank.

Part 2b: Movements persuade decision-makers of liberal member states
To increase the plausibility in my argument that movements were successful in persuading decision-makers, I first of all draw on the technique of *cognitive mapping*. Wolff (2009) introduced cognitive maps to political science research as a technique for actor-centric process-tracing. While cognitive maps—an instrument from the realm of psychology—have been part of political science research (Axelrod, 1976; Shapiro, Bonham, & Heradstveit, 1988), Wolff (2008) showed the potential of cognitive maps in rendering causal mechanisms plausible. In short, cognitive maps depict specific interpretations of the social world by linking central concepts and causal assumptions held by a given actor. Instead of providing a comprehensive worldview of an actor, cognitive maps are specific in that they are focus on the perception of specific problems. Their character as maps allows visualizing a simplified version of core assumptions, values, and relationships. Drawing on such cognitive maps helps to verify the plausibility that a movement frame relates to already held beliefs by decision-makers, thus increasing the cogni-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

tive ease and experiential commensurability (for an elaboration of both terms, see) of these proposals. Where we are confronted with ideas or concepts that we already know or believe in, we experience cognitive ease and find it easier to believe the idea. The material laying ground for the cognitive maps of critical decision-makers consists of biographies, newspaper and journal articles written by close observers, as well as interviews, press statements and speeches. Secondly, I hold that the nature of the response by decision-makers tells us something. Are challenges made by TSM welcomed, treated as reasonable but misguided, or rejected and treated as beyond the pale? For instance, do decision-makers suddenly employ the same language and metaphors of the TSM in describing the problem, attributing causality, in their evaluations and their actions?

Scope Conditions (Part 2)

As elaborated upon in the previous section, I theorize that disruptive movement tactics produce a crisis at the MDB in virtue of the fact that important MDB member states begin to worry about the MDB's legitimacy. This crisis, which is the product of disruptive tactics, is an important scope condition for the effectiveness of conventional tactics, as it leads to enhanced *access* to decision-makers and produces the necessary degree of uncertainty—two conditions that are integral to the success of conventional tactics. To use the access effectively, the degree of epistemic/moral authority of the movement matters as a scope condition for movement success in Part 2 of the causal mechanism.

According to a classical definition provided by Seeger, Sellnow, and Ullner(1998), organizational ***crisis*** refers to "specific, unexpected, and non-routine events or series of events that [create] high levels of uncertainty and threat or perceived threat to an organization's high priority goals" (p. 231). Crisis at MDBs then typically threatens to undermine the MDB's legitimacy or funding, or both, as both. Crisis caused by TSM activity becomes evident if the MDB adopts measures that are directly connected to movement activity, but unusual in the sense that they exit the path of everyday routine. For instance, if the MDB switches from ignorance and avoidance of movement demands to fiercely rejecting them, while simultaneously engaging in internal investigations, these would be indications of crisis at the MDB. Alternatively, if MDB management gets caught by member state representatives and/or media at lying or covering-up misconduct, this would constitute crisis.

I operationalize ***access*** as the degree and quality of direct interaction between movement representatives and decision-makers. Indicators are the quantity and quality (time, setting, purpose, involvement of public) of direct meetings. To collect data on access, I take two routes. : The first is an institutional or *de jure* route, asking, "What are the formal channels of access guaranteed by the constitution of the actor in question?" This measure includes official channels of access to those sections of a member state that are responsible for policy-making in MDBs in virtue of their respective constitutions. I thus look at the political system of MDB member states and additional formal channels. The second criterion seeks to capture access (e.g., private encounters) through informal channels. For instance, I seek to include access that is based on personal relationships between activists and decision-makers. To collect

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

data on this second channel, I asked movement and government representatives in interviews about important contacts and/or encounters.

I assess the degree of ***moral or epistemic authority*** among movement constituents by looking at the availability and social status of movement members. Concretely, the share of academics among movement constituencies, as well as the presence of organizations with scientific credentials (e.g., professors and PhD holders among their ranks and people who publish in academic journals) will serve as a measure of epistemic authority. Meanwhile, the availability of public intellectuals and organizations with high moral reputation among the general public (i.e., in mainstream media) serves to indicate moral authority.

