Ideational Change and the Emergence of the International Norm of Truth and Reconciliation Commissions
Michal Ben-Josef Hirsch

European Journal of International Relations published online 12 July 2013
DOI: 10.1177/1354066113484344

The online version of this article can be found at:
http://ejt.sagepub.com/content/early/2013/08/22/1354066113484344

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Standing Group on International Relations of the ECPR

Additional services and information for European Journal of International Relations can be found at:

Email Alerts: http://ejt.sagepub.com/cgi/alerts

Subscriptions: http://ejt.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> OnlineFirst Version of Record - Aug 23, 2013

OnlineFirst Version of Record - Jul 12, 2013

What is This?
Ideational change and the emergence of the international norm of truth and reconciliation commissions

Michal Ben-Josef Hirsch
Belfer Center for Science and International Affairs, Harvard Kennedy School, USA

Abstract
What role does ideational change play in norm emergence? While there has been some attention to changes in the application of norms, most scholars refer to the ideas that are associated with a norm’s practice as being fixed. I argue that ideational change is a causal mechanism that facilitates norm emergence. In particular, I propose three types of content change that capture changes in the ideas associated with the goals expected to be attained by the application of the norm (‘logic of consequences’), with its morality (‘logic of appropriateness’), and with its relations with similar or alternative practices (‘specification’). These changes in the rational and moral reasoning and argumentation that frame the practice that is associated with an emerging norm are likely to make this practice congruent with more contexts and appealing for more states. To illustrate the content change proposition, this article traces the emergence of the international norm of truth and reconciliation commissions. In the debates that followed the South African Truth and Reconciliation Commission, truth and reconciliation commissions shifted from being seen as a political compromise to being regarded as a ‘holistic’ tool for social and political reconstruction and came to be associated with multiple democratizing effects. Truth and reconciliation commissions also shifted from being the ‘weaker alternative’ to trials to a practice that is morally equal and complementary to the judicial option. Taken as a whole, these changes in the expected utility, morality, and specification of truth and reconciliation commissions facilitated their emergence and consequent institutionalization as an international norm.

Corresponding author:
Email: michal_BJH@hks.harvard.edu
Keywords
Appropriateness, international norms, norm emergence, transitional justice, truth and reconciliation commission

Introduction

Why, in the past decade, have truth and reconciliation commissions (TRCs hereafter) come to be regarded as the ‘staple’ of transitional justice mechanisms and an ‘imperative’ for societies emerging from repressive regimes or intra-state conflict (Bloomfield et al., 2003: 3; Hayner, 2001: 250; Roht-Arriaza, 2006: 4)? While the practice of truth commissions began nearly four decades ago, it was only in the early 2000s that the number of initiated TRCs grew more significantly.1 Early truth commissions (from the 1980s through the mid-1990s) were largely variants of a regionally diffuse and temporally specific model associated with the political transitions in Latin America. By most accounts, these early truth commissions were deemed suitable only for a specific political setting (i.e. negotiated transition) and for specific types of human rights violations (i.e. torture and disappearances). In fact, these early truth commissions were described as ‘a poor alternative to justice’ (Mendez, 1997: 267) and legal scholars were quite concerned that this model might spread to other settings (McAdams, 1997; Maclean, 2006).

Currently, even though there is enduring criticism regarding their underlying assumptions (Daly, 2008; Mendeloff, 2004; Snyder and Vinjamuri, 2004) and scant evidence-based assessment of their successful impact (Ben-Josef Hirsch et al., 2012; Olsen et al., 2010; Wiebelhaus-Brahm, 2010), TRCs not only are spreading worldwide, but also have been institutionalized at the international level. They have been endorsed, recommended, and supported by the United Nations (OHCHR, 2006) as well as by leading non-governmental organizations, including Amnesty international (Amnesty International, 2007), Human Rights Watch (Dicker and Keppler, 2004), and the International Center for Transitional Justice (ICTJ, 2012). As a consequence of this institutionalization, TRCs appear to have become a permanent feature of transitional and post-conflict practices.

The international institutionalization of TRCs has not gone unnoticed and there has been a growing body of literature that describes and explains the spread of TRCs. Often, TRCs have been associated with the so-called ‘justice cascade’ and with the principle of universal legalism that emerged following World War II and the Nuremberg trials and that accelerated following the political transitions in Latin America in the late 1980s and early 1990s (Lutz and Sikkink, 2001). Other explanatory factors have included the demands of domestic groups for accountability and justice (Elster, 1998; Huntington, 1991: 221–230; O’Shaughnessy and Dobson, 1999; Sieff and Vinjamuri, 1999), the demonstration effect of the South African TRC (Gready, 2011), and the advocacy of transitional justice practitioners and experts (Ben-Josef Hirsch, 2007). The most common causal mechanism that is said to account for the international spread of TRCs is human agency, that is, either domestic groups and state leaders, or international advocacy networks.

While these political actors are undoubtedly an important part of the story, in order to understand the shift from what was largely a regional institution to an emerging international norm, we must focus our attention on the ideas, or rationalizations, that made TRCs increasingly appealing to both domestic and international political actors. I argue
that the debates that followed the South African TRC changed the ideas about why, when, and for whom a TRC would be desirable. TRCs shifted from being seen as a political compromise to being regarded as a ‘holistic’ tool for social and political reconstruction and associated with multiple democratizing effects. TRCs also shifted from being the ‘weaker alternative’ to trials to a practice that is morally equal and complementary to the judicial option. Taken as a whole, I argue that these ideational changes facilitated the emergence of an international norm of TRCs.