4.3.5 Member State Incentives Toward the MDB (Part 3)

In Part 3 of the causal mechanism, decision-makers of liberal MDB member states use material incentives to induce reform at the MDB (Part 3a). As brokers between member state pressures (the primary locus of perceived responsibility) and MDB well-being, EDs) on the Board of Directors will push the whole Board to adopt those reforms necessary to mitigate critique and to avoid funding cuts (following a logic of consequences[Part 3b]).

Part 3a: MDB member states use material incentives to induce reform

Material inducements or even coercion of the MDB on behalf of a member state becomes evident in press statements or legislation enacted by decision-makers of member states. In principle, there could be positive incentives in the form of increased financial contributions attached to the adoption of a reform. Yet, Ipredict material inducements will generally come in the form of threats to withdraw parts of the MDB's funding. Specifically, I will look at publicly voiced threats to withdraw funding, either in the form of legislation, through media, or, in the course of World Bank —member state discussions concerning the next round of funding. I triangulate this information with my interview material, as most World Bank member states' discussions in relation to future funding take place behind closed doors. Inducements turn into coercion where the MDB has no choice but to comply (Anderson, 2008; see elaboration in analytical framework section). Whether this threshold is met depends on the shares a member state holds and threatens to withdraw. Clear indicators for that threshold do not exist, which is why I rely on a careful assessment against the background of general circumstances to make this judgment call (i.e., that coercion rather than inducement took place).

Part 3b: The MDB Board of Directors collective worries about a funding cut and adopts political and institutional reforms

I assess worries of the MDB Board of Directors by observing their reactions to incentives from member states closely. Sudden changes in rhetoric, excuses, bold promises for reform and other unusual actions in direct response to the threat of incentives or sanctions of member states are indicative here. Ultimately, however, the adaption of a reform the Board of Directors did not want in the beginning and in light of clear in-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

dication that it is not convinced of the reform (e.g., leaks to the press), is the clearest indication for the Board of Directors 's fear. Assessing the adoption of reform itself is a straightforward task, as new policies and institutional reforms are publicly available in the respective MDB policy documents and resolutions.

Scope Conditions (Part 3)

The necessary scope conditions for an effective use of material incentives toward the MDB are straightforward. Such incentives only work if the MDB is highly dependent on those member states using the incentive, for example, when there are power asymmetries on the Board of Directors . Closely related to this scope condition, the degree of counter-mobilization among those opposing the direction of change should not exceed a certain threshold. In short, if a move of a powerful coalition of member states provokes an effective balancing of power, the effect might not be produced. Balancing of power may also occur through the MDB itself by engaging in strategies to derail or co-opt the push for reform.

Power asymmetries on the Board of Directors , the highest decision-making organ of each MDB, exist by virtue of differential shares that member states hold. As these financial shares—what each member state contributes to the MDB's overall budget—translate into voting rights, more shares means more votes. The higher the discrepancy in terms of shares among member states, the higher the degree of power asymmetry. Since the degree of power asymmetries in virtue of shares is relative, no precise measure exists determining the threshold separating a low from a high degree of power asymmetry. I argue that power asymmetry is high if one member state has the ability to determine the fate and circumstances of the MDB as an institution. For instance, if the funding cut of one member state risks the survival of the MDB as it stands, or if one country possesses so many shares that they can exercise a veto with regard to any Board decision, the power asymmetry between that country and those who lack such ability is high.