The literature on norm emergence has generated a number of important insights about the actors and mechanisms that account for the emergence and selection of international norms. While there has been some attention to changes in the application of norms, most scholars refer to the ideas that are associated with a norm’s practice as being fixed. Especially during a norm’s emergence, scholars’ focus on norm entrepreneurs assumes, either explicitly or implicitly, that the goals and morality that are associated with a given norm are presented by these norm entrepreneurs early on in the process of norm emergence and remain largely unchanged. This static conceptualization is empirically unwarranted as it does not provide an accurate description of the emergence of international norms. Moreover, this static conceptualization is particularly problematic when it comes to debates over the effects of norms. Both advocates, who argue that ‘norms matter,’ and their critics tend to aggregate cases from different stages of the norm life-cycle as a way of demonstrating their effects or lack thereof. Results are likely to be skewed if different logics for the norm’s application have operated in these different stages.

This article challenges this static conceptualization by proposing that change in the content of a norm — that is, in the ideas that constitute the norm — is a causal mechanism that facilitates norm emergence. In particular, I propose three types of content change that capture changes in the ideas associated with the goals expected to be attained by the application of the norm (‘logic of consequences’), with its morality (‘logic of appropriateness’), and with its relations with similar or alternative practices (‘specification’). These changes in the rational and moral reasoning and argumentation that frame the practice that is associated with an emerging norm are likely to make this practice congruent with more contexts and appealing for more states.

This article proceeds as follows. In the next section, I argue that ideational changes have been largely overlooked as a causal mechanism that facilitates norm emergence. I then introduce my content change proposition and consider the main alternative challenge to it. In the following section, I demonstrate my argument through an analysis of the Truth vs Justice debate that took place during and following the South African TRC (1995). The Truth vs Justice debate revolved around the question of whether the South African TRC model was generalizable and universally applicable. In the conclusion, I focus on the empirical implications of this analysis for how we study and evaluate the effects of TRCs and on the empirical and theoretical implications of this argument for the broader study of international norms.

**Change and the emergence of international norms**

When it comes to conceptualizing change, the international norms literature generally identifies two types of change: first, ‘change in the broadly accepted standards of
international behavior,’ that is, the process of norm emergence and the dynamic of change from no norm to norm or from one norm to another (e.g. from practicing slavery to condemning it) (Adler, 2004); and, second, change in a norm’s effects and/or effectiveness, that is, in the ways and/or degree to which a norm changes other features of the political landscape, such as states’ behavior and/or the strategies political actors use within and between states (Berman, 2001: 234; Dimitrakopoulos, 2005; Kacowicz, 2005). Beyond these two rather narrow understandings of change, attention to changes in the goals and morality associated with a norm and its application has been limited (Barkin and Cronin, 1994; Goertz, 2003: 49–51; Sandholtz, 2008).

In line with Finnemore and Sikkink’s (1998) widely cited Norm-Life-Cycle model, most of the studies on international norms accept, either implicitly or explicitly, the notion that international norms develop in a linear progression. Accordingly, norms go through the process of being introduced by norm entrepreneurs, they then cascade and are institutionalized, until finally they are internalized (Saurugger, 2010). From their ‘introduction’ by norm entrepreneurs, the ideas associated with a norm are mostly regarded as fixed. Consequently, the tendency has been to focus on how norm entrepreneurs promote a norm and not on how the process of norm promotion is actually prompted and preceded by a process of ideational contestation, debate, and discussion that refines and selects the ideas that are associated with the norm’s practice, which ultimately come to define an international norm.3

More recent literature on international norms highlights changes that occur during the norm’s diffusion. These arguments focus on the process of norm ‘localization,’ in which global and local actors change the framing and content of an international norm in order to resist, adapt, or incorporate the norm (Acharya, 2004; Capie, 2008; Cardenas, 2007; Dixon, 2013; Florini, 1996). This focus on change and strategies of resistance or adaptation, while important for our understanding of compliance and non-compliance with international norms, still understates the role of ideational change as a causal mechanism in the early stages of a norm’s emergence. Since this literature focuses on norm diffusion, less is understood about processes of contestation during a norm’s emergence. This omission implies that between the norm’s initial articulation by norm entrepreneurs and its diffusion, the content of the norm remains largely unchanged. This is inaccurate and contributes to analytical misunderstandings of the determinants of norm emergence and to biased assessments of the effects of a norm across time.

I propose that understanding the changes in the ideational content of norms provides a more accurate description of international norms’ emergence.4 For example, changing ideas about slavery — from the logic of declining profitability to the logic of moral progress — took place during and facilitated the emergence of the international norm of the abolition of slavery (Ray, 1989). Similarly, changing ideas about sovereignty — from territorial-based to populations-based legitimacy — took place during and contributed to the emergence of sovereignty as an international norm (Barkin and Cronin, 1994). Furthermore, I propose that a theory that captures changes in the ideational content of norms not only provides a better description of the emergence of international norms, but also better explains how norms emerge.

Before further developing my content change proposition, it is pertinent to clarify the relationship between ideational change and agency. In explaining the emergence
of international norms, the International Relations literature to date has established the critical role of different types of norm entrepreneurs, including advocacy networks (Keck and Sikkink, 1998; Schrad, 2010), epistemic communities (Haas, 1997), and international organizations (Finnemore, 1996). The content change proposition I develop here contends that agency is a necessary yet insufficient mechanism in explaining the emergence of international norms. Norm entrepreneurs are necessary precisely because they frame, articulate, debate, and facilitate the spread of ideas. However, a complete causal explanation must also account for the content of those ideas. For the purpose of this article, I therefore bracket the issue of norm entrepreneurship and focus on the role of ideational change in the emergence of international norms.5

The content change proposition

The focus on change in the ideational content of norms is closely linked to the definition of norms. Following the most widely accepted definition, international norms are defined as 'a standard of appropriate behavior for actors (states) with a given identity’ (Finnemore and Sikkink, 1998: 891). A central component of this definition is the notion of appropriateness, which involves both a cognitive and an ethical assessment of why a specific practice is positive or good and what or for whom it is good (March and Olsen, 1998: 951). The importance of actors’ assessments in adopting a behavior and/or a policy option is grounded in the empirical and conceptual literature on the role of persuasion in norm emergence and diffusion (Kratochwil, 1989; Payne, 2001; Risse, 2000). As the argument goes, rational and moral reasoning and argumentation are both causally significant for their roles in creating new majority views (Crawford, 2002; Kornprobst, 2007). Building on these insights, I propose that change in the content of an emerging norm that makes the practice associated with the norm more persuasive or appealing is likely to facilitate the adoption of this practice as a new standard of appropriate behavior.

I propose three types of change that are likely to make a practice more persuasive and to make the ideas and practices associated with a norm congruent with more contexts and for more states. These changes are in the rational and moral reasoning and argumentation that frame the practice that is associated with the emerging norm. By themselves, these changes are neither necessary nor sufficient; yet, their cumulative effect makes the emergence of an international norm more likely.

The first two types of change I propose correspond with March and Olsen’s logics of action (1998): the logic of consequences and the logic of appropriateness. The logic of consequences refers to a rational calculation with regard to the expected returns attained by a specific course of action. The logic of appropriateness refers to the values and identity that are associated with a specific course of action:

- **Type I: Ideational change in expected returns**: broadening the range of goals — moral and practical — perceived to be attained by adopting the norm. The logic here is straightforward: the more a state or its leadership might expect to gain by adopting a particular practice the more likely it is to adopt that practice. In the same vein, the less a state or its leadership has to lose by adopting the practice the more likely it is to adopt it.
• **Type II: Ideational change in values and identity**: elevation of the normative status of a practice and expanding the range of identities it is associated with; thereby, attaching more positive moral values and desired identity to the practice and/or to the actors who adopt the practice. In the same vein, a moral degradation may be associated with the non-adoption of a practice and of those who do not adopt it when deemed appropriate.

Beyond the utilitarian and normative logics of actions, actors’ assessments are also framed by how a specific practice relates to other practices that are associated with similar or alternative goals and values:

• **Type III: Specification**: the specific practice becomes more clearly distinguished from other similar or alternative practices. While emerging norms are related to and ‘grafted’ onto a broader normative formwork (Kelley, 2008; Price, 1998), a norm’s specification and differentiation from similar or alternative practices plays a role in its emergence. This type of change relates to the proposition that the clearer and more specified a norm is, the more likely it is to be followed (Legro, 1997: 34). Accordingly, state leaders are more likely to adopt a practice when they understand what is expected from them and what they can achieve by adopting the practice.

The three types of ideational change outlined above generate important observable implications. In the first place, norms are not merely a behavioral regularity or a standardized behavior (Axelrod, 1986: 1097; Thomson, 1993). Instead, norms — by definition — add value to behavior and necessarily relate to what is considered acceptable or not acceptable by a larger community; that is, norms carry a sense of shared moral assessment. Accordingly, since norms are collective, or intersubjective, we should find evidence of shared moral assessment in the discourse or rhetoric that precedes behavior (Bjorkdahl, 2002: 13; Klotz and Lynch, 2007: 19).

Second, the first or repeated occurrence of the practice associated with a norm is an insufficient measure for an emerging or spreading international norm. For example, recorded cases of international elections observation began in 1962; yet, we do not consider this practice to have been an international norm before the early 1990s (Hyde, 2011). Therefore, explanations of when and how a norm emerged should include analysis over time of the speech and written interpretations and specifications of the goals and values associated with the behavioral practice that is prescribed by the norm. On the basis of such empirical investigation, we should be able to distinguish between a pre-norm stage, in which there are no shared assessments about the universal applicability of a practice for actors within a given identity (e.g. democracies), and a stage of norm emergence, in which there is a shared assessment about the universal applicability of a practice for actors within a given identity.

**Alternative explanation**

The main alternative to the *content change proposition* is that any identified changes in a norm’s ideational content are not an explanatory mechanism in the process of norm
emergence but are merely an indication that the norm is being adapted to various contexts. These changes may reflect a process of rational adaptation or ‘localization,’ according to which norms change because they spread as opposed to spread because they change (Acharya, 2004; Capie, 2008; Checkel, 1997; Cortell and Davis, 2000).

The timing of the change in discourse and argumentation is one way to adjudicate between the content change and the adaptation propositions. If the change in discourse is concurrent or follows the international adoption of a norm, it may be dismissed as a rational adaptation and an indication of how local actors merely reconstruct or even ‘hijack’ international norms (Subotic, 2009). But if my proposition is valid, the change in discourse would precede the growing adoption of a norm’s behavioral manifestation. Timing, however, may not be sufficient to determine the direction of causality between the content change and the spread of the norm. The critical evidence, in this case, is the ‘accretion’ of the content change. Accordingly, if the rational adaptation proposition is valid, actors’ argumentations in support of the practice would vary by case, and different arguments would be developed in different contexts even in later cases. If my content change proposition is valid, change in the argumentation is likely to be cumulative and will consistently reflect the new ideas that helped constitute the norm. These new ideas will not only frame the practice of the norm across different cases, but also be consistent with how the norm is subsequently institutionalized.

The universal appropriateness of truth and reconciliation commissions

How is it that the regionally bounded and ‘poor alternative’ truth commissions became a universally applicable and desirable mechanism? This shift in the meaning of TRCs is a good case for demonstrating and evaluating the content change proposition with its three types of change. These changes should be observable in speech and written materials and should precede the international spread and institutionalization of TRCs. One could argue that the selection of one case study makes general inferences problematic (King et al., 1994). While there are important limitations to the examination of one case, the analysis of a single case can generate valuable suggestions for refining and improving new theoretical propositions (Eckstein, 1975; George and Bennett, 2005; Van Evera, 1997). Moreover, there are three criteria that justify the selection of this case. In the first place, the international spread of TRCs provides a unique opportunity for studying a norm as it emerges. Since this case is contemporary, it provides access to rich empirical data.