Scope Condition (perpetual): The degree of counter-mobilization

I argued in Chapter 3.3 that counter mobilization by the MBD bureaucracy and MDB member states formed a perpetual scope condition representing the dynamic interaction between the transnational social movement and movement addressees. Accordingly, counter mobilization is present along all three parts of the mechanism. Counter-mobilization on behalf of the *MDB bureaucracy* involves three dimensions: avoidance, defiance and manipulation (see chapter 3.3). Since counter mobilization assumes intentionality, I will only talk of *avoidance* as an act of counter mobilization if it is a deliberate reaction to TSM demands. To observe avoidance, I thus need to proof that a) TSM demands were perceived and known by the MDB in virtue of apparent indications (e.g. extensive media coverage or letters to the president), and that b) the MDB chose not to react to these demands as if they did not exist. Next, *defiance* involves explicit and subtle forms of counter mobilization. Instances of the former are open rejections of the demands (e.g. in press statements or during public events). Subtle defiance are attempts at cooptation, either by opening-up to TSM consultation without translating TSM de-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

mands into tangible decision making outcomes (Cooptation 1), or through the inclusion of only moderate movement constituencies, which has the important side effect to divide the movement into moderate and "radical" constituencies (Cooptation 2). Evidence of defiance-cooptation 1 consists in a) the existence of MDB-TSM consultations in combination with b) the lack of integration of TSM demands into policy or institutional reforms. While it is admittedly difficult to collect data on this scope condition, I relate the extent of TSM engagement by the MDB to the degree of MDB responsiveness to TSM demands over time (i.e., for the period of four years in each of my cases). I supplement this measure with assessments by observers and those actively involved in the process of TSM-MDB exchange in interviews. Evidence for defiance-cooptation 2 is incentives to moderate movement parts and/or sanctions to more "radical" parts (Deitelhoff & Daase, 2017; O'Brien et al. 2000). Incentives in turn are invitations or funding to attend consultation rounds, the legitimation of movement parts through joint conferences, pictures or even the employment of selected movement representatives (Anderl, 2018). Manipulation, finally, are attempts to influence the MDB's environment, specifically external demands, to mitigate pressures (Oliver, 1991; see chapter 3.3). On a structural level, MDB bureaucracies may use their knowledge creation powers and seek to influence the very rules and norms that regulate MDB – TSM interaction. Moreover, the MDB bureaucracy may either align with external allies that support its goals (including governments of MDB member states) to mitigate TSM critique.

Counter-mobilization by the *MDB member states* is arguably easier to capture and operationalize, as it is typically expressed more openly on the respective MDB executive boards, using either the "power of the purse" or "voice and vote" (Park, 2017). To operationalize member state counter mobilization expressed in terms of the "power of the purse", I look for instances where member states threaten to withdraw (parts of) their funding. To operationalize counter mobilization in terms of "voice and vote", I look for traces of member state dissent in minutes of board meetings, public statements, or input they provide to the MDB in the course of formal MDB-TSM consultations. Such open declarations of member state opposition will be subject to a qualitative content analysis in MAXQDA. To get at the more hidden opposition voiced by member states, I asked my interviewees about the opposition they experienced during negotiations. A drastic form of MS counter mobilization is the creation of novel institutions, or the threat thereof. Such opting-out has been described by Morse and Keohane (2014) as counter multilateralism.

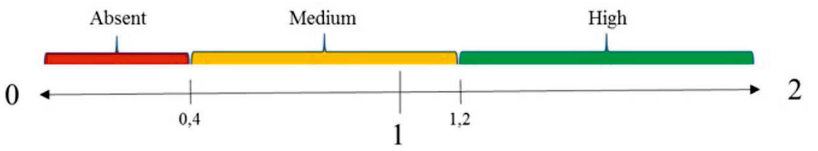
4.3.6 Rules of Aggregation

To assess the final outcome in both cases, I assign an aggregate value between 0.0 and 2.0 to the resulting human rights accountability. Then, values from 0.0 to 0.4 count as an *absence* of accountability, 0.5 to 1.2 as *limited* and 1.3 to 2.0 as *comprehensive* human rights accountability.[5]

5 Naturally, the cut-off points are somewhat arbitrary, as the difference between 1.2 and 1.3 is by no means bigger than that between 1.1 and 1.2. If a final value is close to a cut-off point, I will make that transparent in the discussion of my outcomes.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

Graph 4: Cutting points (rounding up or down)



Source: own illustration.