Second, the international norm of TRCs emerged within the international periphery and therefore provides some control against the enduring challenge of a power-based explanation. By and large, scholars study norms emerging within and diffusing from dominant members in the international system. These norms and the process by which they spread often reflect the political and economic interests and influences of these dominant states, as well as their morality and values. Therefore, it is often impossible to control for the argument that norms spread due to the influence of power or coercion, which is the main alternative realist-based critique of the norm literature (Mearsheimer, 1994; De Nevers, 2007). This has been an inherent challenge for those arguing for the
independent role of norms in world politics. While power politics is undoubtedly part of why TRCs became an international norm, the ideas that frame this norm formed with little if any involvement of strong states in this process. Therefore, this case minimizes the potential effects of power relations and/or coercion on the emergence of this norm and leaves room to study the ideational process in relative isolation.

A third selection criterion is that the international norm of TRCs is a case of a prescriptive norm, which allows us to learn more about the role of ideas. When we study the spread of prescriptive norms, that is, norms that advocate rather than ban social or political practices (Goertz, 2003: 10; Nadelmann, 1990), we are likely to learn more about states’ positive motivations, such as the practical and moral merits associated with a norm, as opposed to negative motivations, such as the threat of sanctions. Accordingly, the emergence and spread of prescriptive norms, such as the norm of TRCs, are expected to reveal more about the role of ideational change in such processes.

Content change and truth and reconciliation commissions

The practice of truth commissions began in the early 1980s in Argentina (1983). There, the truth commission addressed the real need to find out information about the disappeared and was primarily intended as a preparatory stage for forthcoming criminal prosecutions. Due to political constraints, however, the trials did not follow and the truth commission became, de facto, the main mechanism for addressing Argentina’s legacy of human rights violations. Similar truth commissions were adopted in other Latin American transitions, such as Uruguay (1985), Chile (1990), and El Salvador (1992). In these transitions, members of the former regime remained in positions of power and were able to ward off attempts to hold them accountable for human rights violations. These truth commissions were forged as political compromises between the constraining power of the former regime and domestic and international demands for justice and accountability.

By the mid-1990s, there were no indications of an international norm in the making or that truth commissions were or should be a mechanism with universal rationale and applicability. The Latin American truth commissions were largely framed as a regional institution that was specific to the political conditions and types of human right violations in Latin America. Within the region, there was an apparent dynamic of emulation and institutional transfer, but there were no claims of universal applicability. In fact, legal scholars were concerned that this model might spread to other settings (McAdams, 1997; Maclean, 2006).

Change in the rationalization and framing of truth commissions began with the South African TRC (1995). Like the Latin American truth commissions of the late 1980s and early 1990s, the political process leading to the South African TRC ‘was a manifestation of political compromise and negotiated settlement’ (Christie, 2000: 76) and was motivated by the need to balance between the feasibility and stability of the democratic transition and demands for justice and accountability (De Lange, 2000). However, the people who conceptualized and designed the South African TRC introduced two types of unique features in the design, goals, and framing of the commission. The first type is institutional, including quasi-legal powers, scope of mandate, and transparency and publicity.
The second type of unique features is those of content; specifically the idea of assigning the truth-seeking process with the goals of reconciliation and healing. These unique features of the South African TRC along with a widespread perception of success even before its operation was completed made South Africa an ideal-type model for all subsequent TRCs (Ben-Josef Hirsch, 2009).

A key question during and following the TRC process in South Africa was whether this was a unique case or whether the principles it introduced were universally applicable. The question of the general applicability of the South African model stood at the heart of the so-called Truth vs Justice debate, in which scholars and practitioners, many of whom personally took part in the TRC in South Africa, debated the theoretical, philosophical, and practical dimensions of this question. While criticism of and doubts about the South African model persist to this day, several points of consensus had emerged among the growing community of scholars and practitioners that came to be the leaders of the transitional justice field. A practical consensus that emerged was that TRCs may achieve multiple and important psychological, political, and legal goals, including many of the goals of criminal prosecutions. A second, normative consensus that emerged was that finding truth is in itself a form of justice. This ideational shift elevated the normative status of the truth-seeking process. Finally, in light of the TRCs’ broader practicality and elevated moral value, TRCs increasingly came to be seen as distinct from trials and as a complementary rather than an alternative tool for dealing with the past.

In what follows, I trace in more detail the arguments that shaped these points of consensus and demonstrate how they correspond with the three types of ideational change I outlined earlier. I then demonstrate how these changes in the utilitarian and normative logics for and specification of TRCs facilitated their broader applicability and drove the emergence of the TRCs international norm.

Type I: Ideational change in expected returns

According to Alex Boraine, who was the Deputy-Chair of the South African TRC, ‘dealing with the past is inescapable … for the sake of justice, for stability and the restoration of dignity of victims’ (2000: 6). Indeed, these three rationales — psychological, legal-normative, and political-practical — commonly came up in the Truth vs Justice debate. In particular, they redefined and broadened the list of positive effects that are associated with the practice of TRCs.

The psychological rationale. At the individual level, the process of truth-telling was reconceived to have healing or therapeutic qualities (Goldstone, 2000: 65; Ntsebeza, 2000). In support of these arguments, proponents cited studies of trauma and torture victims and Holocaust survivors, which found that the process of producing a systematic testimony that records the torture that victims experienced helps victims channel anger as well as provide a form of catharsis (Roht-Arriaza, 1995: 19).

This psychological rationale was utilized to argue for the merits of TRCs relative to trials (Crocker, 1999; Humphrey, 2003). Alex Boraine (2000: 292) argued that in trials, victims are in a foreign setting with its own jargon and rules that are alien to them,
whereas in TRCs, victims’ hearings are designed to be embracing and supportive. According to transitional justice scholars-advocates Laurel Fletcher and Harvey Weinstein (2002), the lengthy litigation in trials may be therapeutically counterproductive for victims. In TRCs, on the other hand, victims are not cross examined and can share their feelings and emotions in a positive setting.

The psychological rationale also echoed the idea that the healing power of truth not only affects individual victims, but also fulfills the collective need for healing and catharsis (Cobban, 2002; Kritz, 2005). The ethics scholar Elizabeth Kiss (2000: 72) argued that ‘telling the truth about their wounds can heal the wounded — and perhaps listening to such stories can help heal societies.’ Advocates of this argument, like the legal scholar Martha Minow (2000: 63), described post-repression societies as ‘wounded’ and traumatized and in need of ‘recovery’ and ‘healing,’ especially as a precondition for national reconciliation (Baralou, 2005: 4). In a legal process, according to Minow (2000a: 26), ‘[the] reconstruction of relationship, seeking to heal the accused, or indeed, healing the rest of the community, are not the goals in any direct sense.’ In contrast, the process of truth-seeking, according to journalist Tina Rosenberg (1995: xviii), can help nations ‘face up to and understand past events before they can put them aside and move on to normal life.’

The legal-normative rationale. Punishment, deterrence, and the prevention of future violence are the three legal goals commonly associated with criminal prosecution. In the Truth vs Justice debate, advocates of TRCs argued that trials are not the most or the only suitable tool for attaining these goals. According to Martha Minow (2000b: 253), TRCs ‘punish by shame’ via the public exposure of crimes and can also be an effective deterrent (Sarkin, 1999). Others argued that trials reinforce group demarcation and that trial witnesses tend to ‘take sides’ and view members of their own group as victims of injustice (Fletcher and Weinstein, 2002: 592). TRCs, on the other hand, were argued to remove incentives for revenge and break the actual or potential cycle of resentment and violence. In particular, moral questions and the morality of TRCs became a focus for many advocates of TRCs, who argued that their practice of participation and reciprocity is the basis for their morality (Azzam et al., in Steiner, 1997: 17–22; Gutmann and Thompson, 2000). As stated by Jose Zalaquett, who was a Commissioner in the Chilean TRC and advised the South African TRC’s preparatory process: ‘Following a major breakdown of the rule of law and basic civic values, a society must reconstruct its moral underpinning. Truth Commissions can be a part, perhaps the cornerstone, of such a process of moral reconstruction’ (cited in Steiner, 1997: 15).

The practical-political rationale. In the Truth vs Justice debate, the list of practical and political arguments in support of TRCs was particularly long. New practical arguments focused on the efficacy and efficiency of TRCs, especially vis-a-vis trials. Human rights violations on a massive scale generally involve and implicate a large part of the population (Graybill and Lanegarn, 2004). Moreover, countries in transition often lack the institutional (e.g. judicial system) and social foundations needed to support widespread prosecutions. Trials are, therefore, likely to be too costly and impractical. Advocates argued that TRCs, on the other hand, can begin functioning relatively quickly during or
following a transition and are less expensive and time-consuming than trials. In addition, countries in transition are more likely to require the reintegration and rehabilitation of former perpetrators and collaborators into the broader society, functions which a truth commission can fulfill more successfully (Villa-Vicencio, 1998).

Beyond these practical logics, the Truth vs Justice debate set forth broader political rationales for TRCs. Jose Zalaquett argued that ‘the commission’s purpose is political in the broad sense that it helps to lay foundations for a new political system or to reconstruct a broken one’ (cited in Steiner, 1997: 28). The operation of a TRC was said to foster the democratic elements of reciprocity and deliberation (Gutmann and Thompson, 2000: 38) and could set the foundations for a stronger civil society within the new democratic state (Kritz, 2002: 61). This partnership could create a public sphere with shared commitment to norms that denounce human rights violations.

Finally, scholars argued that TRCs could produce a unifying national historical narrative, which is supposedly a crucial first step in nation-building (Christie, 2000: 117). Criminal prosecutions, with their emphasis on individual liability, are only interested in the truth that is relevant for establishing guilt and therefore fail to tell the whole story of past abuses. TRCs, on the other hand, increase the quantity and quality of information about past abuses (Greenawalt, 2000; Slye, 2000). Personal suffering is publicly acknowledged and written into the national narrative (Andrews, 2003). Moreover, the official historical narrative can become the basis for public education, which socializes the new post-repression generation by teaching the lessons of the past and promoting a culture of human rights (Sarkin, 1999).

Taken together, these new arguments about the positive political effects associated with TRCs broadened the perception of TRCs’ democratizing role. In addition to strengthening the accountability, legitimacy, and stability of a new regime in a cost-effective way, TRCs also came to be understood as helping to constitute and advance a ‘true’ democratic political culture (Rosenberg, 1999). As a result, TRCs were framed as having multiple democratizing effects, including: advancing the creation of a public sphere via a deliberative dialogue over truth; fostering a culture of human rights through a shared commitment to norms that denounce human rights violations; promoting reconciliation, which is an element of democratic coexistence; and establishing an official historical narrative that helps construct a state’s new democratic identity.

In sum, the ideational change that resulted from the Truth vs Justice debate led to the perception that TRCs have a broader range of goals and that there is more to be attained by TRCs. Moreover, one of the most significant aspects of this debate was that TRCs were positioned for the first time on equal ground with their judicial alternative. It was argued that TRCs not only achieve broader goals, but may also do better than trials in achieving these goals. According to law professor Naomi Roht-Arriaza (2006: 4), TRCs were no longer ‘a second-best alternative where trials were unavailable,’ but instead could ‘accomplish things no trial could provide.’

Type II: Ideational change in values and identity

The second type of content change focuses on change in the ‘logic of appropriateness’ of TRCs, which is positive change in their normative status and in the values and identity
that are associated with states which adopt them. Since the early truth commissions in the mid-1980s, the trade-offs between trials and truth commissions had been mainly conceptualized as a choice between justice and truth (Llewellyn, 2006). The consensus was that trials represented ‘real’ justice (Seils, 2002). In the Truth vs Justice debate, therefore, the potential of TRCs to achieve justice was avidly debated. The concept of restorative justice entered the discourse in the South African TRC report. It offered a way out of the ethical dichotomy between justice and truth and became the main conceptual and normative framework for addressing a justice-based critique (Crocker, 2000; Kiss, 2000; Teitel, 2003: 76; Van Zyl, 2000; Villa-Vicencio, 2000).