To get at this aggregate value, I equally assign a value between 0.0 and 2.0 to each dimension of my outcome. If several dimensions exist and the assigned values on these dimensions diverge, it is also possible that the aggregate value lies in between these numbers (e.g., adopting the aggregate value of 1.5 if the assigned values on two dimensions were 1 and 2). More demanding are calculations where one concept is operationalized across more than two dimensions (e.g., with the assigned values 2, 0, 1, 2, and 2). Again, I calculate the average value, adding all values and dividing the outcome by the number of dimensions (in the example: 2 + 0 + 1 + 2 + 2 = 7; then: 7 / 5 = 1.4), whereby all aspects of a single dimension count equal. All values obtained in each dimension are then rounded up where there is a 5 or higher behind the decimal point, while I rounded the value down where the number behind the decimal point was ≤ 4. Each dimension is assigned one value, based on the questions and indicators listed in each table.

4.4 Data Collection and Analysis

Naturally, a great deal of the interaction between TSMs and critical actors for IO decision-making (at the World Bank or member states) occur behind closed doors, particularly when the social movement engages in conventional tactics. Similarly, interactions between member states and the World Bank often remain obscure. For example, there is a tacit understanding at the World Bank that initiatives opposed by powerful stakeholders typically do not make it on the agenda of the Board of directors. To avoid a negative vote, U.S. deputies, EDs and World Bank management negotiate a consensus behind the scenes (Babb, 2009).

4.4.1 Data Collection

I was fortunate to do several field research trips from 2015 to 2017. Among them were participations in TSM strategy meetings (notably a one week strategy meeting in Frankfurt in May 2016) and World Bank –TSM consultations on human rights policy reform in Berlin (either organized by the Berlin representation of the World Bank , or by the German Institute for Human Rights). Most importantly, though, were two field research trips to Washington, D.C. The first took place in June and July of 2015, the second one in March andApril of 2017 (as a visiting scholar at American University's Washington Col-

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

lege of Law). The two research stays in Washington D.C. allowed me to interview critical people for both cases more than once, sometimes even three to four times. On the side of the World Bank , I interviewed staff working for the offices of several EDs on the Board of Directors, senior management in charge of the accountability reform process, the head of the World Bank 's legal department responsible for operational policies, several former and current members of the World Bank Inspection Panel as well as senior members of the Independent Evaluation Group (IEG;see Appendix for a full list with the time, location, and positions of those interviewed). My second research stay also coincided with the World Bank's Annual Spring Meeting 2017. In the context of this Annual Spring Meeting, the Civil Society Policy Forum provided an excellent opportunity to meet leading TSM representatives, executive directors and World Bank staff working on safeguards policies, as it also provides a unique opportunity to observe TSM – Board of Directors interactions during CSO Roundtables first hand. Moreover, Annual Spring Meetings provide a unique opportunity to take a look behind the scenes. For instance, several networking events hosted by NGOs or subunits of the World Bank (e.g. a reception organized by the Inspection Panel) allow for extensive participant observation, but also for a range of background conversations. Similarly (why conceal it?), house parties that take place during Annual Spring Meetings allow good opportunities to have background conversations with TSM, member state and World Bank representatives. Finally, in Washington D.C., I was allowed to take part in "Tuesday Group" meetings. The Tuesday Group is a regular networking format that brings Washington, D.C.-based NGOs together with U.S. decision-makers—primarily from the Treasury, but also from State Department and Congress (see Chapter 6 for an elaboration). The flip side of such informal background conversations is that the acquired data cannot be cited, and these conversations do not appear in the Appendix).

A number of challenges are common to using interviews as a method to reconstruct causal mechanisms. I here restrict myself to discussing the most pressing ones, before outlining potential remedies. First, access is not complete. While I gained access to key people within the social movement and the World Bank, getting access to key parliamentarians in the most important member states (most notably, U.S. Congress) was difficult. While not all potential interviewees agree to do an interview, those who do may have an agenda on their own. In other words, why do interviewees agree to do the interview in the first place? As George and Bennet (2005) put it, "The analyst should always consider who is speaking to whom, for what purpose, and under what circumstances" (p. 99). To avoid systematic biases, researchers should thus be aware of the agenda interviewees might have. To assess whether the information gained through interviews accurately reflects reality, one needs to consider whether the selection of sources is systematically biased in one direction. In my work, I talked to slightly more activists than World Bank staff, and to considerably more activists and World Bank staff than decision-makers of member states. This is one of the reasons for rating evidence gained from relevant documents higher than my interview material overall (see below). In addition, different people tend to have different perceptions regarding the course of an event. For instance, while I interviewed several individuals who occupy key roles in the process I am interested in, their perspective on the same event may differ considerably (i.e., interviewees tend to systematically overstate their importance in bringing