The restorative approach to justice developed in the United States domestic criminal system. It focused on the idea that a process of dialogue between victims and offenders is more suitable than trials for restoring the victim’s sense of personal autonomy and at the same time allows for the reintegration of offenders and the rebuilding of the community (Dzur, 2003; Moosa, 2000: 120). Key elements of the restorative justice approach echoed with TRC practices and it became a common normative frame of reference. The legal scholar Martha Minow (2000a: 91), for example, argued that ‘unlike punishment, which imposes penalty or injury for violation, restorative justice seeks to repair the injustice, to make up for it and to effect corrective changes in the record, in relationships, and in future behavior.’ And the moral theologian Nigel Biggar (2003: 7) noted that ‘[i]t is common to think of criminal justice primarily, even wholly, in terms of the punishment of the perpetrators … this is a mistake, because justice is primarily not about the punishment of perpetrators, but about the vindication of the victims.’

Before the South African TRC and before the Truth vs Justice debate, the understanding was that truth commissions sought ‘justice to the extent possible’ (Zalaquett, cited in Weissbrodt and Fraser, 1992: 602). This reflected the consensus in the early 1990s that truth commissions might achieve ‘less’ justice and were morally acceptable only as a necessary political compromise. The restorative justice framework changed that consensus. The healing of individuals and society, along with deliberation and reconciliation, and the incorporation of perpetrators into the newly reconstructed society were not only distinct goals of the truth-seeking process, but also conceptualized as elements of justice — a restorative justice. Accordingly, it was argued that ‘the truth or more precisely the process of truth-seeking, truth-telling, and truth-publishing is in itself a form of justice’ (Godwin Phelps, 2004). In this discourse, the TRC process came to embody a restorative justice that is morally equal, and in some aspects even superior, to a retributive justice (Leebaw, 2003). The TRC, therefore, has become a morally desirable practice and not a moral compromise.

**Type III: Specification**

The third type of content change is in how a specific practice relates to other practices that are associated with similar or alternative goals and values. As noted above, a key aspect of the Truth vs Justice debate was the changing assessment of the utility and morality of TRCs relative to trials. The Truth vs Justice debate did not place TRCs and trials in all-out competition, but rather framed the two mechanisms as having different relative merits (Leebaw, 2011). In this assessment, there was not only an emerging
consensus, but also strong opposition. Significant opposition came from human rights activists and scholars, who argued that human rights violations necessitate a retributive approach and trials (Orentlicher, 1991). Many legal scholars and human rights activists reaffirmed and endorsed the ‘duty to prosecute,’ especially in the context of discussions during the early to mid-1990s over international tribunals and the International Criminal Court (Bass, 2002).9

This opposition and subsequent dialogue led many scholars and practitioners to argue for the ‘fallacy of the dichotomy of truth versus justice’ (Lutz, 2006: 326). As the argument goes, since truth and justice are morally equal, and since TRCs and trials effectively fulfill multiple different goals, then both are necessary components of transitional justice. This so-called complementarity principle focuses not on whether to have TRCs or trials, but on how to combine and sequence them (Fletcher and Weinstein, 2002; OHCHR, 2006: 5; Roht-Arriaza, 2006; Villa-Vicencio and Doxtader, 2003). What we observe is not merely a division of labor (Robinson, 2003); instead, the complementarity principle accepts the distinct political and moral rationales — that is, the specification — of trials and TRCs.

Sierra Leone offers a good example of the emerging complementarity consensus in practice. The Sierra Leonean TRC was part of the negotiated peace agreement (1999) between the government of Sierra Leone and the Revolutionary United Front of Sierra Leone. Initially, the initiative for a TRC was met with criticism from many legal experts, and when in 2002 the United Nations and the government of Sierra Leone established the Special Court for Sierra Leone, it was assumed by some that a TRC was no longer needed. However, the decision in Sierra Leone was to have both a TRC and trials. This decision confirmed the distinct goals of TRCs and trials. While the Special Court emphasized justice through punishment, the mandate of the TRC was to promote reconciliation and healing à la South Africa. As the Chair of the Commission, Joseph C. Humper, noted, the Special Court and the TRC are ‘going to the promised land, but by different roads’ (cited in Schabas, 2003: 1038). Notably, the TRC was not seen as a ‘second fiddle,’ since a public and conscious decision was made by the United Nations Office of Legal Affairs that neither of these institutions would receive primacy over the other (Kritz, 2002: 69).10

All in all, the complementarity consensus reaffirmed the distinctiveness of TRCs from trials by specifying their distinct functionality and expected outcomes. It demonstrated that TRCs had come to occupy a specific ‘niche’ in the set of mechanisms for dealing with the past.

Alternative explanation

As noted above, the main alternative proposition to the content change proposition is that the observed changes in the ideational content of the TRC norm are merely an indication of the rational adaptation of TRCs to various contexts. Rational adaptation to local conditions was undeniably a part of this process. Since the early truth commissions in the mid-1980s, designers of subsequent TRCs learned from and refined the model based on previous experiences. For the most part, however, the changes in discourse that I have identified precede the growing adoption of TRCs, they accrue across time, and do not randomly vary across cases.
As shown in Figure 1, the growing proliferation of TRCs followed the South African TRC. As I have demonstrated, the content changes also followed the South African TRC, which initiated changes in the discourse about the goals, values, and specific functions of TRCs. In the design process and the mandate of subsequent TRCs, there are often direct references to the South African precedent, as well as to the new framing it initiated.

Beyond the demonstrative effect of the South African TRC, it took some time for the new ideas associated with the TRC model to take over. For example, the truth commissions in Guatemala (1994), Burundi (1995), Ecuador (1996), and Nigeria (1999), which took shape during the operation of the South African TRC, included no reconciliation clause and did not attempt to incorporate any of the other unique features of the South African TRC. In fact, their framing and designs shared more with earlier Latin American commissions. From 2000 onward, however, out of 32 commissions initiated, all but one case (Lebanon (2000)) of initiated TRCs reflects the new ideas about the TRC model in their goals and design.

The growing applicability of truth and reconciliation commissions

In accordance with my proposition, the Truth vs Justice debate that followed the South African TRC yielded changes in the ideas associated with the goals, values, and the specification of the TRC model. Together, these changes made TRCs congruent with more settings and appealing to more states. This outcome is demonstrated in the growing consideration and application of TRCs in at least three new contexts. The first one is the context of conflict resolution (Borer, 2006; Hayner, 2000).
Beyond reconciliation, the conflict resolution literature highlights that TRCs provide acknowledgment of and vindication for victims and, therefore, may alleviate underlying feelings of resentment and the desire for revenge (Lund, 1998). TRCs can also remove the ‘contested past’ that often is in itself an important underlying cause of conflict (Bassiouni, 1996: 23). Moreover, a TRC can become part of the negotiation game and may signal positive intentions as well as facilitate and legitimize a compromise (Hayner, 2000: 361).

All in all, these arguments draw on the rationalizations for TRCs in a transitional context and apply them to a conflict resolution context. This application demonstrates the broadening of the goals of the truth-seeking process. While early truth commissions in Latin America were designed for revealing the ‘hidden truth’ of state repression, the current framework since the early 2000s also emphasizes ‘contested truth’ and finds TRCs to be appropriate for a wider range of settings. In practice, there are several examples that demonstrate the emphasis on reconciliation and the incorporation of truth-seeking into the trade-offs in peace negotiations, including: the Sierra Leone TRC (2000); the Burundi TRC (negotiated in 2000 and established in 2005); the TRC in the Democratic Republic of Congo (negotiated in 2002 and established in 2006); and the Liberian TRC (2004). Beyond intra-state conflicts, similar logics have been used to argue for TRCs in the context of international conflicts as a way to advance inter-state or even regional reconciliation (in Israel–Palestine (see Dudai, 2007) and in former Yugoslavia (see Initiative for REKOM, 2011)).

The second context in which we observe the applicability of TRCs to settings other than political transitions is to cases of gross human rights violations (e.g. genocide), for which trials have been an imperative and for which TRCs were deemed inappropriate in the past. The TRC initiatives for Bosnia and Herzegovina provide a compelling illustration. When the conflict in the former Yugoslavia ended in the mid-1990s, the option of a TRC was not considered and the International Criminal Tribunal for the Former Yugoslavia (ICTY) was established. This choice reflects the consensus at the time that cases of gross human rights violations called for prosecution (Neier, 1998). Highlighting this fact, a United States Institute of Peace (USIP) initiative for a truth commission in 1997 was forcefully rejected and deemed unnecessary or even harmful by senior officials of the Hague tribunal and human rights activists (Louise Arbour, ICTY Chief Prosecutor, cited in Hayner, 2001: 207; Freeman, 2004; Kritz, 2002; Neier, 1999: 42). By 2000, however, when a revised proposal for a joint Bosniak–Croat–Serb TRC was presented by USIP and a grassroots initiative called the National Coordinating Committee for Establishment of the Truth and Reconciliation Commission, reactions were different (Mertus and Sajjad, 2005). The proposal cited the South African model and emphasized the goals of ethnic reconciliation and establishing a common historical narrative. With these goals in mind, the former ICTY President Claude Jorda, for example, endorsed the TRC as ‘complementary [to] and distinct’ from the ICTY (Jorda, 2001). The TRC proposal was drafted into legislation in Bosnia several months later and was re-endorsed by local party leaders in 2005. As this example demonstrates, a TRC became a viable option when it was no longer perceived as an alternative to trials or as an impediment to justice. In contrast, local leaders considered and accepted this option when it carried the promise of achieving goals...
that the ICTY was not designed to achieve, including long-term reconciliation and nation-building. In other words, the content change traced above re-shaped the perceived viability of the TRC option in this case.

Finally, a third context in which we find the growing applicability of TRCs is in mature democracies seeking to account for historical injustice. While historical truth commissions in mature democracies have precedents (Hayner, 2001: 17–19, 298–299), the more recent ones, like the TRC of Canada (2008), draw on the broader rationales for a TRC that followed the South African TRC; specifically, on the nexus between a truth-seeking, public acknowledgment process and reconciliation.

Conclusion

In the last decade, TRCs have become a permanent feature of transitional and post-conflict settings. This article traced the emergence of the TRCs norm to the South African TRC and the debates that followed it, in which the framing of the rationale for TRCs has changed significantly. In the new framing, TRCs came to be associated with multiple political, psychological, and moral goals and they came to be understood as a tool that strengthens democratization. The framing of the truth-seeking process’s relative merits has shifted from emphasizing what TRCs can achieve in comparison with trials, to emphasizing what TRCs can achieve that is different from and better than trials. In accordance with my proposition, the three types of ideational change, namely, changes in the expected returns, in values and identity, and the specification of TRCs, made it more likely that TRCs would be adopted by a growing number of states and recommended by a growing number of international actors and organizations.

My analysis carries at least three important implications for the study of the emergence and effects of international norms. The first implication of my content change proposition is that norm emergence is not a linear goal-oriented process predetermined by norm entrepreneurs who set out to create an international norm. Rather, norm emergence is a dynamic process in which ideas, practices, and consensus change significantly. Our ability, therefore, to predict which international norms will emerge and which will not survive is limited. At most, the study of international norms should focus on uncovering changes as they occur, thereby offering a better understanding of the actual processes through which the ideational and normative international environment is shaped.11

The second implication focuses on changes during different stages of the norm life-cycle. The content change proposition identifies ideational change as a causal mechanism that shapes norm emergence. The recent norm localization literature understands change as a state strategy — or an intervening variable — during norm diffusion. In practice, these two conceptualizations of change may complement each other. Overall, they suggest that ideational change is an enduring feature of the norm life-cycle, which contrasts with the prevailing view of the content of international norms as relatively static.