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

about an outcome). To deal with these diverging perceptions which all represent a piece of the mosaic, I tried to get as many perspectives from people with different affiliations as possible. While the motives to misrepresent a given event, and the fact that perceptions diverge(even where people have the best intentions to tell "the truth", the use of interview material as evidence for an objective reality is particularly challenging with regard to historical events. The more time elapses between the event and the interview, the more researchers need to calculate the imperfections of human memory, leading people to recall an event selectively. In some instances, interviewees change their interpretations, and thus even their own memory of an event over time through the exchange with other people. Due to these potential biases in interview accounts, the analysis of primary documents is at the core of my research. In particular, I draw on four types of documents:

- official position papers of the transnational social movement (typically in the form of "Joint Letters" or "Position Statements" that are submitted to World Bank management or the President),
- press Statements by the World Bank ,
- minutes of parliamentary hearings (particularly U.S. Congressional subcommittee hearings, but also hearings from other member states), and
- newspaper articles from major newspapers in powerful member states (e.g., *Financial Times*, the *Washington Post*, the *New York Times* , and *The Guardian)*

Information of several scope conditions required me to look at movement documents. For instance, collecting data on issue characteristics requires looking at the core demands the transnational social movement pushed for in their activism toward the MDB in joint movement letters as well as strategy papers. In addition, I consulted two further types of documents to collect data on the outcome of TSM activity:

- operational social and environmental policies,
- board resolutions laying ground for institutional reform (i.e., the resolution establishing the World Bank Inspection Panel)

Access channels to these six types of documents vary. I accessed joint letters of TSM engagement via the websites of those TSM constituencies coordinating the campaign. Press statements of the World Bank are available on the World Bank 's website. I accessed newspaper articles via the LexisNexis search engine. Finally, I accessed minutes of parliamentary hearings via parliamentary archives[6] or research services, such as the Congressional Research Service. Of course, these documents are not neutral. Notably, newspaper articles inevitably represent an extract of reality. Yet, the documents are rather close to my objects of interest in a given part of the causal mechanism. For instance, TSM joint statements are traces of the TSM frames and demands, while World Bank safeguard policies and the Board Resolution establishing the Inspection Panel are

6 For the U.S. Congressional archive, I combined an online search with several visits to the U.S. Library of Congress in Washington, DC.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

a fairly direct accounts of the degree of legalization of World Bank human rights accountability (for a more detailed analysis, see chapter 5).

Summing up, my data corpus consists of a vast array of documents (specifically legal documents such as board resolutions and policies, press statements, minutes of parliamentary hearings, and the memoires of decision-makers), newspaper articles, notes from participant observations and— last but not least—interviews. As a general rule, I preferred documents over interviews where these documents were available and a close representation of the phenomenon of interest. While interviews are an imperfect source to gain access to real-world phenomena (i.e., outside the perceptions of those interviewed which are real-world phenomena as well), they may provide the only available source filling blank spaces and are therefore unavoidable. To deal with the problems of biases in the data collected, the best solution is the combination of three strategies in the analysis of my results (see empirical Chapters 6 - 8):

- a critical reflection on the scope and direction of the bias,
- a triangulation of different sources of the same type (i.e., interviews with several stakeholders in the process), and
- a triangulation of observations across different types of sources (i.e., interviews, newspaper articles and minutes of congressional hearings).