The third implication pertains to empirical research on the effects of norms. As my analysis of the TRC norm demonstrates, earlier applications of truth commissions operated within different international normative environments, under different rationales, and with different goals from later ones. Therefore, not differentiating between early and
late cases of TRCs is likely to yield biases in the evaluation of their effects. The bias of aggregating early and late cases of truth commissions is mostly apparent in large-n studies of TRCs. Sikkink (2011), for example, includes in her ‘justice cascade’ data set all truth commissions since 1974. She analyzes the correlation between truth commissions and the recurrence of human rights violations in order to determine the deterrence effect of truth commissions. While deterrence might have been an important goal of early truth commissions, when they were conceived as a weaker alternative to criminal trials, in the later truth commissions, deterrence was not a central goal.12

I propose that in order to avoid these biases, researchers interested in the effects of an international norm must contextualize their dependant variables — that is, the expected effects of the practice associated with the norm — in accordance with changes in the goals and values that are associated with the norm. At a minimum, large-n studies, which are particularly prone to the bias of aggregating early and late cases of a norm’s application, should include a time-series analysis that captures the changes in the norm’s content across time.

As a final point, my proposition calls for new directions in the international norms research agenda. In theory, rigid ideational frameworks that do not allow for change are likely to yield weaker international norms. On the other hand, the more a norm is flexible and adaptable to change, the more it is likely to be congruent with more settings, and hence likely to be stronger. In addition, and somewhat counter-intuitively, the less initial consensus there is about a practice and the more it is subject to debate, the greater the likelihood is that a norm that emerges out of these debates will be stronger. Addressing these propositions will require developing empirical measures for how flexible or rigid ideational frameworks are and what constitutes weak and strong norms. This is a task worth pursuing as it will allow us to make better predictions about the longevity of international norms and to better understand the structure and content of the international normative environment.

Acknowledgements

I would like to thank the participants of the 2011/2012 Conflict, Security and Public Policy Seminar, Belfer Center for Science and International Affairs, Kennedy School of Government, Harvard University. Special thanks go to Sarah Bush, Ehud Eiran, Kelly M. Greenhill, Melissa Willard-Foster, and mostly to Jennifer M. Dixon whose questions, comments, and suggestions have contributed to this final version. I also thank two {EJIR} reviewers for their constructive suggestions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1 The ‘universe’ of initiated truth commissions between 1974 and 2008 include 52 cases, of which 29 were initiated in the last decade (Ben-Josef Hirsch, 2009).
2 Specifically, this static conceptualization creates a validity problem in the coding of variables in large-n studies. For example, in many longitudinal studies, the coding of democracy ignores that since 1815 democratic values and practice have changed over time (Oren, 1995).
Recent exceptions are studies on agenda-setting within advocacy networks (Carpenter, 2011). But even these studies focus on political competition within and between norm entrepreneurs, not on how the ideas themselves play into the process of agenda-setting.

While some constructivist scholars, especially those identified as ‘thick constructivists,’ may object to the distinction between ideas and norms, this distinction is both analytically and methodologically justified for the purpose of describing and explaining the causal role of ideational change (Barkin, 2010: 148–153; Klotz and Lynch, 2007).

The distinction between agents and ideas is evidently for analytical purposes. In practice, it is agents — in this case, mostly practitioners and scholars — who initiate, debate, and spread the new content, which facilitates the emergence of international norms. I explore the role of agency elsewhere (Ben-Josef Hirsch, 2007, 2009).

The Truth vs Justice debate took place in roundtable discussions and academic conferences as well as over the pages of numerous articles. The Truth vs Justice title is that of an edited volume by Robert I. Rotberg and Dennis Thompson (2000). This book originated in a May 1996 roundtable discussion sponsored by the World Peace Foundation and the Human Rights Program at the Harvard Law School. Two years later (May 1998) it was followed with a conference in South Africa and with the subsequent publication of the edited volume. Additional conferences were held and also resulted in edited volumes relevant to this debate (Barahona de Brito et al., 2001; Biggar, 2003; Villa-Vicencio and Verwoerd, 2000).

These rationales follow Rama Mani’s (2002) typology of the goals of truth commissions.

Not everyone in this debate agreed that there is a direct causal link between truth-seeking and democratization. But even critics argued that truth-seeking can set in motion a dynamic for wider democratic reforms (Barahona de Brito et al., 2001: 32).

This opposition highlights that the consensus described here is by no means hegemonic and was, and still is, contested. Nevertheless, it captures the framing by leading TRCs scholars and practitioners as well as the framing of the international institutionalization of TRCs.

While in principle it was decided in Sierra Leone that neither the TRC nor the Special Court receives primacy over the other institution, in the competition over funding, trials were markedly privileged (Dougherty, 2004).

Krook and True’s (2012) study of gender equality norms illustrates the research agenda I promote here.

Small-n studies can also err by applying post-hoc analysis which seeks to evaluate effects that were not even conceptualized when the truth commission operated. Joanna R. Quinn (2004), for example, uses the restorative justice framework to evaluate the impact of the 1986 Ugandan Commission of Inquiry. Given that the concept of restorative justice was introduced to the logic of truth commissions only in the report of the South African TRC more than a decade after the Uganda truth commission, it is of little surprise that Quinn finds no lasting impact.

References


**Author biography**

**Michal Ben-Josef Hirsch** is a Research Fellow with the International Security Program, Belfer Center for Science and International Affairs at the Kennedy School of Government, Harvard University, USA. In 2009–2012, she was a post-doctoral Fellow at the Schusterman Center for Israel Studies, Brandeis University and a Lecturer at the Politics Department and the Coexistence and Conflict Program, Heller School for Social Policy, Brandeis University. Her research interests include International Relations theory with a focus on the role of ideas and norms in world politics, historical justice and memory, and transitional justice.