4.4.2 Data Analysis

I engaged in a qualitative content analysis in MAXQDA to deduct the most important information from my interview transcripts, legal and organizational documents, hearings and newspaper articles (Mayring, 2010; Schreier, 2012). Specifically, I derived the codes for my analysis deductively from my research question and the operationalization of my concepts. For example, I developed a code with the label "issue" that involved all indicators specified in the operationalization (e.g., nature and content, specificity, salience to MDB) as subcodes. Then, I scrutinized the empirical material and structured it according to the codes of interest to me (Schreier, 2012). To evaluate my evidence, I used tests common to process-tracing studies (Collier, 2011). In the section "Operationalizations" (4.3) I specified the empirical manifestations (or indicators) of the components of my causal mechanism. Yet, developing indicators is different from analyzing the information in light of these indicators. In process tracing, as in other scientific works, a common differentiation exists between "observations" and "evidence." Whereas the former refers to raw data before it has been scrutinized with regard to its content and accuracy, "evidence" is a term reserved for data after it has been evaluated. Evidence thus enjoys a superior epistemic status compared to observations, as only the former entails inferential value with regard to the phenomenon under investigation. At a minimum, the evaluation of observational data involves a reflection concerning the accuracy of the observations (i.e., by using contextual knowledge to interpret whether a piece of information could be expected, or not) as well as a reflection on potential sources of error. Using Bayesian logic of inference helps to update the confidence in a new piece of information. "Bayes' theorem" is a term borrowed from statistical analysis that describes the probability of an event based on prior knowledge of contextual factors related to that

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

event. The theorem helps to update subjective beliefs in a given statement in a rational way, considering new pieces of evidence in relation to already existing evidence. In process tracing, the Bayesian theorem helps to evaluate information concerning evidence for each part of the causal mechanism. Specifically, it provides guidance to update our belief that a particular part of the mechanism took place in light of the evidence. To do the updating, we need to know our prior confidence in the hypothesized part of the mechanism and which evidence we expect to find if the mechanism was present. The prior confidence in the causal mechanism I hypothesized stems from the logical plausibility of the mechanism (logic) and existing research. A reflection with regard to our expectations to find particular pieces of evidence helps us to determine the degree of *certainty*, as well as the *uniqueness* of the evidence at hand. Finding evidence means finding empirical fingerprints of the causal mechanism. With regard to some empirical fingerprints, I was certain to find them before looking at the data. For instance, I expected to find joint letters addressed to the World Bank as indications of coordinated TSM advocacy. While the presence of such evidence only confirms prior beliefs that were already strong, the lack of such evidence has strong disconfirming power: if I do not find it, there is a high likelihood that the theorized mechanism is not complete. Uniqueness, on the other hand, refers to empirical predictions that other theories do not make. If my theorized causal mechanism predicts to find a rare piece of evidence, and we actually find it, this has great confirmatory potential regarding the theory. We can determine the degrees of certainty and uniqueness of the evidence with the help of our contextual knowledge. Against this background, the following four "tests" then allow to scrutinize and evaluate the evidence (Collier, 2011):

- **Straw-in-the-wind test** (low uniqueness, low certainty). This is the weakest of the four tests, neither necessary nor sufficient to confirm a hypothesis.
- **Hoop test** (high certainty: necessary to confirm hypothesis). If the hypothesis fails the hoop test, this disconfirms the hypothesized mechanism.
- **Smoking gun test** (high uniqueness: sufficient to confirm hypothesis). If the causal mechanism does not leave traces of a smoking gun, this does not decrease our confidence in the causal mechanism due to the high uniqueness.
- **Doubly decisive test** (high certainty, high uniqueness). This is the most demanding test, both necessary and sufficient to confirm a hypothesis.

In practice, I then used these tests to evaluate the degree of confidence I could have in a given causal link. To give an example, there were competing accounts regarding the proposal for an inspection panel at the World Bank (see Chapter 6). While World Bank staff claimed that they had worked on plans for an inspection panel, movement representatives firmly asserted authorship for the idea. To test these competing "hypothesis" (which mattered a lot for the approval/disapproval of my causal mechanism), I evaluated the evidence for both based on the four tests.

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -

https://doi.org/10.14361/9783839451830-008 - am 14.02.2026, 07:38:44. https://www.inlibra.com/de/agb - Open Access -